引用格式：张鑫港,闫浩文,张黎明.一种用于DEM数据认证与篡改定位的感知哈希算法[J].地球信息科学学报,2020,22(3):379-388. [Zhang X G, Yan H W, Zhang L M. A perceptual hash algorithm for DEM data authentication and tamper localization[J]. Journal of Geo-information Science, 2020,22(3):379-388.] **DOI:**10.12082/dqxxkx.2020.190336

一种用于DEM数据认证与篡改定位的感知哈希算法

张鑫港[1,2,3],闫浩文[1,2,3*],张黎明[1,2,3]

1. 兰州交通大学测绘与地理信息学院,兰州 730070；2. 地理国情监测技术应用国家地方联合工程研究中心,兰州 730070；
3. 甘肃省地理国情监测工程实验室,兰州 730070

A Perceptual Hash Algorithm for DEM Data Authentication and Tamper Localization

ZHANG Xingang[1,2,3], YAN Haowen[1,2,3*], ZHANG Liming[1,2,3]

1. Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730070, China; 2. National-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring, Lanzhou 730070, China; 3. Gansu Provincial Engineering Laboratory for National Geographic State Monitoring, Lanzhou 730070, China

Abstract: As a type of fundamental and important geographic data, the integrity of DEM data cannot be ignored. The commonly used technology for data integrity authentication is mainly based on traditional cryptography and digital watermarking technology. The former is very sensitive to the change of every bit of data, suitable for accurate authentication of text data; while latter is mostly based on data carrier for authentication, seldom considers if DEM data content changes or not, and needs additional secure channels and communication media. In this paper, based on the requirement of authenticity and integrity of DEM data and the shortcomings of related authentication algorithms, a DEM data authentication algorithm was designed based on the Perceptual Hashing technology, which can achieve tamper localization. Perceptual hashing is a kind of method that maps multimedia data unidirectionally into perceptual summary sets (i.e. hash sequences). It inherits the characteristics of traditional Hash functions such as unidirectionality, anti- collision, and summarization, and is robust to the operation of content preservation, so it can better meet the requirements of DEM data authentication. The main ideas of this algorithm are as follows: Based on the characteristics of a large amount of DEM data and abundant details, the DEM data is divided into regular and non- overlapping grids. Feature extraction is the key of Perceptual Hashing algorithm. In this paper, the discrete cosine transform was used to extract features and generate the eigenvector matrix. Then the eigenvector matrix was digested. Next, the simplified eigenvector matrix was scrambled by using a Logistic chaotic system to meet the security requirements of authentication. Followingly, the scrambled matrix was quantized and coded to generate perceptual hash sequence. In the data authentication stage, the relative error of elevation between the original data and the data to be validated was

收稿日期：2019-06-26；**修回日期**：2019-12-29.

基金项目：国家自然科学基金项目(71563025、41761080)；甘肃省高等学校产业支撑引导项目(2019C-04)；兰州交通大学优秀平台支持项目(201806)。[**Foundation items:** National Natural Science Foundation of China, No.71563025, 41761080; Industrial Support and Guidance Projects of Colleges and Universities in Gansu Province, No.2019C-04; Funded by Lanzhou Jiaotong University Excellent Platform (LZJTU EP), No.201806.]

作者简介：张鑫港(1997—)，男，山东济宁人，硕士生，主要从事空间数据安全的研究。E-mail: 1441861764@qq.com

***通讯作者**：闫浩文(1969—)，男，甘肃武威人，博士，教授，主要从事地图自动综合、空间关系、空间数据安全、微地图等研究。E-mail: haowen2010@gmail.com

calculated firstly. Subsequently, the perceptual hash sequence of the original data and the data to be validated was normalized to measure the Hamming distance. Combined with the determination threshold, the DEM data was authenticated. The scope of tampering would be located on the "grid unit" mentioned above. The algorithm has strong robustness against DEM data format conversion, watermarking embedding and other attacks. It is sensitive to various operations of changing contents, and can recognize and locate minor tampering of DEM data. Compared with the traditional DEM authentication algorithm, this algorithm innovatively regards "content" as the sole criterion of identity determination, which effectively compensates for the traditional digital watermarking method's excessive dependence on information carriers.

