

含油气区域自动提取的多剖面分析方法¹⁾

周冠雄 胡肃非 邹海明

(华中工学院)

提 要

作者在文献[1]中给出了一个用于地震勘探剖面图象岩性分析的模式识别专家系统,它包含低级处理及高级处理系统。本文给出一个岩性变化的多剖面自动分析方法。它是低级处理系统的内容,为高级处理系统提供剖面图岩性变化区域的分类描述,亦可单独运行,输出垂直剖面上的岩性分布,直接提取含油气区域。应用模式识别技术判断含油气区域,目前的方法都是在单一种剖面图上实现的,其效果离实用尚有很大距离。本文的方法综合利用多种剖面图象信息,在理论和实用上有更高价值。在 CYBER 系统上应用该方法处理实际测线数据的结果,输出了岩性分布图,其中的含油气区域与人工解释及当前出油井位的情况非常吻合。

关键词——剖面图,模式识别,特征抽取,含油气区域,聚类。

一、前 言

地震勘探资料的解释在确定含油气区域及钻井位置上有着决定性的意义,而剖面图岩性信息对于取得正确的解释结论又是最为关键的。目前,在地震勘探中应用模式识别与人工智能技术已成为该领域技术发展的一个主要方向之一,其研究工作大体上是沿着两个方向进行的:一是采用专家系统方法^{[2],[3]};二是采用模式识别方法^[4-7]。前者是基于知识的方法,后者是基于特征的方法。作者在文献[1]中介绍了一个基于特征和知识的模式识别专家系统的研究工作,它综合应用模式识别与人工智能技术,既能处理非量化的描述性知识,又能处理量化的数字图象,因而更适用于解决石油地震剖面图自动解释这一复杂问题。本文给出的岩性变化多剖面图综合分析方法,是该系统的低级处理子系统的内容。它为高级处理子系统提供剖面图象的岩性变化区域的分类描述,亦可单独运行,实现垂直剖面上的岩性分析,直接提取可能的含油气区域。该方法已在 CYBER 处理中心用 FORTRAN 语言实现。处理某油田实际测线的结果,输出了相应的岩性分布剖面图,其中所标示的可能含油气区域与目前地质专家们人工解释的实际结果及已有钻井位置非常吻合。

本文于1987年4月1日收到。

1) 国家自然科学基金资助项目。

二、剖面图相关分析

设计多剖面岩性自动分析系统的一个重要问题是如何选择输入剖面。从解释理论及模式识别的角度来看,应选择那些与岩性关系密切且相互间相关性小的剖面图。这样既可利用更多的信息,又可避免系统运行时需要面对过于庞大的计算量。

对目前用于人工解释的八种剖面图的产生机理及其提取的实际作业过程进行的分析表明,在瞬时速度、波阻抗、反射系数、瞬时振幅、瞬时相位、瞬时频率、低频吸收系数及高频吸收系数八种剖面图中,瞬时速度、瞬时频率与吸收系数剖面是最理想的输入剖面。

首先,瞬时速度、波阻抗与反射系数在理论上具有直接的推演关系。一般认为,地震道是地震子波与界面反射系数相褶积的结果,即

$$x_n = b_n * r_n. \quad (1)$$

其中, x_n 为地震记录, b_n 为子波序列, r_n 为反射系数序列。为了求取反射系数,须先求取一子波序列 a_n , 并有关系式

$$r_n = a_n * x_n. \quad (2)$$

而根据反射系数序列 r_n 及地表的波阻抗 z_0 , 就可求出地表下各层的波阻抗

$$z_{i+1} = z_0 \prod_{n=0}^i \frac{1 + r_n}{1 - r_n}. \quad (3)$$

地层的波阻抗与瞬时速度之间则存在关系式

$$z = \rho v. \quad (4)$$

其中, ρ 为地层密度, v 为瞬时速度。

式(1)–(4)所给出的关系式正好与 cyber 系统从野外作业数据提取瞬时速度剖面的信息流程相一致。因此,对瞬时速度、波阻抗及反射系数三种剖面而言,最基本的是瞬时速度剖面图。

