



(12) 发明专利

(10) 授权公告号 CN 108055042 B

(45) 授权公告日 2021.03.16

(21) 申请号 201711044322.6

US 2005013361 A1,2005.01.20

(22) 申请日 2017.10.31

US 5426512 A,1995.06.20

US 5510903 A,1996.04.23

(65) 同一申请的已公布的文献号

申请公布号 CN 108055042 A

辛欣等.“一种改进的基于离散余弦变换的声音压缩算法”.《计算机与现代化》.2016,第2016年卷(第10期),第1-5页.

(43) 申请公布日 2018.05.18

(73) 专利权人 华北电力大学(保定)

地址 071003 河北省保定市莲池区永华北大街619号华北电力大学

Hongyan Yang等.“Anti-jamming range imaging using slope-varying LFM signal based on compressed sensing”.《2015 3rd International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)》.2015,第199-203页.

(72) 发明人 黄建才

(74) 专利代理机构 北京知联天下知识产权代理

事务所(普通合伙) 11594

代理人 王冲 吴鑫

Kan Luo等.“Compressed sensing for wireless pulse wave signal acquisition”.《2013 Seventh International Conference on Sensing Technology (ICST)》.2014,第345-350页.

(51) Int.Cl.

H03M 7/30 (2006.01)

审查员 李晓阳

(56) 对比文件

CN 106612437 A,2017.05.03

CN 106597386 A,2017.04.26

CN 102394657 A,2012.03.28

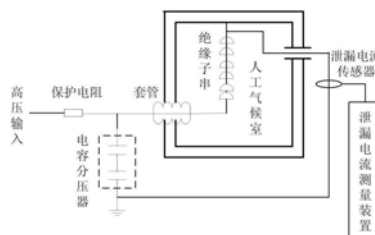
权利要求书2页 说明书8页 附图5页

(54) 发明名称

基于斜率分布的离散波形数据压缩方法

(57) 摘要

本发明涉及一种基于斜率分布的离散波形数据压缩方法,该方法包括:首先,将原始离散波形数据转换为斜率序列,并由这些斜率值组成一个集合;其次,根据斜率序列的分布,选择若干个类别中心,将其作为代表性斜率值;再次,在误差允许的范围内,由各代表性斜率值逼近原始离散数据波形中各相邻点的线段,并将各个线段连接在一起,实现对原始离散波形的逼近,进而实现原始离散波形数据的稀疏表示;然后,对原始离散波形数据的稀疏表示进行压缩编码,实现波形信号的压缩处理。结果表明,本发明所提方法在保留有用成分的情况下,可有效实现波形数据的压缩。



1. 一种基于斜率分布的离散波形数据压缩方法,其特征在于,所述方法包括如下步骤:

步骤1:将原始离散波形数据转换为斜率序列;步骤1具体包括,步骤11:采集待分析的离散波形数据,设其长度为N,并令 $i=1$;步骤12:若 $i \leq N-1$,转步骤13,否则转步骤16;步骤13:取波形上的相邻数据点 (x_i, y_i) 和 (x_{i+1}, y_{i+1}) ,两点构成一条线段,令该线段所在的直线为 L_i ;步骤14:计算直线 L_i 的斜率 $k_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$;步骤15: $i=i+1$,转步骤12;步骤16:构造集合 $S = \{k_1, k_2, \dots, k_N\}$ 表示斜率序列;

步骤2:根据斜率序列的分布,选择M个类别中心;步骤2具体包括,步骤21:根据压缩的要求选择类别数,记为M;步骤22:采用Kmeans聚类算法,将集合S中的元素分为M类,计算每类的平均值,并将其设为类别中心,记为 $C_j, j=1, 2, \dots, M$,并令集合 $C = \{C_1, C_2, \dots, C_M\}$;

步骤3:以各类的类别中心为斜率值,逼近原始离散波形中各相邻点之间的线段,所有线段连接在一起,实现逼近整个离散波形的目的,从而实现原始离散波形的稀疏表示;步骤3具体包括,步骤31:令斜率序列组成的集合为 $R = \{\}$,令逼近原始离散波形的新的坐标集合 $V = \{\}$,取原始离散波形的坐标值 (x_i, y_i) ,此处 $i=1$,并取 $V = V \cup \{(x_i, y_i)\}$;步骤32:令 $j=1$;步骤33:若 $j \leq M$,转步骤34,否则转步骤36;步骤34:计算截距 $b_j = y_i - C_j \times x_i$ 和逼近之后的纵坐标 $y_{i+1}^j = C_j \times x_{i+1} + b_j$;步骤35: $j=j+1$,转步骤33;步骤36:取 $\text{MIN}(\text{ABS}(y_{i+1}^j - y_{i+1}))$,其中, $j=1, 2, \dots, M$,并记此时对应的斜率为 C_k ,将 y_{i+1}^j 赋值给 y_{i+1} ;步骤37:更新集合 $R = R \cup \{C_k\}, V = V \cup \{(x_{i+1}, y_{i+1})\}$;步骤38:令 $i=2$,若 $i < N$,则转步骤32,否则转步骤39;步骤39:求得的R为逼近原始离散波形所用的斜率序列,求得的V为新波形的坐标序列;

