

基于 Hausdorff 距离的测井曲线尖峰深度校正^①

王 辉¹, 李国家², 潘俊辉¹, 高 兴¹

¹(东北石油大学 计算机与信息技术学院, 大庆 163318)

²(大庆油田工程有限公司 市场开发部, 大庆 163712)

摘 要: 提出了一种基于 Hausdorff 距离的测井曲线尖峰深度校正方法. 先利用极值和变量数列分析法选取合理的尖峰, 应用双向 Hausdorff 距离准确选取相似曲线段位置, 再应用单向 Hausdorff 距离确定最佳匹配尖峰, 进而计算尖峰深度校正值. 实验表明, 该方法能实现快速、精准的尖峰深度校正, 为油田的油气层开发提供可靠准确的数据资料.

关键词: Hausdorff 距离; 尖峰; 深度校正; 自然伽马; 变量数列分析

Peak Depth Correction of Logging Curves Based on Hausdorff Distance

WANG Hui¹, LI Guo-Jia², PAN Jun-Hui¹, GAO Xing¹

¹(School of Computer & Information Technology, Northeast Petroleum University, Daqing 163318, China)

²(Department of Market Development, Daqing Oilfield Engineering Co.Ltd, Daqing 163712, China)

Abstract: An effective method of peak depth correction based on the Hausdorff distance is put forward in this paper. The automatic identification of logging curve peak made up the blank of automated processing of logging curve peak depth correction. In this method, the reasonable peak is selected by using the extremum and variable sequence analysis method, and the position of similar curve segment is selected according to the bidirectional Hausdorff distance, then the best matching peak is defined through computing the unidirectional Hausdorff distance. Finally, the peak depth correction of logging curves is computed. The experimental results demonstrate that the method of peak depth correction has better speed and high accuracy, and it provides reliable and accurate data for petroleum reservoir development.

Key words: hausdorff distance; peak depth correction; natural gamma curve; variable sequence analysis

射孔是利用射孔器携带射孔弹等爆炸器材, 下放到井内储油层, 引燃后打开套管形成通道而流出油气. 受两次测井的电缆伸缩程度、采样间距不同等因素影响, 导致测井曲线产生了深度误差^[1]. 射孔定位是一项复杂而精细的工作, 人工校深主要参考两种方法: 一是在下套管前后测得的自然伽马(GR)曲线上, 选择相对应的砂岩层界面深度作差值; 二是选择相对应的致密岩层或泥岩层尖峰深度作差值. 两种方法相辅相成, 计算两种方法得到的差值平均值, 将其作为全井深度校正值. 测井曲线尖峰深度校正是射孔校深的重要工艺, 尖峰的合理选取和深度识别的准确性是尖峰深度校正的关键. 然而, 沿用人工方式计算深度校正

值, 工作量大且计算中存在一定的人为干扰因素. 在计算机技术迅速发展的信息化时代, 只有些少数的研究提出了一些数学模式和设计思想, 却没有完全实现对深度校正的计算机自动处理, 取得精确的校正值, 完全代替手工作业^[2]. 本文主要针对利用计算机, 自动实现尖峰深度校正进行深入探讨和研究.

豪斯多夫(Hausdorff)距离是匹配点特征的一种方法, 它不需要建立点之间的一一对应关系, 只是计算两个点集之间的相似程度, 进行集合间距离的度量, 可以有效地处理很多特征点的情况^[3]. 近年来, Hausdorff 距离多用于模式识别领域的图像匹配^[4]与目标跟踪^[5]的研究. 研究表明, Hausdorff 距离从很大程

① 基金项目:国家自然科学基金(61170132)

收稿时间:2011-12-30;收到修改稿时间:2012-02-23

度上减小了噪声和孤立点对精度以及稳定性的影响,能提高图像匹配的准确率和计算速度,计算量小且鲁棒性强。测井曲线是由数据点组成的集合,数据量比图像数据少,点特征简单,更适合应用 Hausdorff 距离。本文先利用极值法确定岩性突变位置,结合变量数列的峰度与偏度选择满足深度校正要求的曲线尖峰,再应用双向 Hausdorff 距离来确定最佳对比段位置,通过计算单向 Hausdorff 距离寻找最(佳)匹配的尖峰点,最终计算尖峰深度校正值。

1 Hausdorff 距离

Hausdorff 距离,是针对两个有限闭集间距离度量的数学定义^[6]。其数学描述如下:

