文章编号:1671-8860(2008)06-0652-04　　文献标志码:A

基于经验模式分解的数字高程模型数据伪装方法

刘水强[1]　陈继业[2]　朱鸿鹏[1]

(1　邵阳学院网络信息中心,邵阳市邵水西路,422000)
(2　邵阳学院理学与信息科学系,邵阳市邵水西路,422000)

摘　要:提出了一种基于经验模态分解(EMD)的数字高程模型数据伪装技术。首先利用 SHA-256 单向 Hash 函数产生由种子控制的伪随机序列,扩充序列后再用经验模态分解生成用于伪装的 DEM 数据,伪装后的 DEM 数据具有较高的视觉欺骗性。针对 DEM 数据提出了直方图的概念,通过修改直方图,在伪装的 DEM 数据中可逆地嵌入水印。本文方法可在提取水印后完全恢复伪装 DEM 数据,以及使用种子可完全还原秘密 DEM 数据,算法安全性高。

关键词:信息伪装;数字高程模型;经验模态分解;直方图;可逆数字水印
中图法分类号:P237.3; P231.5

在信息安全领域,为了保证具有重要价值的信息不被非法用户注意和利用,有必要对其进行伪装。传统的信息伪装(information disguising)通常等同于信息隐藏(information hiding)[1,2]。本文研究的信息伪装通过将被保护的信息变换成可辨识的有意义信息,且整个过程无需引入载体,从而实现信息的安全存储和传输。

目前国内外有关真正意义上的信息伪装技术的研究处于起步阶段。文献[3]中将待伪装的灰度图像置乱之后加权值隐藏在另一个用于伪装的灰度图像中,所得到的伪装图像质量存在较严重的噪声且不能完全恢复待伪装的图像。文献[4]针对 DEM 数据提出基于模糊关系的信息伪装技术,但是生成的用于伪装的 DEM 数据不逼真,而且不能识别版权与隐藏重要参数。本文提出一种新的基于经验模态分解 EMD (empirical mode decomposition)的 DEM 数据伪装技术,不但伪装数据隐蔽性好,能在不需要原始数据的情况完整恢复出秘密数据,而且与可逆数字水印[4,5]相结合可以声明 DEM 数据的版权与隐藏重要参数,进一步提高了 DEM 数据的安全性。

1　基于 SHA-256 和二维 EMD 的 DEM 数据生成

SHA(secure hash standard)是由美国国家标准技术研究所和美国国家安全局联合设计的 Hash 算法。SHA-256 是 SHA 的一类,是 MD 结构的迭代 Hash 函数,能将任意长度的消息压缩成 256 bit 的 hash 值,每一 hash 值由 8 个 32 bit 的寄存器联结构成。因此,作为单向散列函数的 SHA-256 有抗碰撞性,即使每秒万亿次的计算机也很难攻破[6]。EMD 是 Hilbert-Huang 变换的核心部分[7],它将变化剧烈的数据分解成若干个固有模态函数 IMF(intrinsic mode function)与一个残余项的和。IMF 相当于数据信号中的高频成分,残余项相当于数据信号中的低频成分,变化相对缓和。EMD 对于变化比较剧烈的部分反映比较敏感,这是由于数据变化较剧烈,容易导致得到的上下包络在该部分差距较大。所以处理随机生成的 DEM 数据时,EMD 方法对芒刺状的高频部分非常敏感,IMF 可以捕捉到较多高频信息,通过几次分解去掉 IMF 便可得到变化缓和的低频成分。EMD 是基于数据本身的局部特性来分

收稿日期:2008-03-28。
项目来源:湖南省科技计划资助项目(06CK3047);湖南省教育厅重点资助项目(06A066);湖南省教育厅科研项目(06C753)。

解的,所以是自适应的,在去掉高频成分后可保持原始信号的总体特征。二维情形的 EMD 通常采用基于 Delaunay 三角化和分片三次多项式插值的曲面拟合方法[8,9]。

DEM 数据是描述地面高程空间分布数字地形模型 DTM (digital terrain mode)的一种表示形式。本文充分利用了 SHA-256 单向 Hash 函数和 EMD 的特点,构造了具有良好的模拟效果,随机模拟生成 DEM 数据的算法。

1) 构造 SHA-256 单向 Hash 函数[7]。对每一个待生成的伪装 DEM 数据取 1 个种子 seed,由构造的单向 Hash 函数生成伪随机序列,作为初始数据 v,然后细分网格进行分片三次多项式插值,扩充 v 得到初始 DEM 数据 Y_0。