步骤4:采用哈夫曼方法对原始离散波形的稀疏表示进行压缩编码,实现信号的最终压缩处理;步骤4具体包括,步骤41:统计集合C中各个元素在集合R中出现的次数 $W_j (j=1, 2, \dots, M)$,并令集合 $W = \{W_1, W_2, \dots, W_M\}$;步骤42:为集合中W构造哈夫曼树T;步骤42具体包括,步骤421:根据集合W构造哈夫曼树的初始集合 $T = \{T_1, T_2, \dots, T_M\}$,其中,每棵二叉树 T_i 中包含斜率值 C_i 及其权值 W_i 的根节点,它的左右子树均为空;步骤422:如果集合T中只有一个元素,则转步骤425,否则转步骤423;步骤423:在集合T中选取两个根节点权值最小的树作为新构造的二叉树的左右子树,新二叉树的根节点的权值为其左右子树的根节点的权值之和;步骤424:从集合T中删除上述两个根节点权值最小的树,并把新二叉树的根节点权值加入集合T中,并转步骤422;步骤425:令集合 $A = \{\}$,并取 $A = A \cup T$;步骤426:若 $A \neq \{\}$,则转步骤427,否则转步骤4212;步骤427:从A中取元素 T_i ,若该节点的左子树不为空则转步骤428,否则转步骤429;步骤428:令 T_i 的左子树为 T_j ,则在集合T中将 T_i 指向 T_j 的连线上设置标号0,并计算 $A = A \cup \{T_j\}$;步骤429:若节点 T_i 的右子树不为空,则转步骤4210,否则转步骤4211;步骤4210:令 T_i 的右子树为 T_k ,则在集合T中将 T_i 指向 T_k 的连线上设置标号1,并计算 $A = A \cup \{T_k\}$;步骤4211:计算 $A = A - \{T_i\}$,并转步骤426;步骤4212:得到最终的哈夫曼树,其中,每个叶子节点包含一个斜率值及其权值,从根节点到该叶子节点的所有连线上的标号序列即为该叶子节点的编码;步骤43具体包括,步骤43:根据哈夫曼树T,为集合R构造压缩编码;步骤431:令编码串 $Z = \emptyset$;步骤432:若集合 $R \neq \{\}$,则转步骤433,否则转步骤435;步骤433:从R中取一个元素,并在哈夫曼树T中查找对应该元素的叶子节点,并得到该元素的编码 z_i ,并令 $Z = Z + z_i$ (此处的“+”表示连接);步骤434:从集合中删除该元素,并转步骤432;步骤435:编码串Z即为压缩编码,结束;

步骤5:对信号进行分析时,对压缩编码进行解压,并恢复用于逼近原始离散波形的数据;步骤5具体包括,步骤51:求解解压后的斜率序列;步骤51具体包括,步骤511:令集合 $R = \{\}$;步骤512:若 $LENGTH(Z) \neq 0$,转步骤513,否则转步骤515,其中,函数 $LENGTH$ 表示取长度;步骤513:从编码串 Z 的起始位置开始,查找第一个能和哈夫曼树中某个叶子节点一致的编码,并取该节点对应的斜率值 C_j ;步骤514:计算 $R = R \cup \{C_j\}$,并将 Z 中已识别的编码去掉,转步骤52;步骤515:最终得到的集合 R 就是解压后的斜率序列,结束;步骤52:利用原始离散波形的第一个数据点和解压后的斜率序列,构造用于逼近原始离散波形的数据;步骤52具体包括,步骤521:取原始离散波形的第一个数据点,将其坐标赋给 (x_i, y_i) ,并令 $i = 1$,集合 $X = X \cup \{(x_i, y_i)\}$;步骤522:若 $R \neq \{\}$,转步骤523,否则转步骤527;步骤523:取 R 中第一个元素对应的斜率值 C_j ,并计算 $R = R - \{C_j\}$;步骤524:计算截距 $b = y_i - C_j \times x_i$;步骤525:计算下一个相邻数据点的坐标,设等间隔采样,并令采样间隔为 d ,则下一个数据点的横坐标 $x_{i+1} = x_i + d$,纵坐标 $y_{i+1} = C_j \times x_{i+1} + b$;步骤526:计算 $X = X \cup \{(x_{i+1}, y_{i+1})\}$,计算 $i = i + 1$,转步骤522;步骤527:最终得到的集合 X 就是解压后的数据集合,结束。

基于斜率分布的离散波形数据压缩方法

技术领域

[0001] 发明涉及属于数据压缩的技术领域,具体涉及一种基于斜率分布的离散波形数据压缩新方法。

背景技术

[0002] 波形数据在工业领域的设备监测中占有非常重要的地位,以电力系统为例,很多监测系统均采用波形数据展开对电力设备的监测,比如变压器、断路器、避雷器、绝缘子等等,均通过波形信号展开对其状态的诊断。

[0003] 采用波形信号监测工业设备的过程中会产生大量数据。以电力系统中的绝缘子为例,可以采用泄漏电流波形信号诊断绝缘子的状态,但是,污秽绝缘子在潮湿情况下会产生大量的报警波形数据,而输电线路绝缘子众多,一旦发生雨雪天气,同一时间段内产生的波形数据将非常巨大,这对数据传输和存储都提出了极高的要求。因此,提高数据传输效率和增强数据存储能力成为一项必要的措施。

[0004] 数据压缩方法可以将原始的数据压缩为相对较少的数据,可以有效提高数据传输的有效性和增强存储空间的利用率。数据压缩方法分为无损压缩和有损压缩,采用前者方法压缩信号之后的成分不会发生损失;采用后者方法压缩信号之后成分有所取舍。由于有损压缩舍掉了其中一部分成分,所以,往往压缩效果更好。若在保证不影响信号分析效果的前提下,有损压缩成为一种很好的方法。

[0005] 当前,已存在许多压缩方法,比如,直接对原始信号采用RLE (Run Length Encoding) 方法进行压缩;再比如,采用小波变换的方法将原始信号变换到时频域后,采用适当的编码方案实现信号的压缩。不同的压缩方法适用于不同特征的信号,即,为了实现更佳的压缩效果,不同的信号应根据自身特点,选择更好的压缩方法。

发明内容

[0006] 本发明的目标是提供一种压缩离散波形数据的新方法,以便提高传输通道和存储空间的利用率。为此本发明专利中提出了一种适合于压缩离散波形数据的方法,该方法根据离散波形信号的斜率分布,选择少数斜率值作为依据,构建各个线段,达到逼近原始信号的目的,从而实现离散波形信号的压缩。