假设空间中存在两个点集 $A = \{a_1, a_2, \dots, a_p\}$, $B = \{b_1, b_2, \dots, b_q\}$, 则集合、间的 Hausdorff 距离定义为:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (1)$$

$H(A, B)$ 也可以叫做集合 A 与集合 B 间的双向 Hausdorff 距离。公式 1 中的 $h(A, B)$ 和 $h(B, A)$ 分别为集合 A 到集合 B 和集合 B 到集合 A 的单向 Hausdorff 距离(directed Hausdorff distance):

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (2)$$

$$h(B, A) = \max_{b \in B} \min_{a \in A} \|a - b\| \quad (3)$$

$\|\bullet\|$ 表示距离范数,可以取 L_1 (和范数)、 L_2 (欧氏范式)和 L_∞ (极大范数)等。当 $a \in A$ 为集合 A 所有点中距离集合 B 任一点最远的点时,就以 a 与其在 B 中的最近邻点间的距离,作为集合 A 到集合 B 的单向 Hausdorff 距离 $h(A, B)$; $h(B, A)$ 同理。而 A 与 B 间的 Hausdorff 距离则取 $h(A, B)$, $h(B, A)$ 的最大值。两个点集间的 Hausdorff 距离越小,二者相似程度越大。

2 尖峰深度校正基本原理

在泥岩层或致密岩层处的尖峰两侧的数据点,在其所在的曲线段中分布比较均匀,且尖峰幅值比较高,整体上尖峰形态高耸,能够很好的描述泥岩层或致密岩层的正向异常。因此,应选择此类特征尖峰完成尖峰深度校正计算。

基于 Hausdorff 距离的测井曲线尖峰深度校正结合尖峰特征,利用极值点计算、变量数列分析和

Hausdorff 距离的计算来完成尖峰深度校正度计算,其基本思路如下:

(选定下套管前测得的自然伽马曲线为标准曲线,下套管后测得的自然伽马曲线为对比曲线。)

① 利用极值和变量数列分析法,结合标准曲线尖峰特点,选取标准曲线上满足深度校正条件的尖峰。

② 根据标准曲线上确定的尖峰深度,应用双向 Hausdorff 距离,在对比曲线上确定相似曲线段位置。

③ 计算单向 Hausdorff 距离,在相似曲线段内确定与特征点最匹配的尖峰,将两尖峰深度作差值,作为此尖峰对的深度校正值,实现曲线尖峰深度校正。

3 曲线尖峰深度校正步骤与方法

3.1 测井曲线归一化处理

由于两曲线的量纲及测井值大小的不同,计算前,利用公式 4 对标准曲线和对比曲线进行归一化处理:

$$q_i = \frac{q_i - q_{\min}}{q_{\max} - q_{\min}} \quad (4)$$

其中, q_i 为测井曲线在第 i 个采样点处的测井值; q_{\min} 为处理井段的最小测井值; q_{\max} 为处理井段的最大测井值。

3.2 应用变量数列分析法选取标准曲线尖峰

测井曲线是地层物理性质随井深变化的记录,致密岩层或泥岩层尖峰在自然伽马曲线上表现为高尖,幅度很大,在微梯度与微电位曲线上表现为两曲线重合,没有幅度差,如图 1。



图 1 适用于尖峰深度校正的合理尖峰图

结合这些特点,对标准曲线上尖峰选取步骤如下:

3.2.1 利用极值点确定标准曲线尖峰

测井曲线的极值点往往表示测井曲线具有较大的变化,即附近存在着地层突变。通过计算标准曲线极值初步确定曲线尖峰数列 $Q_0 = q_1, q_2, \dots, q_i, \dots, q_{N_0}$, 其中, q_i 为尖峰点在全井数据点中的采样点深度, N_0 为

尖峰个数. 计算所有曲线尖峰幅值平均值 $\bar{w} = \frac{\sum_{i=1}^{N_0} w(q_i)}{N_0}$,

$w(q_i)$ 为点 q_i 的幅值. 将 Q_0 中各极值点幅值与 \bar{w} 逐一做比较, 通过筛选排除小于 \bar{w} 的极值点, 形成新的极值点集合 $Q_1 = \{q_1, q_2, \dots, q_{N_1}\}$, 其中 N_1 为新集合极值点个数.