2) 对 Y_0 进行二维 EMD。先确定 DEM 数据的所有局部极值点(极大值和极小值);用三次多项式函数连接所有的局部极大值点,得到上包络 φ_k,对应地连接所有的局部极小值点,得到下包络 ϕ_k,从而得到局部均值 v_k,其中,$v_k=\frac{\varphi_k+\phi_k}{2}$。由于 IMF 应该具有局部零均值,从 DEM 数据中去掉均值,令 $Y_{k+1}=Y_k-v_k$,此处,k 取经验值 4,即对 Y_0 进行 4 次 EMD 就可得到效果比较逼真的 DEM 数据,如图 1。

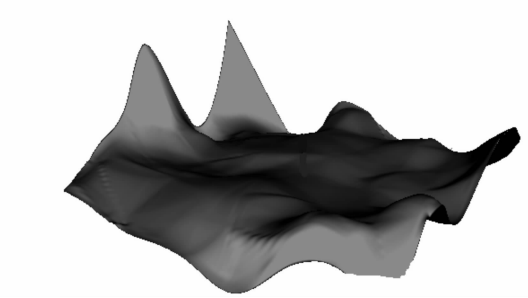


图 1　模拟 DEM 数据

Fig. 1　Secret DEM Data

2　信息的伪装

2.1　生成伪装数据

若秘密的 DEM 数据记为 $\boldsymbol{X}$,即矩阵 $\boldsymbol{X}=(x_{ij})$;随机生成的 DEM 数据记为 $\boldsymbol{Y}$,即矩阵 $\boldsymbol{Y}=(y_{ij})$;伪装后的 DEM 数据记为 $\boldsymbol{Z}$,即矩阵 $\boldsymbol{Z}=(z_{ij})$。X、Y、Z 有相同大小 $N_1\times N_2$,$z_{ij}=y_{ij}+w x_{ij}$,其中 w 是取值范围在 0.1～0.9之间的权值。

2.2　可逆水印算法

伪装后的 DEM 数据记为 Z,每个坐标点的取值 $z_{ij}\in[m,M]$,其中,$m=\min\limits_{i,j}z_{ij}$ 和 $M=\max\limits_{i,j}z_{ij}$ 都是浮点数。将 $[m,M]$ 等分成 n 个小区间 $h_k=[m+kh,m+(k+1)h)$,从左到右依次有 $k=0,1,2,\cdots,n-1$,每个小区间的长度 $h=\frac{M-m}{n}$。将海拔高度值 z_{ij} 落于第 k 区间 h_k 的坐标点的数目记为 $h(k)$,称为高度关于区间 k 的频数。$h(k)$可按下述方法确定:设置初始值 $h(k)=0$,对 DEM 数据按从上到下、从左到右的顺序进行扫描,若 $z_{ij}\in h_k$,则置 $h(k):=h(k)+1$。重复上述操作,最后输出 $h(k)(k=0,1,2,\cdots,n-1)$。记整幅数据的扫描频数总体结果为 $H(Z)$,则 $H(Z)=\{h_k\times h(k)\mid k=0,1,\cdots,n-1\}$,$H(Z)$为 Z 的直方图。为方便起见,将 $h(k)\in H(Z)$ 与 $h_k\times h(k)\in H(Z)$用作同义语。

秘密 DEM 数据的版权信息、重要参数(地理坐标等)将作为水印信息嵌入伪装后的 DEM 数据中。水印嵌入前先进行加密处理,只有较高权限的用户才持有其密钥 key1,加密后的水印是均匀分布的 0、1 序列。先考虑最简单的情况,假设伪装后的 DEM 数据的直方图 $H(Z)$中只有一个最大值,记为 $h(a)$,不妨有 $0<a<n-1$。水印嵌入的步骤如下。

1) 对伪装后的 DEM 数据进行区间划分,顺序扫描数据 Z,生成 Z 的直方图,记为 $H(Z)$。

2) 在 $H(Z)$中找到最大值 $h(a)$,$0<a<n-1$,假设 a 左边的坐标点个数少于右边坐标点个数。

3) 将 $H(Z)$中所有序号 $k\in[0,a-1]$的区间都统一向左移动 1 个区间单位,即序号 $k\in[0,a-1]$中的区间的坐标点的海拔高度值 z_{ij} 都减少 h,使得落在第 $a-1$ 个区间的坐标点个数为 0,得到数据 Z'。规定水印嵌入时区间移动的记号为 s,左移记为 $s=0$,右移 $s=1$。这里是左移的情况,于是 $s=0$。