[0007] 本发明提供了一种基于斜率分布的离散波形数据压缩方法,其特征在于所述方法包括如下步骤:

[0008] 步骤1:将原始离散波形数据转换为斜率序列;

[0009] 步骤2:根据斜率序列的分布,选择M个类别中心;

[0010] 步骤3:以各类的类别中心为斜率值,逼近原始离散波形中各相邻点之间的线段,所有线段连接在一起,实现逼近整个离散波形的目的,从而实现原始离散波形的稀疏表示;

[0011] 步骤4:采用哈夫曼方法对原始离散波形的稀疏表示进行压缩编码,实现信号的最终压缩处理。

- [0012] 进一步地,所述步骤1包括下列步骤:
- [0013] 步骤11:采集待分析的离散波形数据,设其长度为N,并令 $i=1$;
- [0014] 步骤12:若 $i \leq N-1$,转步骤13,否则转步骤16;
- [0015] 步骤13:取波形上的相邻数据点 (x_i, y_i) 和 (x_{i+1}, y_{i+1}) ,两点构成一条线段,令该线段所在的直线为 L_i ;
- [0016] 步骤14:计算直线 L_i 的斜率 $k_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$;
- [0017] 步骤15: $i=i+1$,转步骤12;
- [0018] 步骤16:构造集合 $S = \{L_1, L_2, \dots, L_{N-1}\}$ 。
- [0019] 进一步地,所述步骤3包括下列步骤:
- [0020] 步骤31:令斜率序列组成的集合为 $R = \{\}$,令逼近原始离散波形的新的坐标集合 $V = \{\}$,取原始离散波形的坐标值 (x_i, y_i) ,此处 $i=1$,并取 $V = V \cup \{(x_i, y_i)\}$;
- [0021] 步骤32:令 $j=1$;
- [0022] 步骤33:若 $j \leq M$,转步骤34,否则转步骤36;
- [0023] 步骤34:计算截距 $b_j = y_i - C_j \times x_i$ 和逼近之后的纵坐标 $y_{i+1}^j = C_j \times x_{i+1} + b_j$;
- [0024] 步骤35: $j=j+1$,转步骤33;
- [0025] 步骤36:取 $\text{MIN}(\text{ABS}(y_{i+1}^j - y_{i+1}))$,其中, $j=1, 2, \dots, M$,并记此时对应的斜率为 C_k ,将 y_{i+1}^j 赋值给 y_{i+1} ;
- [0026] 步骤37:更新集合 $R = R \cup \{C_k\}$, $V = V \cup \{(x_{i+1}, y_{i+1})\}$;
- [0027] 步骤38:令 $i=2$,若 $i < N$,则转步骤32,否则转步骤39;
- [0028] 步骤39:求得的R为逼近原始离散波形所用的斜率序列,求得的V为新波形的坐标序列。
- [0029] 本发明的基于斜率分布的离散波形数据压缩方法,适用于解决离散波形数据在传输中需要较大带宽和在存储中需要较大空间的问题,通过该方法可以有效降低数据传输带宽和数据存储空间,达到提高利用率的目的。

附图说明

[0030] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例描述中所需要使用的附图作一简单地介绍。显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

- [0031] 图1为根据本发明的实验原理图;
- [0032] 图2为根据本发明的实验中采集到的完整泄漏电流波形图;
- [0033] 图3为根据本发明的泄漏电流波形图;
- [0034] 图4为根据本发明的选择7个斜率时的哈夫曼树;
- [0035] 图5为根据本发明的选择7个斜率时解压后的波形图;
- [0036] 图6为根据本发明的原始波形的频谱和选择7个斜率时压缩后的频谱;
- [0037] 图7为根据本发明的选择5个斜率时的哈夫曼树;
- [0038] 图8为根据本发明的选择5个斜率时解压后的波形图;

- [0039] 图9为根据本发明的原始波形的频谱和选择5个斜率时压缩后的频谱；
 [0040] 图10为根据本发明的选择3个斜率时的哈夫曼树；
 [0041] 图11为根据本发明的选择3个斜率时解压后的波形图；
 [0042] 图12为根据本发明的原始波形的频谱和选择3个斜率时压缩后的频谱。

具体实施方式

[0043] 方法的基本思路是：首先，将原始离散波形数据转换为斜率序列，即，取原始离散数据波形相邻点所在直线的斜率值，并由这些斜率值组成一个集合，求所有斜率值的目的是为了确定少数代表性斜率值，为压缩打下基础；其次，根据斜率序列的分布，选择若干个类别中心，将其作为代表性斜率值；再次，在误差允许的范围内，由各代表性斜率值逼近原始离散数据波形中各相邻点的线段，并将各个线段连接在一起，实现对原始离散波形的逼近，进而实现原始离散波形数据的稀疏表示；然后，对原始离散波形数据的稀疏表示进行压缩编码，实现波形信号的压缩处理；最后，对信号进行分析时，对压缩编码进行解压，并恢复用于逼近原始离散波形的数据。

[0044] 本发明的技术方案为：基于斜率分布的离散波形数据压缩新方法包括如下步骤：

[0045] 步骤1：将原始离散波形数据转换为斜率序列；

[0046] 步骤2：根据斜率序列的分布，选择M个类别中心；

[0047] 步骤3：以各类的类别中心为斜率值，逼近原始离散波形中各相邻点之间的线段，所有线段连接在一起，实现逼近整个离散波形的目的，从而实现原始离散波形的稀疏表示；