现在,再来分析瞬时振幅与吸收系数剖面。吸收系数反映了地震波在传播过程中的衰减现象。不同岩性的地层对地震波的吸收程度不同。在均匀非完全弹性介质中,地震波的瞬时振幅随距离的增加呈指数衰减,并有关系式

$$A = A_0 e^{-\alpha r}. \quad (5)$$

其中, A 为经过吸收衰减后的振幅, A_0 为地震波的初始振幅, r 为波的传播距离, α 为吸收系数。

吸收系数 s 表述了平均振幅曲线相对于时间的梯度关系

$$ds_i = \frac{2 \times 10^6}{L_p} \log A_{i+1} \quad (6)$$

其中, L_p 为计算振幅的时窗长度, A_i 及 A_{i+1} 分别为第 i 层及第 $i+1$ 层的平均振幅。

式(6)表明,瞬时振幅与吸收系数之间有直接的推演关系,因而,吸收系数剖面带有瞬时振幅剖面的岩性信息。另外,瞬时相位与瞬时频率之间当然存在着更为直接的关系。因而,作为岩性分析最基本的信息源,应是速度、吸收系数和频率。

应用相关系数

$$\rho_{xy} = \frac{E\{(x - E\{x\})(y - E\{y\})\}}{\sqrt{D\{x\}}\sqrt{D\{y\}}} \quad (7)$$

在 50 道 × 500 毫秒的时窗上对上述四种剖面图的相关性进行分析, 得到低频吸收系数与高频吸收系数的相关系数为 0.813, 说明其相关性很大; 瞬时速度与低频吸收系数及高频吸收系数之间的相关系数分别为 -0.006 及 -0.005; 瞬时频率与瞬时速度、低频吸收系数及高频吸收系数的相关系数分别为 0.120、-0.028 及 -0.029。因而决定采取瞬时速度剖面 and 低频吸收系数及瞬时频率作为岩性鉴别的基本依据。下节介绍使用瞬时速度及低频吸收系数两个剖面图的岩性分析方法, 它不难推广到同时使用三个剖面图的情况。

三、含油气砂岩区域的自动提取

分析系统框图如图 1 所示。

3.1 地震道基元分割及基元特征

抽取与正规化

在瞬时速度剖面上分割地震道波形, 取两个过零点之间的波峰段为地震道基元, 这种分割方法体现了依据同相轴划分地层的思想, 且正好将地震道划分为不同速度的层, 正向波峰段为高速层, 负向波峰段为低速层, 每层的速度由均值给出。对分割后的每一基元, 在瞬时速度剖面上提取瞬时速度 v , 在低频吸收剖面上提取低频吸收系数 s , 作为基元的两个特征分量(若同时亦输入瞬时频率剖面则增设相应的瞬时频率为第三特征分量)并组成基元特征向量

$$\mathbf{c} = (c_1, c_2)^T = (v, s)^T = \left(\frac{1}{n} \sum_i v_i, \frac{1}{n} \sum_i s_i \right)^T \quad (8)$$

其中, i 为基元中地震道上之采样点序号, n 为基元中所含采样点数目。

然后, 对特征分量 c_1 及 c_2 分别实施正规化处理, 令

$$c'_{ki} = \frac{c_{ki} - E\{c_{ki}\}}{D\{c_{ki}\}}$$

$$E\{c_{ki}\} = \frac{1}{m} \sum_{j=1}^m c_{kj} \quad (9)$$

$$D\{c_{ki}\} = \sqrt{\frac{1}{m} \sum_{j=1}^m (c_{kj} - E\{c_{ki}\})^2}$$

$k = 1, 2; j = 1, 2, \dots, m$ (m 为基元数目, j 为基元序号)