Key words: DEM; perceptual hash; grid partitioning; discrete cosine transform; data authentication; root mean square error of elevation

***Corresponding author:** YAN Haowen, E-mail: haowen2010@gmail.com

摘要：DEM数据作为重要的基础地理信息数据，其数据完整性问题不容忽视。基于DEM数据完整性认证的要求，以及相关认证算法的欠缺，本文运用感知哈希技术设计了一种DEM数据认证算法，并可实现篡改定位。因DEM数据具有数据量大、细节丰富的特点，首先对其进行规则格网划分，将其划分为互不重叠的格网单元；然后对格网单元数据进行DCT分解，提取数据的特征信息以生成特征向量矩阵，并对特征向量矩阵进行摘要化处理；随后，使用Logistic混沌系统对简化后的特征向量矩阵进行置乱；对置乱矩阵进行量化、编码后，便可生成感知哈希序列。在数据认证时，首先计算原始数据与待验证数据的高程相对中误差，再将二者的感知哈希序列进行归一化汉明距离度量，结合判定阈值，即可对DEM数据进行数据认证与篡改定位。该算法对DEM数据的格式转换、水印嵌入等攻击有较强的鲁棒性，对各类改变内容的操作具有敏感性，并可实现DEM数据微小篡改的识别与定位。与已有的DEM完整性认证方法相比，将DEM数据的“内容”作为完整性度量的重要标准，在具体应用中更具有实用价值。

关键词：DEM；感知哈希；格网划分；离散余弦变换；数据认证；高程中误差

1 引言

数字高程模型（Digital Elevation Model，DEM）数据是国家重要的基础地理信息数据[1]。目前DEM数据被广泛应用于测绘、地质、水文、气象、工程建设与国防军事等领域[2-7]，与此同时如何鉴定其真实性与完整性也成为备受关注的研究热点。DEM数据在存储、使用与传播等过程中，易受到许多有意或无意的修改或攻击，使数据的完整性难以得到保证。因此，对DEM数据完整性认证技术的研究是极重要的。

数据的完整性指的是数据的精确性与可靠性[8]，即在存储与传输过程中数据的一致性，防止数据被非法用户篡改。常用的数据完整性认证技术主要有：基于传统密码学的认证[9]与基于数字水印技术的认证[10]。其中，传统密码学方法主要通过哈希函数产生数字签名，以该签名作为认证信息实现数据判定。然而密码学方法对数据的每比特变化都十分敏感，即易产生雪崩效应[11]。DEM数据在传输与使用的过程中，可能会经受水印嵌入、格式转换等变化，却并不改变其承载的内容，故密码学方法难以满足DEM数据认证的需求。数字水印技术是一种有效的数据认证手段，它主要通过向数据中嵌入可见的或不可见的版权标记来实现数据认证，已有诸多学者对此进行了研究[12-16]。但数字水印会修改原始数据，同时需要额外的安全信道与可靠的通信介质，其关注点是DEM数据的载体信息。认证从本质上是认证其承载的有效信息是否一致，而非信息载体，所以数字水印技术也没有完全解决DEM数据认证的问题。而感知哈希技术可以为DEM数据认证提供一种新的思路。

感知哈希[17]是将多媒体数据单向映射为感知摘要集（即哈希序列）的一类方法，它继承了传统哈希函数单向性、抗碰撞性与摘要性的特点，且对内容保持的操作具有鲁棒性[18]，能更好地满足DEM数据认证的要求。其特点是哈希值取决于数据所承载的内容，即如果数据的内容不发生明显改变，哈希序列保持不变。近年来一些针对图像的感知哈希算法被不断提出，并被广泛应用于图像检索、数据认证等领域。Ruchay等[19]提出了一种基于级联算法的图像感知哈希算法，首先用短哈希进行初始化，然后对处理后的结果进行全哈希处理，该算法

时间开销较小、识别率高，但不可实现篡改定位。Wang等[20]提出了一种基于感知散列和包聚类算法的复制-移动伪造盲认证方案，可抗击白高斯噪声、高斯模糊与调整对比度等攻击，却只适用于Copy-Move类型的篡改检测与定位。张春艳等[21]运用Henon映射对图像频域进行加密，并结合了离散余弦变换（Discrete Cosine Transform，DCT）与离散小波变换（Discrete Wavelet Transform，DWT）进行哈希特征提取，实现了医疗图像的快速检索。Wang等[22]利用Watson视觉模型提取视觉敏感特征，将基于图像块的特征和基于关键点的特征相结合，生成鲁棒的感知哈希码。该方法在容忍内容保护操作的感知稳健性和检测恶意篡改的感知敏感性之间实现了权衡，有较好的应用效果。Ding等[23]提出了一种基于感知哈希的高分辨率遥感影像完整性认证方法，提取多尺度边缘特征并进行主成分分析得到鲁棒特征，对内容保持的操作具有鲁棒性，可识别微小篡改并定位篡改位置。由此可见，感知哈希技术在数据认证领域有着较好的应用前景，然而目前针对DEM数据的感知哈希算法并不多见。且上述方法[19-23]多针对图像视觉内容的一致性进行认证，不同于DEM数据的使用场景与数据特点。DEM数据认证区别于其他数字图像认证的关键在于，其在量测精度、细节丰富度、数据海量性等方面的要求均高于普通数字图像。故上述方法并不能直接用于DEM数据，本文拟探究感知哈希技术在DEM数据上的针对性应用方案。