[0048] 步骤4：采用哈夫曼方法对原始离散波形的稀疏表示进行压缩编码，实现信号的最终压缩处理；

[0049] 步骤5：对信号进行分析时，对压缩编码进行解压，并恢复用于逼近原始离散波形的数据。

[0050] 所述步骤1包含以下步骤：

[0051] 步骤11：采集待分析的离散波形数据，设其长度为N，并令 $i=1$ ；

[0052] 步骤12：若 $i \leq N-1$ ，转步骤13，否则转步骤16；

[0053] 步骤13：取波形上的相邻数据点 (x_i, y_i) 和 (x_{i+1}, y_{i+1}) ，两点构成一条线段，令该线段所在的直线为 L_i ；

[0054] 步骤14：计算直线 L_i 的斜率 $k_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$ ；

[0055] 步骤15： $i=i+1$ ，转步骤12；

[0056] 步骤16：构造集合 $S = \{k_1, k_2, \dots, k_N\}$ 表示斜率序列。

[0057] 所述步骤2包含以下步骤：

[0058] 步骤21：根据压缩的要求选择类别数，记为M；

[0059] 步骤22：采用Kmeans聚类算法，将集合S中的元素分为M类，计算每类的平均值，并将其设为类别中心，记为 $C_j, j=1, 2, \dots, M$ ，并令集合 $C = \{C_1, C_2, \dots, C_M\}$ 。

[0060] 所述步骤3包含以下步骤：

[0061] 步骤31：令斜率序列组成的集合为 $R = \{ \}$ ，令逼近原始离散波形的新的坐标集合 $V = \{ \}$ ，取原始离散波形的坐标值 (x_i, y_i) ，此处 $i=1$ ，并取 $V = V \cup \{(x_i, y_i)\}$ ；

- [0062] 步骤32: 令 $j=1$;
- [0063] 步骤33: 若 $j \leq M$, 转步骤34, 否则转步骤36;
- [0064] 步骤34: 计算截距 $b_j = y_i - C_j \times x_i$ 和逼近之后的纵坐标 $y_{i+1}^j = C_j \times x_{i+1} + b$;
- [0065] 步骤35: $j = j+1$, 转步骤33;
- [0066] 步骤36: 取 $\text{MIN}(\text{ABS}(y_{i+1}^j - y_{i+1}))$, 其中, $j=1, 2, \dots, M$, 并记此时对应的斜率为 C_k , 将 y_{i+1}^j 赋值给 y_{i+1} ;
- [0067] 步骤37: 更新集合 $R = R \cup \{C_k\}$, $V = V \cup \{(x_{i+1}, y_{i+1})\}$;
- [0068] 步骤38: 令 $i=2$, 若 $i < N$, 则转步骤32, 否则转步骤39;
- [0069] 步骤39: 求得的 R 为逼近原始离散波形所用的斜率序列, 求得的 V 为新波形的坐标序列。
- [0070] 所述步骤4包含以下步骤:
- [0071] 步骤41: 统计集合 C 中各个元素在集合 R 中出现的次数 W_j ($j=1, 2, \dots, M$), 并令集合 $W = \{W_1, W_2, \dots, W_M\}$;
- [0072] 步骤42: 为集合中 W 构造哈夫曼树 T ;
- [0073] 步骤43: 根据哈夫曼树 T , 为集合 R 构造压缩编码。
- [0074] 所述步骤42包含以下步骤:
- [0075] 步骤421: 根据集合 W 构造哈夫曼树的初始集合 $T = \{T_1, T_2, \dots, T_M\}$, 其中, 每棵二叉树 T_i 中包含斜率值 C_i 及其权值 W_i 的根节点, 它的左右子树均为空;
- [0076] 步骤422: 如果集合 T 中只有一个元素, 则转步骤425, 否则转步骤423;
- [0077] 步骤423: 在集合 T 中选取两个根节点权值最小的树作为新构造的二叉树的左右子树, 新二叉树的根节点的权值为其左右子树的根节点的权值之和;
- [0078] 步骤424: 从集合 T 中删除上述两棵树, 并把新二叉树的根节点权值加入集合 T 中, 并转步骤422;
- [0079] 步骤425: 令集合 $A = \{\}$, 并取 $A = A \cup T$;
- [0080] 步骤426: 若 $A \neq \{\}$, 则转步骤427, 否则转步骤4212;
- [0081] 步骤427: 从 A 中取元素 T_i , 若该节点的左子树不为空则转步骤428, 否则转步骤429;
- [0082] 步骤428: 令 T_i 的左子树为 T_j , 则在集合 T 中将 T_i 指向 T_j 的连线上设置标号0, 并计算 $A = A \cup \{T_j\}$;
- [0083] 步骤429: 若节点 T_i 的右子树不为空, 则转步骤4210, 否则转步骤4211;
- [0084] 步骤4210: 令 T_i 的右子树为 T_k , 则在集合 T 中将 T_i 指向 T_k 的连线上设置标号1, 并计算 $A = A \cup \{T_k\}$;
- [0085] 步骤4211: 计算 $A = A - \{T_i\}$, 并转步骤426;
- [0086] 步骤4212: 得到最终的哈夫曼树, 其中, 每个叶子节点包含一个斜率值及其权值, 从根节点到该叶子节点的所有连线上的标号序列即为该叶子节点的编码。
- [0087] 所述步骤43包含以下步骤:
- [0088] 步骤431: 令编码串 $Z = \varphi$;
- [0089] 步骤432: 若集合 $R \neq \{\}$, 则转步骤433, 否则转步骤435;
- [0090] 步骤433: 从 R 中取一个元素, 并在哈夫曼树 T 中查找对应该元素的叶子节点, 并得

到该元素的编码 z_i ,并令 $Z=Z+z_i$ (此处的“+”表示连接);

[0091] 步骤434:从集合中删除该元素,并转步骤432;

[0092] 步骤435:编码串 Z 即为压缩编码,结束。

[0093] 所述步骤5包含以下步骤:

[0094] 步骤51:求解解压后的斜率序列;