得正规化特征向量(仍记为 \mathbf{c})。实施这种正规化处理的目的是为了消除由于各种原因所导致的不正常的特征分量对后续聚类处理所带来的影响。

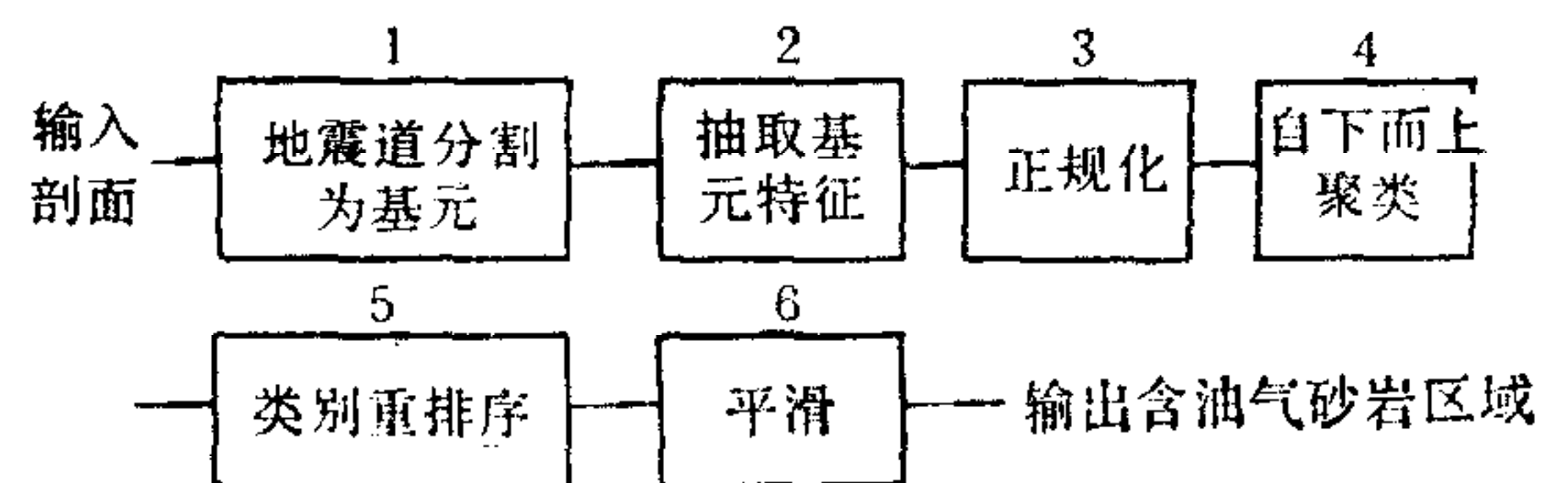


图 1

3.2 基元聚类以提取剖面图中不同岩性的地层

由于无法预知剖面上不同岩性的地层数目,因而是一个未知类别数的聚类问题。聚类从基元出发实行自下而上的合并。首先规定类别数上限 h 及下限 l , 然后将基元分成 h 个类别, 将最近邻距离的两类合并, 并用 k -平均算法计算新的类心, 再合并最近邻的两类……直至将基元合并为 l 类。在此过程中, 应用 PFS 准则决定最佳类别数, 即对 $k = h, h-1, \dots, l$, 均计算其 PFS 值, 取最大 PFS 值为最佳类别数。PFS 值的计算公式为

$$\text{PFS} = \frac{t_r S_B(m-k)}{t_r S_W(k-1)} \quad (10)$$

$$S_B = \sum_{i=1}^k n_i (\mathbf{g}_i - \mathbf{g})(\mathbf{g}_i - \mathbf{g})^T$$

$$S_W = \sum_{i=1}^k S_i$$

$$S_i = \sum_{\mathbf{c} \in \omega_i} (\mathbf{c} - \mathbf{g}_i)(\mathbf{c} - \mathbf{g}_i)^T$$

$$\mathbf{g} = \frac{1}{m} \sum_{\omega} \mathbf{c} = \frac{1}{m} \sum_{i=1}^k n_i \mathbf{g}_i$$

$$\mathbf{g}_i = \frac{1}{n_i} \sum_{\omega_i} \mathbf{c}$$

其中, m 为基元数目, k 为类别数, ω 为总样本集, ω_i 为第 i 类样本集, n_i 为第 i 类中元素的个数, \mathbf{g}_i 为第 i 类均值向量, \mathbf{g} 为总均值向量, \mathbf{c} 为特征向量, S_i 为第 i 类散度矩阵, S_W 为类内散度矩阵, S_B 为类间散度矩阵。

聚类时, 初始参数 h, l 可视分类样点的实际情况而定, 且 $h-l$ 应尽可能小, 建议取 u 为被处理图象区域的直观可视层位数; 初始类心亦可取某一基元特征向量; 距离则可根据聚类策略选用一般欧氏距离或加权欧氏距离。

3.3 类号重排序技术与岩性区域剖面图显示

基元聚类后, 由于类别数目较多, 剖面图被分割为许多小块区域, 每一小块均由类别号及类中心标明其特性。为了便于地震解释专家使用, 使之符合其它解释资料的处理习惯, 岩性区域分割的结果仍应以剖面图形式直观地加以显示。

因此, 提出了一种类别号重排技术对聚类结果实施后处理, 使类别号与岩性相联系而具有相应的物理意义, 并使性质相近的类具有相邻的类别号, 然后按类别号大小作剖面灰度显示或伪彩色显示。