综上所述，针对已有DEM数据认证算法的不足，以及感知哈希在数据认证上的广阔前景，本文拟结合格网划分思想，运用感知哈希技术设计一种对内容操作保持具有鲁棒性，对内容改变操作具有敏感性的DEM感知哈希算法，并充分顾及DEM数据的细节特征，实现篡改定位。

2 DEM数据认证感知哈希算法

2.1 算法思想

对于普通数字图像而言，其感知哈希算法的一般流程为：① 图像预处理，以方便后续的特征提取；② 特征提取，即运用DWT、DCT、SVD分解等方法提取图像的感知特征集；③ 特征简化，为消除感知特征集中的冗余特征，满足特征集摘要性；④ 量化、编码简化后的特征集，生成感知哈希序列；⑤ 相似性度量，将原始图像与待验证图像的哈希序列进行比较，根据哈希相似度判定是否为同一数据，进而实现数据的认证。其中，特征提取是感知哈希算法的关键。特征提取是从图像信息中提取鲁棒性特征的过程，提取到的特征对压缩、噪声添加等合理失真并不敏感。

DEM数据的感知哈希算法作为感知哈希技术的针对性应用，应具有感知哈希的所有基本特点：① 摘要性：能够用尽量少的数据量，尽可能的描述DEM的特征信息；② 鲁棒性：能够容纳不改变内容的合理失真（DEM合理失真的界定将在后文给出）；③ 篡改敏感性：对明显改变内容的操作敏感，能实现检测。

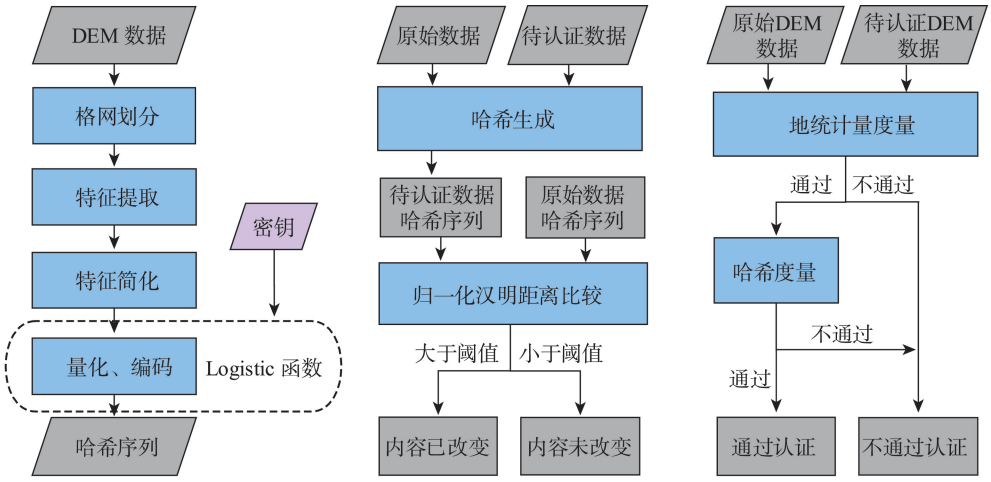
然而，DEM数据亦有其自身特性，针对DEM数据的感知哈希算法设计，需注意以下问题：① 微小篡改识别能力：感知哈希是用尽量小的数据量描述数据整体特征，且多针对全局进行特征提取。如果将现有感知哈希算法直接应用于DEM数据，可能无法识别其中的微小篡改，针对DEM数据的感知哈希算法应能实现对微小篡改的识别与定位。② 像元值精度要求：对普通数字图像而言，即使经历对比度变化、亮度变化等操作，只要不修改图像的内容信息，依然可以将其视为内容真实的数据，感知哈希对这种不改变内容的操作具有鲁棒性。然而DEM数据的像元值（即高程值）变化可能会直接影响地貌分析、水文分析等的正确性，所以认证过程需顾及DEM数据像元值精度的要求，对高程值的变化有更为严格的认证机制。