[0095] 步骤52:利用原始离散波形的第一个数据点和解压后的斜率序列,构造用于逼近原始离散波形的数据。

[0096] 所述步骤51包含以下步骤:

[0097] 步骤511:令集合 $R=\{\}$;

[0098] 步骤512:若 $LENGTH(Z) \neq 0$,转步骤513,否则转步骤515,其中,函数 $LENGTH$ 表示取长度;

[0099] 步骤513:从编码串 Z 的起始位置开始,查找第一个能和哈夫曼树中某个叶子节点一致的编码,并取该节点对应的斜率值 C_j ;

[0100] 步骤514:计算 $R=R \cup \{C_j\}$,并将 Z 中已识别的编码去掉,转步骤52;

[0101] 步骤515:最终得到的集合 R 就是解压后的斜率序列,结束。

[0102] 所述步骤52包含以下步骤:

[0103] 步骤521:取原始离散波形的第一个数据点,将其坐标赋给 (x_i, y_i) ,并令 $i=1$,集合 $X=X \cup \{(x_i, y_i)\}$;

[0104] 步骤522:若 $R \neq \{\}$,转步骤523,否则转步骤527;

[0105] 步骤523:取 R 中第一个元素对应的斜率值 C_j ,并计算 $R=R - \{C_j\}$;

[0106] 步骤524:计算截距 $b=y_i - C_j \times x_i$;

[0107] 步骤525:计算下一个相邻数据点的坐标,设等间隔采样,并令采样间隔为 d ,则下一个数据点的横坐标 $x_{i+1}=x_i+d$,纵坐标 $y_{i+1}=C_j \times x_{i+1}+b$;

[0108] 步骤526:计算 $X=X \cup \{(x_{i+1}, y_{i+1})\}$,计算 $i=i+1$,转步骤522;

[0109] 步骤527:最终得到的集合 X 就是解压后的数据集合,结束。

[0110] 本方法与现有方法的区别在于:

[0111] 本发明所提的方法为有损压缩,与现有的有损压缩方法的显著区别是:从一种全新的视角对离散波形的压缩展开研究,具体来说,将数据波形变换为斜率序列进行压缩,这是以往方法未曾提到过的。现有的对波形进行有损压缩的典型方法有很多,比如:多项式拟合法、均匀抽样插值法、基于小波变换的压缩方法等。其中,多项式拟合法在波形长度增加、波形较复杂的情况下,难以获得满意的拟合结果,会产生较大的误差;均匀抽样插值法在抽样的过程中可能会丢失重要的特征点,从而无法准确反映数据特征;基于小波变换的压缩方法虽然能够取得较好的效果,但需要解决一系列问题,比如选择最佳小波基、选择最佳阈值、选择小波分解的层数等等,这都增加了压缩的困难。本发明所提的方法避免了上述方法的缺陷,实现有效压缩。

[0112] 为了验证本发明算法的有效性,采用所提的方法对实验中采集的波形数据进行压缩。实验中,采集输电线路绝缘子泄漏电流的试验在人工雾室环境下进行,具体的原理如图1所示。高压线路通过保护电阻穿过套管与人工气候室内的绝缘子串连接,同时通过电容分压器进行分压,后部通过测量装置和泄漏电流传感器进行测量。其中,选择的绝缘子的型号

为FXBW4-110/100。试验中采用涂层染污方式,并用硅藻土模拟不溶性物质,用氯化钠模拟可溶性物质,通过涂刷方式将污秽物涂于绝缘子表面。污秽物的等值盐密 $\rho_{ESDD}=0.1\text{mg}/\text{cm}^2$,灰密 $\rho_{NSDD}=0.5\text{mg}/\text{cm}^2$ 。试验过程中采用升压法展开数据的采集工作,且相对湿度设置为90%、温度设置为27℃。在采样频率为10KHz的情况下,如图2所示为试验过程中采集到的完整泄漏电流波形。为了对信号展开分析,截取一部分典型的波形如图3所示,该波形包含了1000个数据点。

[0113] 根据上述所属步骤,将图3的数据波形变换为斜率序列。

[0114] 首先,选择7个类别中心,此处各个类别中心分别设置为:-3/2、-1、-1/2、0、1/2、1、3/2。对图3的波形进行压缩时,各个斜率出现的次数如表1所示。

[0115] 表1各斜率值出现的次数

斜率值	出现的次数
-3/2	0
-1	103
-1/2	161
0	207
1/2	154
1	103
3/2	271

[0117] 按照表1构建哈夫曼树,如图4所示。按照图4的哈夫曼树对波形进行压缩编码。若压缩前数据按Double类型存储,而Double类型的长度为4个字节(32位),则压缩前和压缩后数据占的位数如表2所示。

[0118] 表2编码前和编码后数据占的位数

时间	占用位数
压缩前	31968
压缩后	2996

[0121] 即,压缩前数据占用的空间是压缩后的约10.7倍,即取得了很客观的压缩效果。

[0122] 对压缩编码进行解压,并恢复逼近原始离散波形的数据,如图5所示。原始离散波形和解压后波形的频谱如图6所示,其中图6(a)为原始离散波形的频谱,图6(b)为解压后波形的频谱。由图3和图5的比较可以看出,压缩方法对波形数据的影响很小;由图6(a)和图6(b)的比较可以看出,主要的周期成分都得到了保留。

[0123] 其次,选择5个类别中心,此处各个类别中心分别设置为:-3/2、-1/2、0、1/2、3/2。

对图3的波形进行压缩时,各个斜率出现的次数如表3所示。

[0124] 表3各斜率值出现的次数

	斜率值	出现的次数
	-3/2	175
	-1/2	255
[0125]	0	164
	1/2	219
	3/2	186

[0126] 按照表3构建哈夫曼树,如图7所示。按照图7的哈夫曼树对波形进行压缩编码。若压缩前数据按Double类型存储,而Double类型的长度为4个字节(32位),则压缩前和压缩后数据占的位数如表4所示。