因为岩性分析的目的在于确定含油气砂岩层及其边界, 故以含油气砂岩层的提取为出发点实施类别号重排技术。本系统可视实际情况采用三种重排技术。

(1) P 方式(概率方式)

假定在某一地质区域中, 含油气砂岩层的速度 v 和低频吸收系数均服从正态分布

$$v \sim N(u_1, \sigma_1) \quad s \sim N(u_2, \sigma_2)$$

由正态分布性质可知成立如下关系式

$$\begin{aligned} P(u - \sigma < x \leq u) &= 0.3413, P(u - 2\sigma < x \leq u - \sigma) = 0.1359, \\ P(u - 3\sigma < x \leq u - 2\sigma) &= 0.0215, P(x \leq u - 3\sigma) = 0.0013 \\ P(u \leq x < u + \sigma) &= 0.3413 P(u + \sigma \leq x < u + 2\sigma) = 0.1359, \\ P(u + 2\sigma \leq x < u + 3\sigma) &= 0.0215, P(u + 3\sigma \leq x) = 0.0013. \end{aligned} \quad (11)$$

将上述概率转换为另一形式,分别建立起瞬时速度 v 及吸收系数 s 与含油气概率 P_1 及 P_2 之间的联系。假定当 $u_1 - \sigma_1 < v \leq u_1$ 及 $u_2 - \sigma_2 < s \leq u_2$ 时,其含油气概率分别为 α_1 及 α_2 , 记 $\beta_1 = \alpha_1/0.3413$, $\beta_2 = \alpha_2/0.3413$, 则瞬时速度 v 及吸收系数 s 与含油气概率 P_1 及 P_2 之间的关系如表 1 所示。现设有某类别 ω_j , 其中心为 $c(v_j, s_j)$, 查表 1 可求得 P_1 及 P_2 , 则可计算 ω_j 类的含油气概率为

$$P_j = P_1 + (1 - P_1)P_2 \quad (12)$$

然后按 P_j 之大小实现类别编号。这里的待定参数 α_i, u_i 及 $\sigma_i (i = 1, 2)$ 可借助于剖面图的若干窗口的数据分析确定。

表 1

$v(i=1) \quad s(i=2)$	$P(v)P(s)$	$P_1(i=1) \quad P_2(i=2)$
$(u_i - \sigma_i, u_i]$	0.3413	α_i
$(u_i - 2\sigma_i, u_i - \sigma_i]$	0.1359	$\beta_i \times 0.1359$
$(u_i - 3\sigma_i, u_i - 2\sigma_i]$	0.0215	$\beta_i \times 0.0215$
$(-\infty, u_i - 3\sigma_i]$	0.0013	$\beta_i \times 0.0013$
$(u_i, u_i + \sigma_i]$	0.3413	α_i
$(u_i + \sigma_i, u_i + 2\sigma_i]$	0.1359	$\beta_i \times 0.1359$
$(u_i + 2\sigma_i, u_i + 3\sigma_i]$	0.0215	$\beta_i \times 0.0215$
$(u_i + 3\sigma_i, +\infty)$	0.0013	$\beta_i \times 0.0013$

(2) E 方式(专家方式)

由有经验的地质专家(或专家组)根据具体处理地域的地质背景直接给出 P 方式中之概率表 1, 然后按式(12)提供的概率对类别编号。

(3) W 方式(权重方式)

在本文所研究的地域中,含油气区域是低速、强吸收的砂岩体。根据这一特性,利用速度与吸收系数建立一个直接权重函数

$$W(\omega) = (-v) \times s \quad (13)$$

其中, v 及 s 分别为类心速度分量及吸收系数分量。根据 $W(\omega)$ 的大小实现类别号排序。

经过类别号重排后处理的小区域,再运用多数表决方式进行平滑后,即可以剖面图形式实现灰度显示或伪彩色显示,为解释专家提供处理地域的可能含油气区域剖面图。

四、应用结果及说明

本系统在 CYBER 机上用 FORTRAN 高级语言实现并对实际测线进行了处理。在

此测线上已钻一测试井,证明该区域含有油气,但尚不清楚含油区域的可能边界。系统在 120 道 \times 0.6 秒的图象窗口上运行,输入瞬时速度剖面(图(2))及低频吸收系数剖面(图