3.2.2 筛选致密岩层或泥岩层尖峰

由极值点确定的岩性突变位置, 可能为致密岩层或泥岩层尖峰, 也可能为砂岩层界面. 因砂岩层界面在微梯度与微电位曲线上表现为存在幅度差, 致密岩层或泥岩层尖峰表现为两曲线无幅度差. 需分别计算数列 Q_1 中每个尖峰深度上下 L 米范围作为处理井段, 判断两曲线是否存在幅度差, 从数列 Q_1 中剔除出现幅度差的尖峰深度, 最终确定致密岩层或泥岩层尖峰数列 $Q = \{q_1, q_2, \dots, q_N\}$, N 为尖峰极值点个数.

3.2.3 利用采样点分布选取最佳测井曲线尖峰

尖峰深度校正中, 要求选择对称、陡峭的尖峰进行深度校正. 利用变量数列分析中的偏度(Skewness)来描述数据点分布的对称性, 峰度(Kurtosis)来描述数据点分布形态的陡缓程度. 依次选取数列 Q 中尖峰深度, 并以该尖峰深度为准上下各取 L 米作为处理井段, 分别计算其偏度与峰度.

偏度函数定义为:

$$S = \frac{\sum_{i=1}^n (w(x_i) - \overline{w(x)})^3}{n\sigma^3} \quad (5)$$

其中, n 为 $2L$ 米处理井段内的采样点总数, $w(x_i)$ 为采样点 x_i 的幅值, $\overline{w(x)}$ 为采样点的均值, σ 为采样点方差.

$S = 0$, 分布为对称的, 分布形态与正态分布偏度相同; $S > 0$, 分布为右偏态, 均值右边的值比左边多一些; $S < 0$, 分布则相反.

峰度的函数定义为:

$$K = \frac{\sum_{i=1}^n (w(x_i) - \overline{w(x)})^4}{n\sigma^4} - 3 \quad (6)$$

$K = 0$, 其分布形态的陡缓程度与正态分布相同; $K > 0$, 峰形比正态分布的峰形更陡峭; $K < 0$, 峰形比正态分布的峰形来得平缓. 峰度值越大, 说明在总体均值附近出现一个明显的尖峰.

通过对 Q 中每个尖峰变量所在的处理井段做计

算, 选择 Q 中 K 大于零且 S 接近零的采样点为标准曲线尖峰特征点 a_p , 并记录特征点 a_p 的深度为 d_{a_p} .

3.3 基于双向 Hausdorff 距离的对比曲线段匹配

① 在标准曲线特征点 a_p 深度处, 上下各取 L 米的处理井段内采样点个数为 k 个, 形成采样点集 $A = \{a_1, a_2, \dots, a_{p-1}, a_p, a_{p+1}, \dots, a_k\}$, 再在对比曲线上选取相同深度段作为对比处理井段, 形成采样点集 $B = \{b_1, b_2, \dots, b_m\}$.

② 计算标准曲线段的数据集合 A 与对比曲线段的数据集合 B 的 Hausdorff 距离;

$$\text{计算 } d_{a_i, B} = \sum_{j=1}^m d_{a_i, b_j} \text{ . 其中, } i = \{1, 2, \dots, k\},$$

$j = \{1, 2, \dots, m\}$, d_{a_i, b_j} 为集合 A 中采样点 a_i 与集合 B 中采样点 b_j 的距离测度, $d_{a_i, b_j} = \|w(a_i) - w(b_j)\|$.

依次计算集合 A 中各点与集合 B 的距离测度 $d_{a_i, B}$, 形成数列 $D_{a, B} = \{d_{a_1, B}, d_{a_2, B}, \dots, d_{a_k, B}\}$, 数列中取得最大值的点为集合 A 中距离集合 B 任一点最远的点, 记为 a_{\max} . 计算 a_{\max} 与集合 B 中各点距离 d_{a_{\max}, b_j} , 其中 $j = \{1, 2, \dots, m\}$. 计算的数值形成数列 $D_{a_{\max}, B} = \{d_{a_{\max}, b_1}, d_{a_{\max}, b_2}, \dots, d_{a_{\max}, b_m}\}$. 数列中取得最小值的点为集合 B 中距离 a_{\max} 最近的点, 记为 b_{\min} . 此时, 计算集合 A 到集合 B 的单向 Hausdorff 距离 $h(A, B) = d_{a_{\max}, b_{\min}}$.

同理, 可计算出集合 B 到集合 A 的单向 Hausdorff 距离 $h(B, A) = d_{a_{\min}, b_{\max}}$. 根据两个单向 Hausdorff 距离, 计算双向 Hausdorff 距离 $H_0 = \max(h(A, B), h(B, A))$.