本文基于上述2点要求进行了针对性设计。针对传统图像感知哈希对微小篡改不敏感的问题，采用规则格网划分的方式对数据进行预处理。这样既可以提升特征提取的细节丰富度，又可以将不通过认证的区域定位到格网单元上，即实现篡改定位。针对DEM数据像元值精度的要求，提出基于高程中误差检测的认证方法，即计算原始数据与待认证数据的高程相对中误差，并结合判定阈值实现判断该区域是否发生变化。算法的具体流程如图1所示，下文将对算法流程进行详细说明。

2.2 哈希生成

2.2.1 DEM数据预处理

针对DEM数据图幅尺寸通常较大的问题，将DEM数据划分为规则而不重合的等大小格网区



(a) 哈希生成 (b) 哈希度量 (c) 数据认证

图1 DEM数据认证感知哈希算法

Fig. 1 Perceptual hash algorithm for DEM data authentication

域,可以解决原有感知哈希算法对微小篡改不敏感的问题,并可将篡改范围可被定位于格网范围内。然而格网单元尺寸的选取应依算法使用者的实际需求而确定,如在起伏度计算[24]、坡度坡向分析[25]、水文分析[26]中,其空间分析窗口大小均有不同,故相应格网大小的选取亦应依实际情况考虑,以适应不同的保护等级。更高的格网划分将带来更强的细节特征识别能力,却也会增加计算复杂度,越精细的格网划分便会带来越精细的篡改定位精度,在极端情况下1个像元便是一个格网单元,但这已无实际意义。故本文综合考虑计算效率、篡改定位精度、感知哈希算法特点等因素,将格网单元大小$m\times m$的默认取值设为m=64。算法的使用者在选择格网大小时,亦可选用本文建议格网的2^a($a\in Z$)倍作为格网单元尺寸(如32×32、128×128等),以适应不同的保护级别。这里的保护级别,也可以理解为使用者希望实现的篡改定位精度。

结合前述格网单元尺寸,对原始DEM数据D进行先逐行再逐列的规则格网划分。对于边界处不足m个像素单位的区域,以0高程值进行填充。这样,原始数据D将被划分为$W\times H$的格网单元,划分后的区域记为D_{wh}(w=1, 2, …, W; h=1, 2, …, H),其中w与h为格网单元所在的位置。随后,运用插值重采样算法将各格网单元D_{wh}的分辨率调整为$n\times n$(本实验中,n=32),其目的是保证最终生成的感知哈希序列长度一致,并可以降低随机噪声等的影响。

2.2.2 单元特征提取与简化

内容特征提取是感知哈希算法的核心。本文采用离散余弦变换的方式进行特征提取,对于$M\times N$的矩阵,其离散余弦变换的公式如式(1)所示:

$$F(u,v)=c(u)c(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)\times\cos\frac{\pi(2x+1)u}{2M}\cos\frac{\pi(2y+1)v}{2N} \tag{1}$$

其中,

$$c(u)=\begin{cases}\sqrt{1/M} & u=0\\ \sqrt{2/M} & u=1,2,\cdots,M-1\end{cases} \tag{2}$$

$$c(v)=\begin{cases}\sqrt{1/N} & v=0\\ \sqrt{2/N} & v=1,2,\cdots,N-1\end{cases} \tag{3}$$

式中:x, u=1, 2, …, M; y, v=1, 2, …, N; f(x, y)为格网单元数据在点(x, y)处的像素值(即高程值);$F(u,v)$为DCT变换后的系数。

将各格网单元的数据进行DCT变换后,可发现其左上角的低频系数(即AC系数)区域积聚了大量的能量,可以很好地反映DEM数据的纹理特征。而高频系数部分的值多为0或趋近0,这部分区域存在大量的内容冗余。因此,仅需取图像左上角少量低频系数作为特征矩阵。对于大小为32×32的重采样格网单元D_{wh},选取其左上角8×8区域内的频域系数作为内容特征矩阵F_{wh},即可很好地反映图像的内容特征[27]。如图2所示,为某格网单元离散余弦变换后的结果。