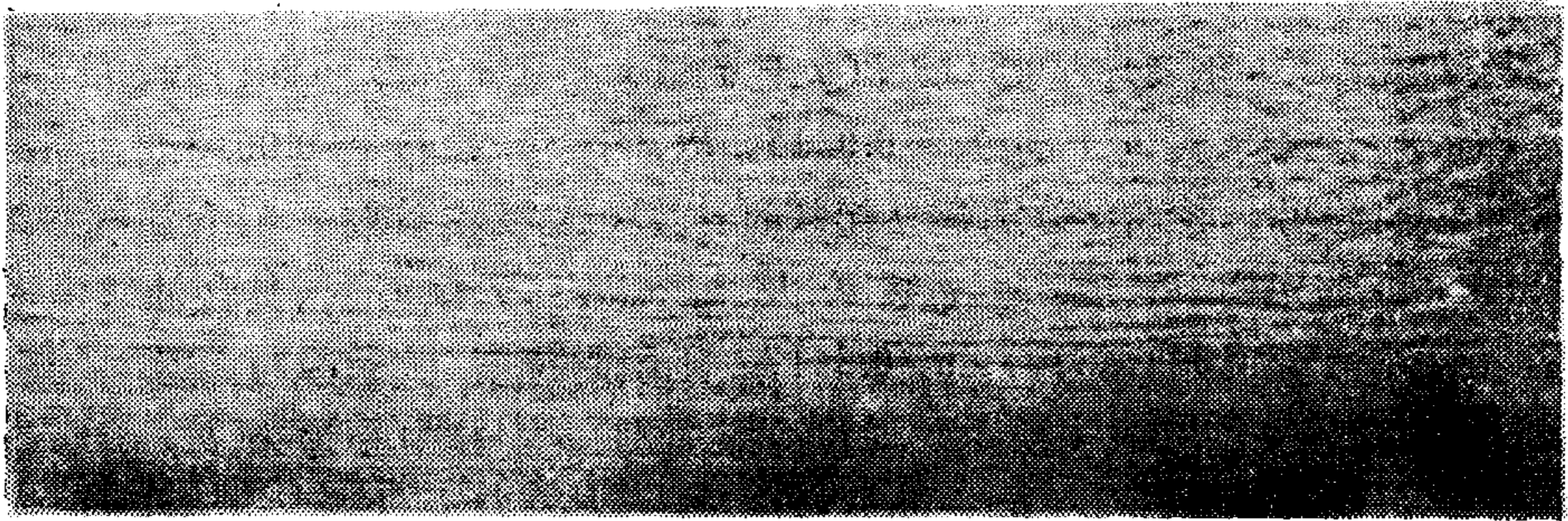


图 2 瞬时速度剖面(输入原始图)