4) 扫描数据 Z',一旦遇到海拔高度值落于区间 a 的坐标点,就对比将要嵌入的水印信息。如果水印信息是 1 bit,则把这个坐标点的海拔高度值改为 $z_{ij}-h$;如果水印信息是 0,则保持这个坐标点的海拔高度值 z_{ij} 不变。

这样,便得到了含有水印信息的 DEM 数据 Z''。把 m、h、a 和 s 称为定位信息,再将定位信息和控制伪随机序列产生的种子 seed 进行加密处理,合法用户(较高、较低权限用户)持有其密钥 key2。水印的提取步骤如下。

1) 对含水印的 DEM 数据 Z''进行区间划分。由于 $s=0$ 区间已向左移动了 h,故从 $m-h$ 开始按步长 h 将 Z''分成 $n+1$ 个小区间,第一个小区

间记为-1。然后顺序扫描，如果遇到海拔高度值落于区间$a-1$的坐标点，提取水印信息1 bit；如果遇到海拔高度值落于区间a的坐标点，提取水印信息0。

2) 扫描Z''，对任何海拔高度值z_{ij}落于区间$[-1,a-1]$的坐标点，将其海拔高度值增加h。

经过以上两步，DEM数据Z便可无损地恢复。数据嵌入量为$C=h(a)$。通常情况下要求嵌入的信息量可能大于$h(a)$，解决的办法是寻找多个最大值$h(a_i)(i=1,2,\cdots,k)$逐步执行。

水印嵌入步骤如下。

1) 对DEM数据进行区间划分，顺序扫描数据Z，生成Z的直方图，记为$H_1(Z)$。找到$H_1(Z)$中的最大值$h(a_1)$，按照2)～4)，记录s_1，生成数据Z_1^*，判断是否有$C\leqslant h(a_1)$。若是则转入4)；否则从原始最低海拔m开始按步长h向两边进行区间划分，顺序扫描原始数据，生成Z_1^*的直方图$H_2(Z)$。

2) 找到$H_2(Z)$中的最大值$h(a_2)$，按照2)～4)，记录s_2，生成数据Z_2''。判断是否有$C\leqslant h(a_1)+h(a_2)$，若是则转入4)；否则，从原始最低海拔m开始按步长h向两边进行区间划分，顺序扫描原始数据，生成Z_2''的直方图$H_3(Z)$。

3) 重复以上步骤，直到$C\leqslant h(a_1)+h(a_2)+\cdots+h(a_k)$，$2\leqslant k\leqslant K$，其中$K$是规定的常数，$h(a_k)\in H_k(Z)$，$H_k(Z)$是$Z_k^*$的直方图。转入4)。

4) 计算终止，输出Z_k''。

称伪装后的DEM数据Z中最低海拔、每个小区间的长度、小区间的序号、$C-(h(a_1)+h(a_2)+\cdots+h(a_k))$及每一步中左右移动的记号$s_t(s_t\in\{0,1\},t=1,2,\cdots,k)$为定位信息。将定位信息和控制伪随机序列产生的种子seed进行加密处理，合法用户(较高、较低权限用户)持有其密钥key2。实际情况下，对大多DEM数据，当$k=2$时，水印的嵌入量已经大于10 000 bit了，故不需要较多最大值的选取。

水印提取步骤如下。

1) 读入小区间的序号及左右移动的记号$s_t\in\{0,1\}(t=1,2,\cdots,k)$由$s_t$中0的个数找到DEM数据$Z$中$m$的位置，按步长$h$向两边进行区间划分，顺序扫描数据$Z_k''$。不妨设$s_k=1$，如果遇到海拔高度值落于区间$a_k+1$的坐标点，提取水印信息1 bit；如果遇到海拔高度值落于区间a_k的坐标点，提取水印信息0。

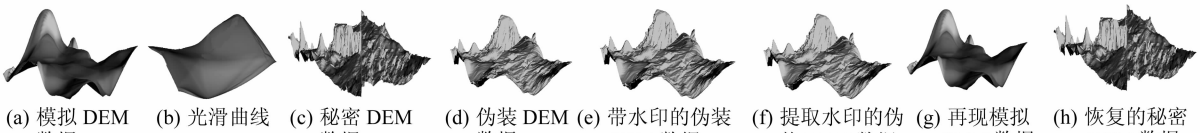
2) 再次扫描Z_k''，对任何海拔高度值z_{ij}落于区间$[a_k+2,n-1+\sum_{t=0}^{k}s_t]$的坐标点，将其海拔高度值减去$h$，得到数据$Z_{k-1}''$。