③ 在对比曲线上, 向上移动 l 米, 每次移动一个采样间距的距离, 形成新的集合 B , 并将计算得到的双向 Hausdorff 距离值记入数列 $H = \{H_1, H_2, \dots, H_m\}$.

同理, 在对比曲线上向下移动 l 米, 得到数列 $H' = \{H_{-1}, H_{-2}, \dots, H_{-m/2}\}$.

④ 将 H_0 与 H 、 H' 中的各数值作对比, 选择其中最小值, 此时两曲线段相似性最大, 并记录下相应的对比曲线集合 $B' = \{b'_1, b'_2, \dots, b'_m\}$, 以及对比曲线

段移动的距离 Δl 。如图 2 所示。

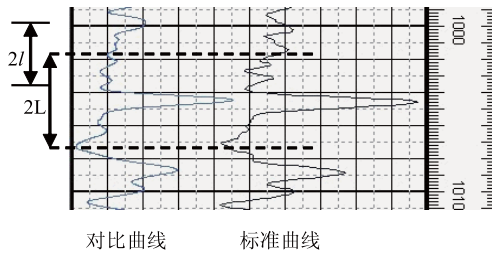


图 2 对比曲线段匹配图

3.4 应用单向 Hausdorff 距离确定最佳匹配点

① 寻找集合 B' 中与特征点 a_p 相匹配的点, 计算 a_p 与集合 B' 中所有点的单向距离测度 $d_{a_p, b_j} = \|a_p - b_j\|$, 其中 $j = \{1, 2, \dots, m\}$. 计算得到数列 $D_{a_p, b'} = \{d_{a_p, b_1}, d_{a_p, b_2}, \dots, d_{a_p, b_m}\}$, 记录数列变量取得最小值时的点 b_q , 为与特征点 a_p 最匹配的点。

② 计算特征点 a_p 与 b_q 的深度差, 即尖峰校正 值, $deep_{a_p, b_q} = d_{b_q} + \Delta l - d_{a_p}$.

选择下一个标准曲线尖峰特征点, 重复以上步骤, 直到特征点的峰度 K 接近于零或所有尖峰均计算完 毕. 再将所有尖峰校正 值与砂岩层顶底界面校正 值作 平均, 作为全井深度校正 值。

4 实验结果与分析

根据前述方法, 编写应用程序(实验环境: VC++6.0, WinXP; 系统参数: Pentium(R)4 CPU3.0、内 存 1G), 对大庆油田实际测井曲线进行自动处理, 表 1 给出了对其中十口井计算后, 与人工测量的对比效 果. 从对比中可以明显看出, 应用极值法和变量数列 分析法选定的标准曲线特征点深度与人工测量误差均 在 0.01 左右, 再利用 Hausdorff 距离算法计算的深度校 正误差均在 0.03 之内, 完全满足实际施工的精度要求, 识别的曲线尖峰深度准确可靠. 同时, 实际生产时, 人工测量并计算一口井的尖峰深度校正需要半小时左 右, 而利用本方法进行自动识别, 仅需要 1 分钟左右。

表 1 实验效果对比分析(单位: 米)

井名	人工测量				Hausdorff 距离				误差	
	标准曲线 尖峰	对比曲线 尖峰	差值	耗时 (s)	标准曲线 尖峰	对比曲线 尖峰	差值	耗时 (s)	差值 误差	耗时 差值(s)
X4-21-F608	997.00	997.11	0.11	1755	997.01	997.09	0.08	53	0.03	1702
X3-20-E934	911.11	911.40	0.29	1824	911.10	911.39	0.29	68	0	1756
X4-3-F122	968.03	968.19	0.16	1768	968.03	968.18	0.15	56	0.01	1712
X12-2-FW27	1156.30	1156.09	-0.21	1790	1156.28	1156.07	-0.21	54	0	1736
X9-1-520	979.90	980.11	-0.21	1855	979.88	979.08	-0.20	69	-0.01	1786
X3-1-E934	910.9	910.74	0.16	1847	910.9	910.73	0.15	67	0.01	1780
X2-10-3E4	975.55	975.60	-0.05	1824	975.53	975.61	-0.08	64	-0.03	1760
X12-3-FB38	969.10	968.89	0.21	1767	969.08	968.88	0.20	54	0.01	1713
X10-2-SFW192	1089.90	1089.68	0.22	1769	1089.88	1089.67	0.21	53	0.01	1716
X4-21-F605	1045.56	1045.38	-0.18	1790	1045.55	1045.36	-0.19	58	0.01	1732