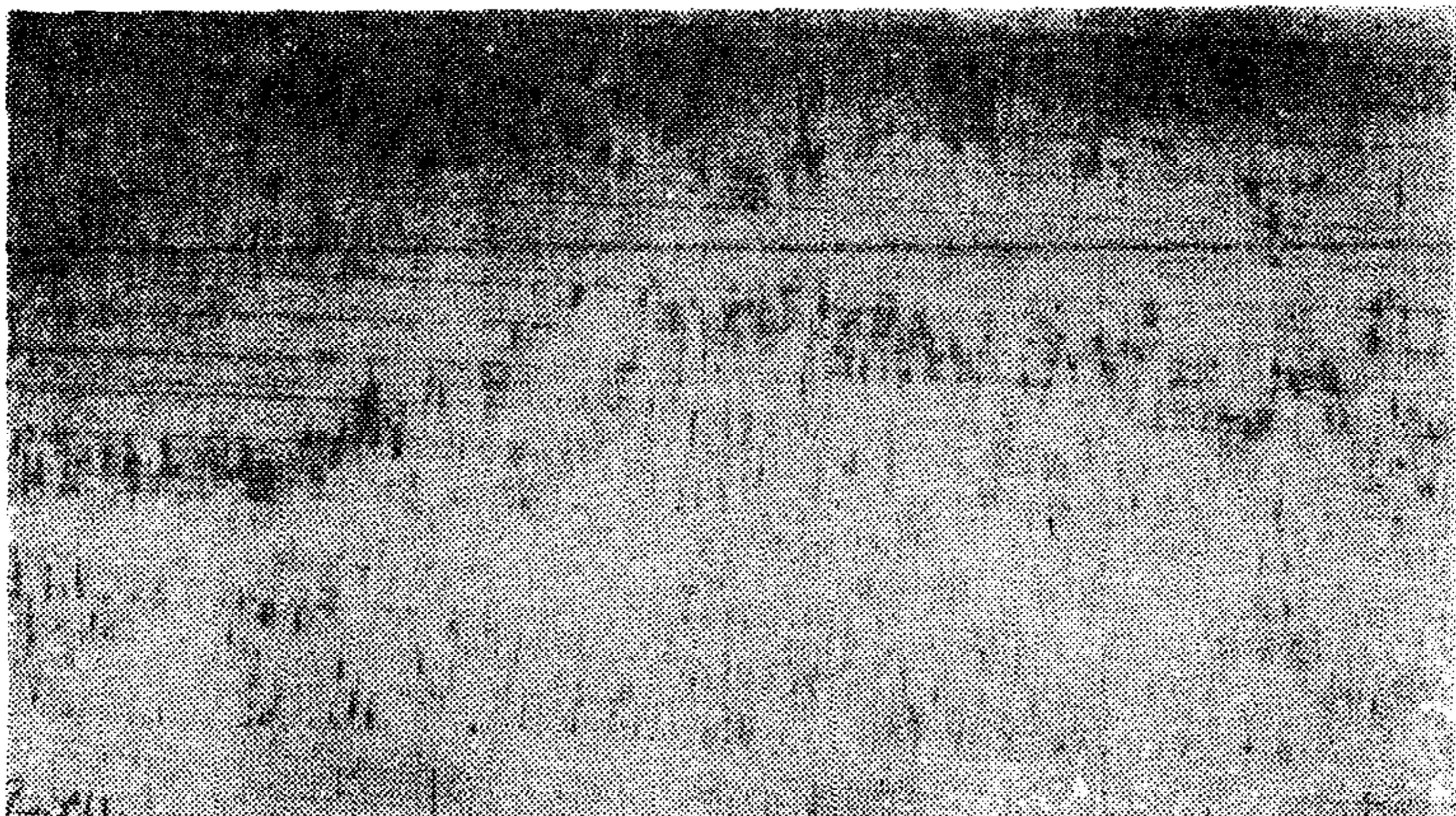


图 3 低频吸收剖面(输入原始图)