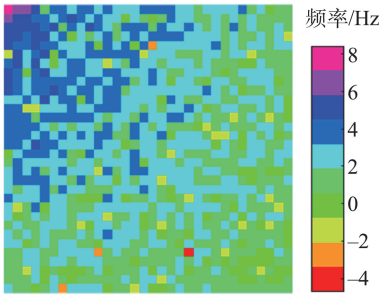


图2　格网单元数据离散余弦变换结果

Fig. 2　Discrete cosine transformation result of grid unit

2.2.3 特征矩阵置乱

DEM数据作为地形信息的连续模拟，不同区域在特征上存在相似性[28]。尤其对于格网划分后的小区域DEM数据，不同格网间特征相似的概率将大大增加。考虑一种应用场景，即恶意修改者运用特征极为相似的DEM块对某处进行替换，则普通哈希算法可能无法识别这种变化。Logistic混沌系统作为一种非线性的动力系统，其对参数与初值极其敏感[29]，常用于矩阵置乱处理。运用Logistic混沌系统对矩阵F_{wh}进行置乱处理后，矩阵将更具随机性与不可逆性，对数据微小的修改亦足以引起置乱矩阵的大幅变化。Logistic混沌系统方程如式(4)所示。

$$Xn+1=\mu Xn(1-Xn)\quad n=1,2,3,\cdots \tag{4}$$

式中：参数$\mu\in(0,4]$，$Xn\in(0,1)$，当$3.5699\cdots<\mu\leqslant 4$时，系统处于混沌状态。其中参数$\mu$即是本算法的密钥，在数据认证阶段，待认证数据应用相同的密钥进行置乱处理。

2.2.4 置乱矩阵的摘要化处理

对于置乱后的特征矩阵，需进行进一步量化、编码以得到最终的感知哈希序列。首先，计算特征矩阵F_{wh}的平均值，记为θ_{wh}。而后对矩阵进行二值化处理，记矩阵中的元素为F_{wh}^{xy}（$x=1,2,\cdots,8$；$y=1,2,\cdots,8$），哈希赋值规则如式(5)所示。

$$Hash(F_{wh}^{xy})=\begin{cases}1 & F_{wh}^{xy}\geqslant\theta_{wh}\\ 0 & F_{wh}^{xy}<\theta_{wh}\end{cases} \tag{5}$$

式中：$Hash\,(F_{wh}^{xy})$为元素(F_{wh}^{xy})对应的哈希码，该步骤将生成一个8×8的二值矩阵。随后将该二值矩阵按逐行拉伸的方式，平铺为一个64位的一维二值矩阵，该矩阵可看作一个64位的二进制数。随后将其转化为16位十六进制数，这便是格网单元D_{wh}的最终感知哈希序列，记为PH_{wh}。原始图像的感知哈希序列PH，为各格网单元感知哈希序列的串联。即$PH=\{PH_{11},PH_{12},\cdots,PH_{wh},\cdots,PH_{WH}\}$。

2.3 数据认证

图像感知哈希与DEM感知哈希的核心差别在于，DEM数据对高程值（即像元值）的修改具有更高的敏感性。例如，将普通数字图像的像元值提升10（即亮度提升3.92%），其内容依然可被判定为真实。然而，考虑DEM数据的应用中常涉及镶嵌操作，如将其中一块DEM数据的高程值提升10 m，将使得空间分析的结果受到质疑（如水文分析对高程的敏感性），故经受这种像元值等量变化的DEM数据应被定义为不真实。而感知哈希的技术特点决定了其对这种像元值的微小变化具有较强的耐受度，不可直接用于DEM数据的真实性认证。鉴于此，本文提出基于高程中误差检测与哈希比较相结合的认证方法。以下认证过程均在格网单元上进行。

2.3.1 高程相对中误差检测

首先，运用高程相对中误差[30]来验证高程的变化量，该步骤计算的是原始DEM数据与待验证DEM数据的相对高程中误差。由国家DEM精度标准[31]可知，不同地貌的DEM数据有不同的中误差要求，由于一份DEM数据中可能包括多种地貌特征，故本算法采用标准中精度要求最高的“平地”误差标准（T_1=0.35 m）作为度量标准。若原始数据与待认证数据的相对高程中误差R超过限差T_1，则直接将该格网单元标记为篡改，不再进行哈希度量；若通过高程中误差检测，则进入哈希度量，即：

$$\begin{cases}\text{进入哈希度量} & R_{wh}\leqslant T_1\\ \text{标为遭到篡改} & R_{wh}>T_1\end{cases} \tag{6}$$