5 结语

通过对实际测井资料的处理效果分析得出, 应用极 值法和变量数列分析法准确选定了测井曲线特征点, 而 Hausdorff 距离能够快速有效识别对比曲线段的准确位 置, 并为寻找最佳相似尖峰提供了有效方法. 同时, 利 用计算机自动识别测井曲线尖峰, 能够为施工人员提供 更精准的全井深度校正 值, 提高射孔准确率和施工质量。

参考文献

- 1 Dashevskiy D. Dynamic depth correction to reduce depth uncertainty and improve MWD/LWD log quality. SPE Drilling and Completion, 2008,23:13-22.
- 2 侯宝成,金海东,田保林,金希卓.射孔自动化校深的数学模

(下转第 170 页)

略,提高网络主动防御的能力,威慑潜在的网络攻击行为.设计通用化的网络追踪溯源技术框架可以在更加广泛的网络空间中应用追踪溯源技术,更好地维护互联网健康发展.

后续还需深入研究非协作网域追踪溯源技术,提高其精准性和实用性,降低其实现复杂度;同时,考虑其隐蔽性;进一步验证并完善非协作追踪溯源整体框架.

参考文献

- 1 彭丹,史志才,陶龙明,马武.IP 追踪技术研究.大连大学学报,2008,29(3):61-66.
 - 2 Savage S, Wetherall D, Karlin A, Anderson T. Practical Network Support for IP Traceback. Department of Computer Science and Engineering University of Washington Seattle, WA, USA,2000.
 - 3 Snoeren AC, Partridge C, Sanchez LA, Jones CE, Tchakountio F, Kent ST, Strayer WT. Hash-Based IP Traceback:BBN Technologies10 Moulton Street, Cambridge, MA 02138,2001.
 - 4 Strayer TW, Jones CE, Tchakountio F, Snoeren AC, Schwartz B, Clements RC, Condell M, Partridge C. Traceback of Single IP Packets Using SPIE, BBN Technologies10 Moulton Street, Cambridge, MA 02138. Proceedings of the DARPA Information Survivability Conference and Exposition (DISCEX'03), 2003.
 - 5 Goodrich MT. Efficient Packet Marking for Large-Scale IP Traceback, Department of Info. & Computer Science, University of California, 2002.
 - 6 Belenky A, Ansari N. IP Traceback With Deterministic Packet Marking. IEEE Commun. Lett.,2003,7(4):162-164.
 - 7 Rizvi B, Fernandez-Gaucherand E. Analysis of Adjusted Probabilistic Packet Marking, IEEE, IPOM2003, 2003,9-13.
 - 8 Kuznetsov V, Simkin A, Sandstrom H. An evaluation of different IP traceback approaches. Department of Computer Science and Electrical Engineering Lulea University of Technology, SE-971 87 Lulea, Sweden, 2003.
 - 9 Cohen D, Narayanaswamy K. Attack Attribution in Non-Cooperative Networks, 2004.
 - 10 Burch H, Cheswick B. Tracing Anonymous Packets to Their Approximate Source. Proc. of the 14th Conf. Systems Administration, Usenix Assoc.,2000,313-322.
 - 11 Staniford-Chen S, Heberlein LT. Holding Intruders Accountable on the Internet. <http://seclab.cs.ucdavis.edu/papers/thumb.ieee95.pdf>. from proceedings of the 1995 IEEE Symposium on Security and Privacy, 1995.
 - 12 Zhang Y, Paxson V. Detecting Stepping Stones. http://www.usenix.org/publications/library/proceedings/sec2000/full_papers/zhangstepping/zhangstepping.pdf. 9th USENIX Security Symposium, Denver, Colorado. August 14-17, 2000.
-
- (上接第 201 页)
- 1 式与系统设计.吉林工业大学自然科学学报,1999,29(95):67-72.
 - 3 Gastaldo P, Zunino R. Hausdorff distance for target detection. Proc. of the 2002 IEEE International Symposium on Circuits and Systems, 2002,5:661-664.
 - 4 雷松泽,姚红革,郝重阳,齐敏.利用 Hausdorff 距离的快速人耳检测.西安工业大学学报,2008,28(3):249-253.
 - 5 陶杰,毕笃彦,吉彦军.一种基于 Hausdorff 距离的目标跟踪算法.微计算机信息,2009,25(2):233-235.
 - 6 Jialing C, James CHC, et al. CT and PET lung image registration and fusion inradiotherapy treatment planning using the chamfer-matching method. Int'l J.Radiation Oncology Biol. Phys, March 1999,43(4):883-891.