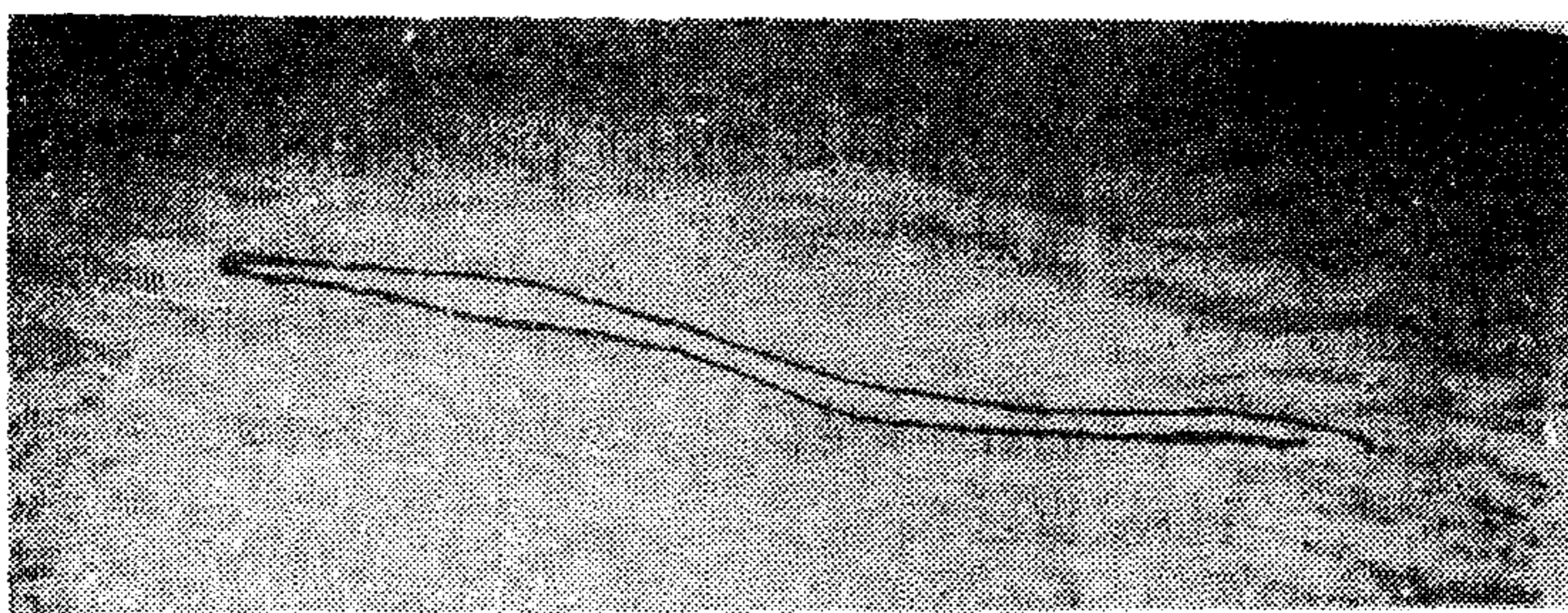


图 4 处理结果宽行打印灰度图

(3)). 瞬时速度剖面用于分割地震道,相应的速度则作为特征 c_1 . 低频吸收系数剖面中相应分割段的吸收平均值作为特征 c_2 . 在 160 道 \times 0.6 秒的窗口上处理的结果,输出的可能含油气区域如图(4)中粗线条所围成的亮带所示. 这一结果与目前人工解释及钻井证实的结果完全吻合. 显然,人们在原始剖面图(2)及图(3)中,是无法清楚地看出可能的岩性层位变化及含油气区域的,而在结果图(4)中,不仅清晰地观察到已有井位的一个含油气区域,而且还显示出了所处理地域的各层位之间的岩性变化. 图(5)是应用W方式类别号重排技术实施处理后,在 RAGOU 上显示的结果,其中最黑的一条图带为含油区.