[0127] 表4编码前和编码后数据占的位数

	时间	占用位数
[0128]	压缩前	31968
	压缩后	2337

[0129] 即,压缩前数据占用的空间是压缩后的约13.7倍,压缩比进一步扩大。

[0130] 对压缩编码进行解压,并恢复逼近原始离散波形的数据,如图8所示。原始离散波形和解压后波形的频谱如图9所示,其中图9(a)为原始离散波形的频谱,图9(b)为解压后波形的频谱。由图3和图8的比较可以看出,压缩方法对波形数据的影响较小;由图9(a)和图9(b)的比较可以看出,主要的周期成分都得到了保留。

[0131] 再次,选择3个类别中心,此处各个类别中心分别设置为:-1/2、0、1/2。对图3的波形进行压缩时,各个斜率出现的次数如表5所示。

[0132] 表5各斜率值出现的次数

	斜率值	出现的次数
[0133]	-1/2	419
	0	163
	1/2	417

[0134] 按照表5构建哈夫曼树,如图10所示。按照图10的哈夫曼树对波形进行压缩编码。若压缩前数据按Double类型存储,而Double类型的长度为4个字节(32位),则压缩前和压缩后数据占的位数如表6所示。

[0135] 表6编码前和编码后数据占的位数

	时间	占用位数
[0136]	压缩前	31968
	压缩后	1579

[0137] 即,压缩前数据占用的空间是压缩后的约20.2倍,压缩比进一步扩大。

[0138] 对压缩编码进行解压,并恢复逼近原始离散波形的数据,如图11所示。原始离散波形和解压后波形的频谱如图12所示,其中图12(a)为原始离散波形的频谱,图12(b)为解压后波形的频谱。由图3和图11的比较可以看出,压缩方法对波形数据的影响较小,且在一定程度上去除了部分噪声,平滑了波形;由图12(a)和图12(b)的比较可以看出,主要的周期成分都得到了保留。

[0139] 由上述3次选择不同的斜率组合进行压缩的结果可以看出,不同的斜率组合的压缩比不同。当压缩比要求较高时,可以选择较少的斜率值进行压缩;反之,选择较多的斜率值。由压缩后的波形及其频谱可以看出,当在保留主要分析成分的前提下,选择较少的斜率值进行压缩,不影响分析效果。