3) 重复以上步骤，直到完全提取水印。

需要注意的是，按上述方法提取水印后的$C-(h(a_1)+h(a_2)+\cdots+h(a_k))$个0不属于水印信息。

3 实验结果

本文仿真中采用的DEM数据尺寸均为512×512，记录高程的数据类型是单精度浮点数。构造SHA-256单向Hash函数作为伪随机序列发生器，生成100×100的伪随机序列，然后将100×100的网格细分512×512，再进行分片三次多项式插值，得到初始的512×512大小的DEM数据。多次实验可知，对初始的DEM数据进行4次经验模态分解就可得到效果比较逼真的DEM数据(图2(a))，若分解5次，得到的数据其三维显示是光滑平整的曲面(图2(b))。实验结果见图2。



(a) 模拟DEM数据 (b) 光滑曲线 (c) 秘密DEM数据 (d) 伪装DEM数据 (e) 带水印的伪装DEM数据 (f) 提取水印的伪装DEM数据 (g) 再现模拟DEM数据 (h) 恢复的秘密DEM数据

图2 实验结果

Fig. 2 Experiment Results

4 结 语

本文基于经验模态分解的DEM数据伪装技术是全新的，由于构造了SHA-256单向Hash函数，通过种子可以再现用于伪装的DEM数据，从而可以完全恢复出秘密DEM数据。算法能够很好地保护DEM数据的版权，安全性高，可实现用户权限的多级管理；另一方面，文中关于信息伪装的算法是全可逆的，可以确保数据恢复无损。但基于4次经验模态分解生成的模拟DEM数据与实际DEM数据之间的差异是客观的，因此，整体

算法造成恢复后的 DEM 与原始 DEM 之间存在差异性,通常以实验来检验其逼真效果,目前尚无系统、成熟的定量描述差异性的理论分析方法。

参 考 文 献

[1] 张雷,平西建,张涛. 一阶统计特征保持的图像信息伪装算法[J]. 计算机辅助设计与图形学学报,2005,17(1):99-104

[2] 张涛,平西建. 基于差分直方图实现 LSB 信息伪装的可靠检测[J]. 软件学报,2004, 15(1):151-158

[3] 杨尚英,朱虹,李永盛. 一种数字图像的信息伪装技术——信息隐藏[M]. 西安:电子科技大学出版社,2001

[4] 罗永,成礼智,陈波,等. 基于模糊关系的 DEM 数据信息伪装技术研究[J]. 模糊系统与数学,2004,18(3):116-120

[5] Celik M U, Sharma G, Tekalp A M. Lossless Watermarking for Image Authentication: a New Framework and an Implementation[J]. IEEE Trans on Image Processing, 2006, 15(4):1 042-1 049

[6] Ni Zhicheng, Shi Yunqing, Ansari N, et al. Reversible Data Hiding[J]. IEEE Trans on Circuits and Systems for Video Technology, 2006, 16(3): 354-362

[7] Gilbert H, Handshuh H. Security Analysis of SHA-256 and Sisters[C]. SAC2003, Ottawa, 2003

[8] Huang N E, Shen Z, Long S R. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis [C]. The Royal Society, London, 1998

[9] Christophe Damerval, Sylvain Meignen. A Fast Algorithm for Bidimensional EMD[J]. IEEE Signal Processing Letters , 2005, 12(10):701-704

第一作者简介:刘水强,教授,主要从事数字图像处理研究。
E-mail:lsq0316@yahoo.com.cn

Information Disguising for Digital Elevation Model Data via Empirical Mode Decomposition

LIU Shuiqiang[1]　*CHEN Jiye*[2]　*ZHU Hongpeng*[1]

(1　Network and Information Center, Shaoyang University, West Shaoshui Road, Shaoyang 422000, China)

(2　Department of Science and Information, Shaoyang University, West Shaoshui Road, Shaoyang 422000, China)

Abstract: A novel information disguising method based on empirical mode decomposition is proposed. The pseudorandom sequence controlled by seeds of the SHA-256 one-way hash function is generated; and digital elevation model data for disguising is achieved by decomposing the expanded pseudorandom sequence via empirical mode decomposition(EMD). The high vision fraudulence is obtained for disguised DEM data. Furthermore, the concepts of the histogram for DEM data is also proposed; and the watermarking was reversibly embedded in the disguised DEM data by modifying its histogram. The disguised DEM data can be completely reconstructed without any distortion from the marked data after the watermark has been extracted. The secret DEM data can be recovered via the seed. The proposed algorithm owns high security.

Key words: information disguising; digital elevation model data; empirical mode decomposition; histogram; reversible watermarking

About the first author: LIU Shuiqiang, professor, majors in digital image processing.
E-mail: lsq0316@yahoo.com.cn