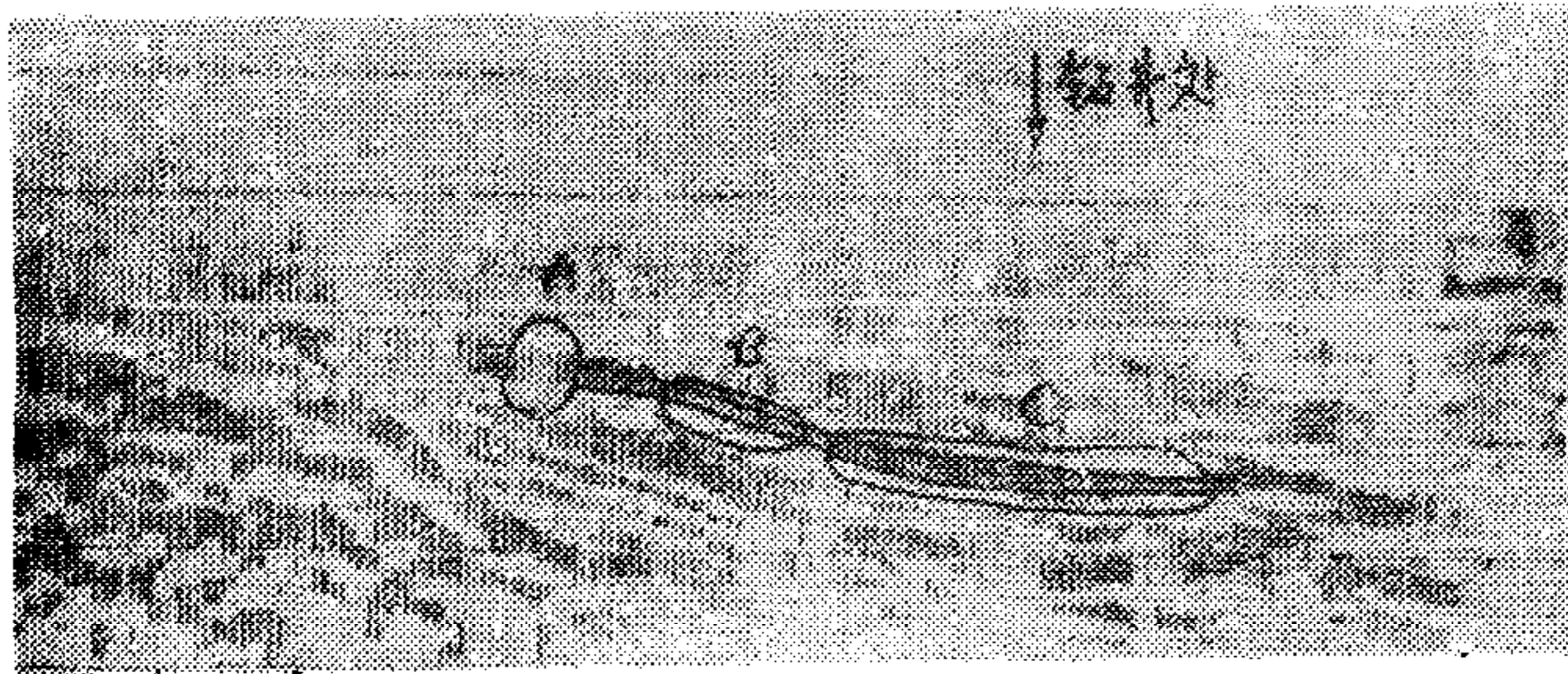


图5 处理结果 RAGOU 图

C 段最黑,含油气可能性最大,且与当前实际出油井位吻合; A 段和 B 段围在同一同向轴内,但其岩性已有所变化,有可能是砂岩层的孔隙度变小了。这一处理结果对判断边界及确定井位很有帮助。在 C 段右边,含油气区延伸不远,可判断其右边界。

聚类参数取 $h = 35$, $l = 10$, 运行过程中 PFS 值曲线变化平缓,说明岩性变化情况复杂,类别数目并不是很固定的,这与实际情况是吻合的,因为同一岩性的岩层性质也会缓慢变化。在该参数区间内最优聚类数为 20, 具有较好的低速、强吸收特性的类别多数落在图(4)的含油区域之中。

初始聚类中心的选择必须保证初始中心通过各种不同岩性的地层,且能加速聚类并提高聚类结果的准确性,为此,取初始类心为 $\mathbf{c}(i-1)s+1$, $i=1,2,\dots,k$, s 为 n/k 之最大整数, \mathbf{c} 为样点特征向量, n 为样本点的数目。

在研究工作中,得到南海油田的支持和帮助,特别是张应波、吴亚振工程师为此付出了辛勤劳动,在此一并致谢。

参 考 文 献

- [1] 周冠雄等,“基于特征和知识的模式识别专家系统”《机器人》(待发表)。
- [2] 管纪文,“勘探专家系统的研制”《计算机科学》第 4 期(1985)23—24。
- [3] 管纪文,“勘探专家系统的研制”《计算机科学》第 5 期(1985)49—57。
- [4] Bois P., “Application of Pattern Recognition to Oil and Gas Exploration” *IEEE, Tr. GE-21* (1983), 21—28.
- [5] Huang K.Y. and Fu, K. S. “Classification of Ricker Wavelets and the Detction of Bright Spots Using a Tree Classifier”, *Proc. of the 3rd Int. Symposium on Computer Aided Seismic Analysis and Discrimination*, 1983, 15—21.
- [6] Hagen, D. C., “The Application of Principal Components Analysis to Seismic Data Sets”, *Proc. of the 2nd Int. Symposium on Computer Aided Seismic Analysis and Discrimination*, 1982, 98—109
- [7] 徐 雷等,“时间序列特征抽取的一种新方法”《信息与控制》第 4 期(1984),18—22。

A MULTI-SECTION ANALYSIS METHOD FOR THE EXTRACTION OF REGIONS CONTAINING OIL-GAS

ZHOU GUANXIONG HU SUFEI ZOU HAIMING

(Huazhong University of Science and Technology)

ABSTRACT

In [1] the author gave a pattern recognition expert system used to analyse the rocky properties of the section image of earthquake surveying. This system includes both the low-level and high-level processing parts. This paper presents a multi-section automatic analysis method for the variation of rocky properties. This method belongs to the low-level processing part, and provides the classification representations of the variation regions of section image rocky properties for the high-level processing part. The method mentioned above can also be used alone, and give the rocky distribution on the vertical section to extract the regions containing oil-gas. At present, the method using the technique of pattern recognition to determine the regions containing oil-gas is realized on single section image, and the performance is not too good, and far from available. The method presented in this paper utilizes multi-section information, hence it has higher value in both theory and practice. This method has been used to process the practical data on CYBER system, and as a result, the picture of the rocky property distribution is presented, and the regions containing oil-gas in the picture are in conformity to the artificial explanation and the location of the oil hole.

Key words ——Seismic; pattern recognition; feature extraction; possible oil and gas deposit; clustering.