[0140] 本发明的压缩方法可以根据要求灵活选择压缩比,达到更佳的压缩效果。

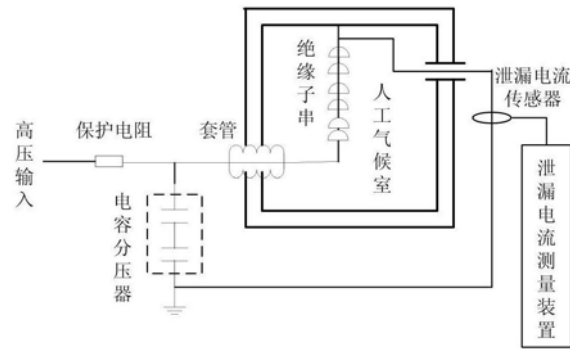


图1

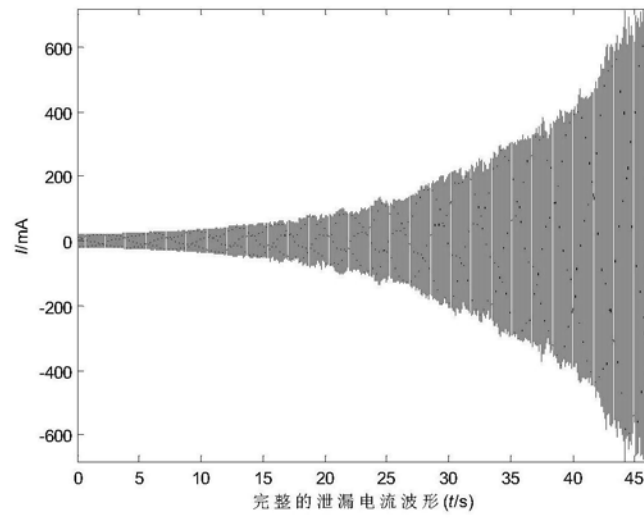


图2

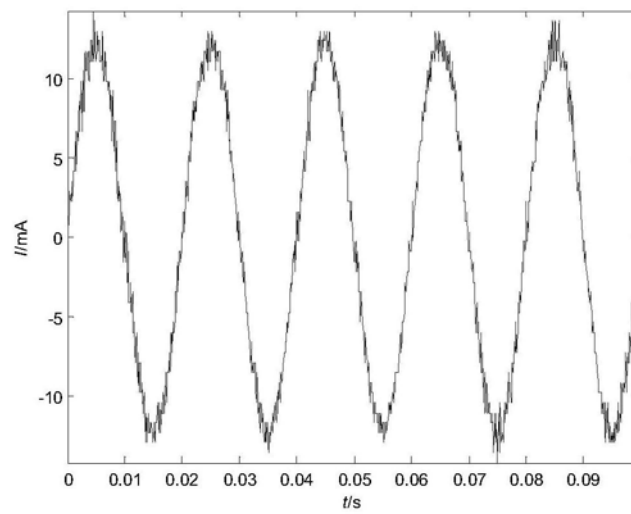


图3

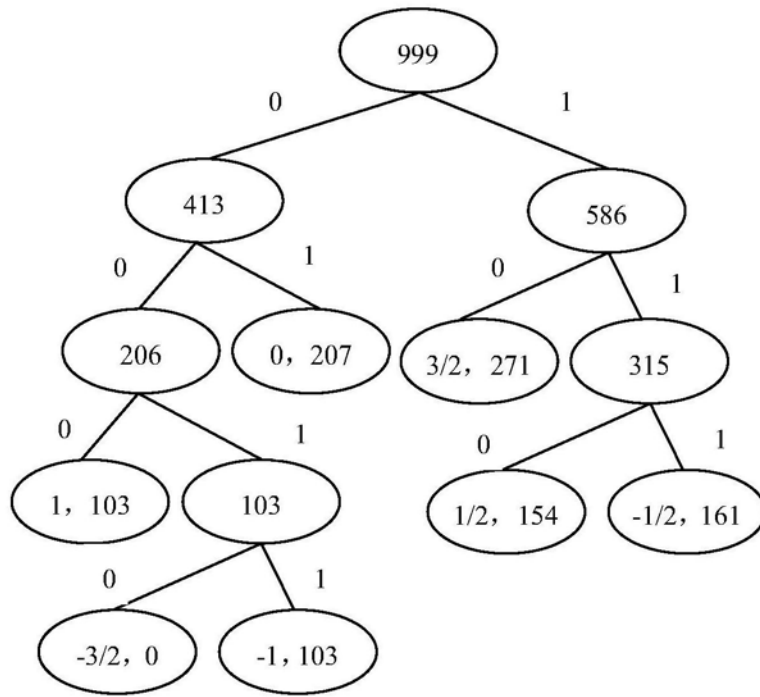


图4

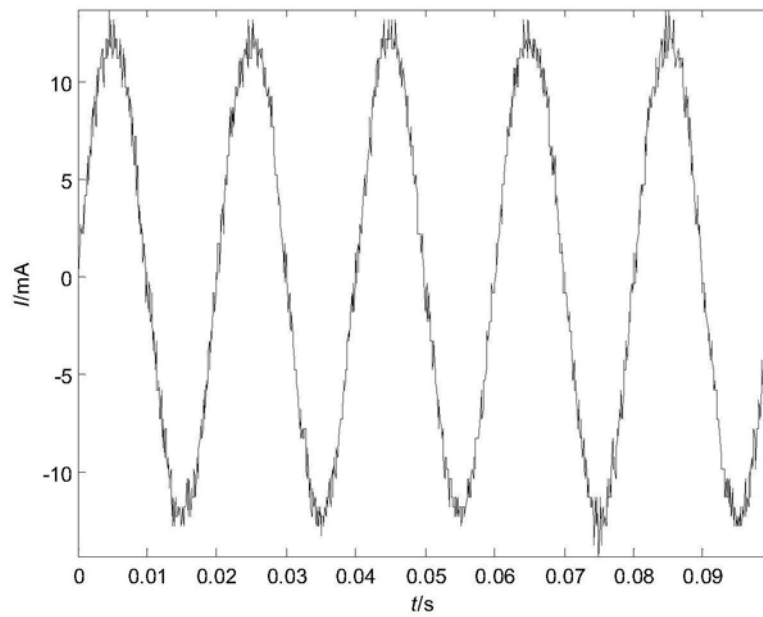


图5

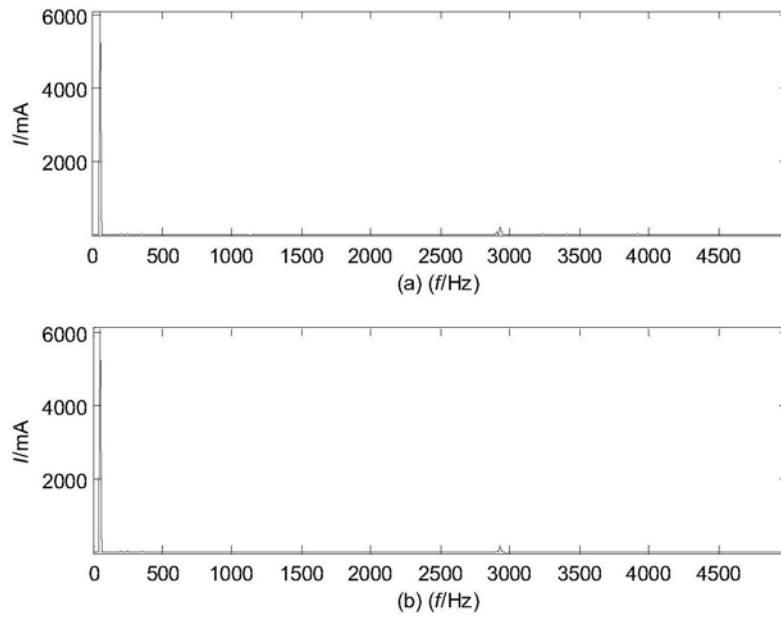


图6

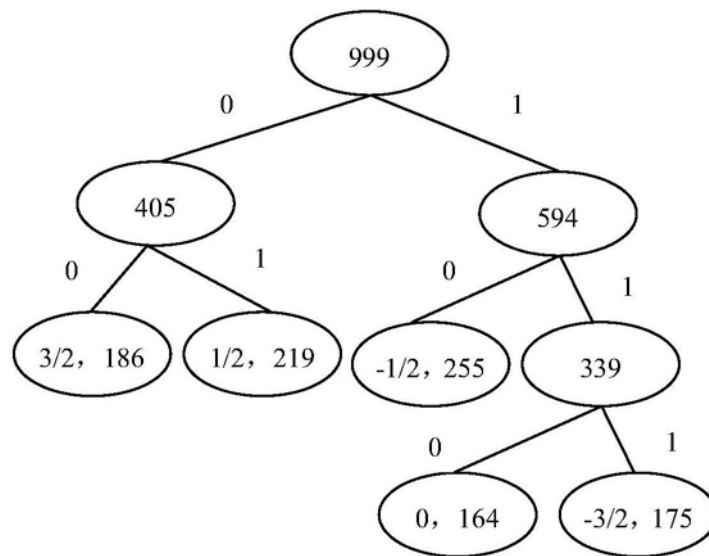


图7

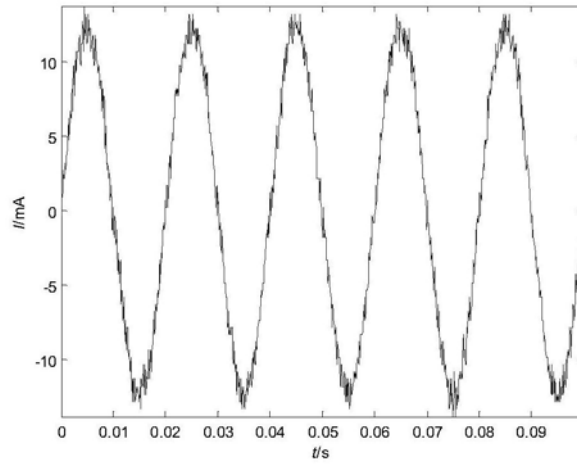


图8

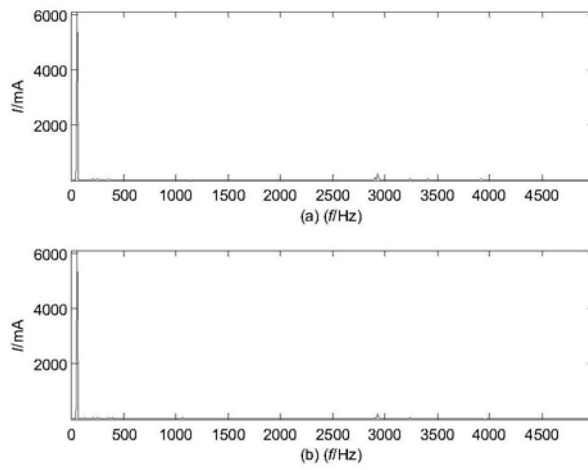


图9

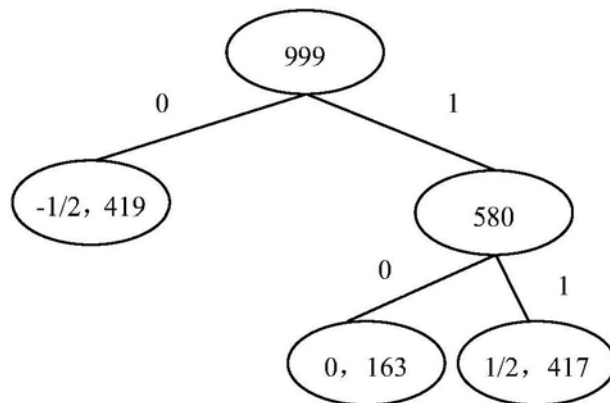


图10

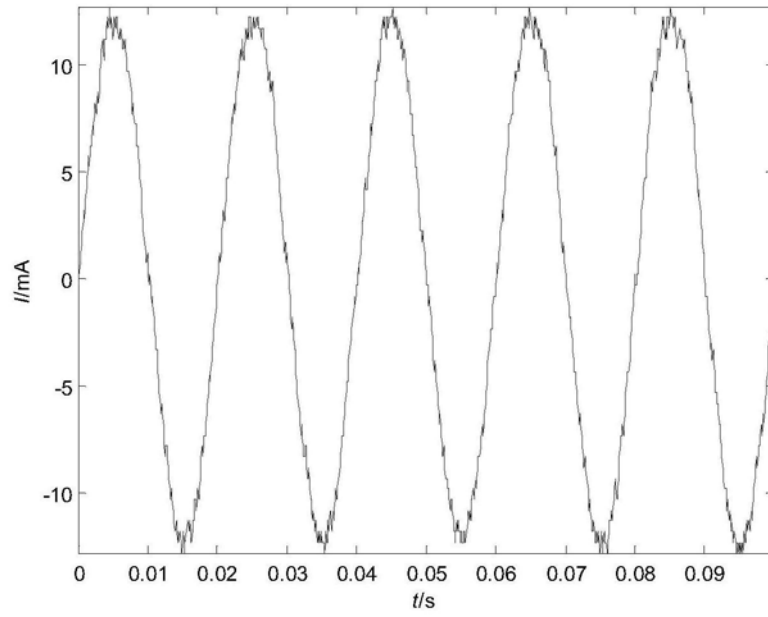


图11

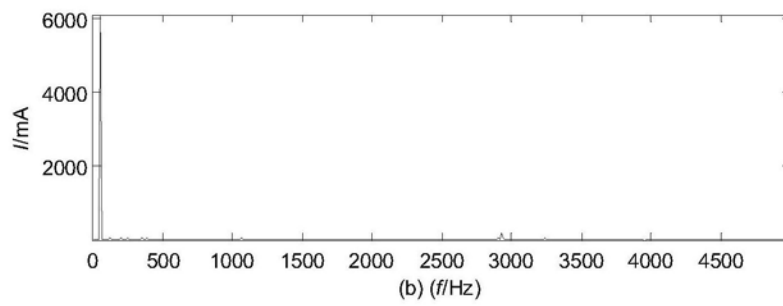
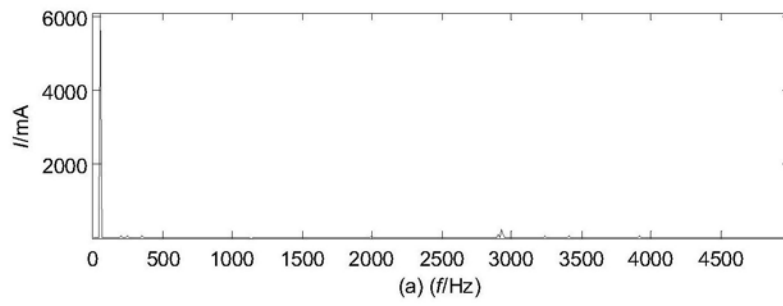


图12