式中：w与h为格网单元的位置。

2.3.2 哈希度量

对于通过相对高程中误差检测的格网单元，用哈希度量的方式对DEM数据的感知一致性进行认证，只有同时通过高程中误差检测与哈希度量的格网单元，才可以通过认证。对于待认证DEM数据，采用相同的哈希生成方法产生各格网单元感知哈希序列，记为PH'（$PH'=\{PH'_{11},PH'_{12},\cdots,PH'_{wh},\cdots,PH'_{WH}\}$）。通过对$PH$与$PH'$之间的差异性度量来进行数据差异性度量，从而实现认证。哈希序列的差异性用归一化汉明距离（Hamming Distance）进行判定，其定义如式(7)所示。

$$\mathrm{HammingDist}(P1, P2) = \frac{1}{N}\sum_{i=1}^{N}\left|P1_i - P2_i\right| \quad (7)$$

式中：$P1$和$P2$为2个长度为N的哈希序列。感知哈希序列汉明距离为0~1之间的浮点数。如果汉明距离小于设定的阈值T_2，则说明格网单元区域内的内容没有发生明显改变，则该格网内的信息通过认证；反之，则说明相应区域的内容发生的较大的变化，即可认定为数据受到了修改或篡改，则该格网认证不通过。结合图像感知哈希算法中的成功经验，并经过实验验证，汉明距离的阈值T_2=0.25时能对精确性与鲁棒性进行较好的平衡。需要指出的是，若有一个格网不通过认证，即可认定数据已不具备“完整性”。然而顾及篡改定位的要求，需对全部格网单元逐一认证后，方可判定篡改的具体影响范围，这在具体应用中具有实用价值。

3 实验与分析

3.1 数据概况与预处理

为验证本算法的广泛有效性，本文选用如图3所示的2份DEM数据进行实验与分析，这两份数据均包括了沟壑、平地、丘陵、山脉等地貌，能较全面地反映DEM数据的各类空间特征。其中，图3(a)为一块中国西部ALOS DEM数据，空间分辨率12.5 m，大小为1500像元×1500像元，经纬度范围为100°4′16″E—100°17′18″E、38°16′6″N—38°26′6″N；图3(b)为一块中国西南部SRTM DEM数据，空间分辨率为90 m，大小为2000像元×2000像元，经纬度范围为98°10′0″E—99°50′0″E、31°40′0″—33°20′0″N运用前述格网划分原则对两份数据分别进行格网划分：将图3(a)数据划分为24×24的格网区域，将图3(b)划分为32×32的格网区域，结果如图3(c)与图3(d)所示。

3.2 鲁棒性测试

对内容保持操作的鲁棒性，以及对改变内容操作的敏感性，是本算法的主要特征。根据对DEM数据实际应用场景的分析，本文定义了3种不改变内容或不影响使用的场景：① 水印嵌入；② 格式转为img；③ 格式转为dat。其中，水印嵌入采用最低有效位(Least Significant Bit, LSB)算法进行测试，该种嵌入算法对高程值的影响极小，可将其定义为不改变内容的操作。img格式与dat格式是常见的栅格地理数据格式，常用于存储DEM数据。本测试结果如表1所示，可以发现本算法对内容保持的操作(水印嵌入、格式转换)具有较强的鲁棒性。

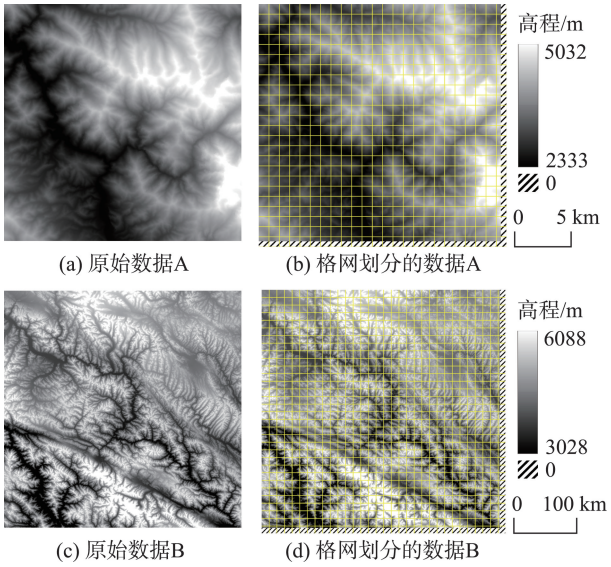


图3 原始实验数据及格网划分数据

Fig. 3 Original experimental data and grid partitioned data

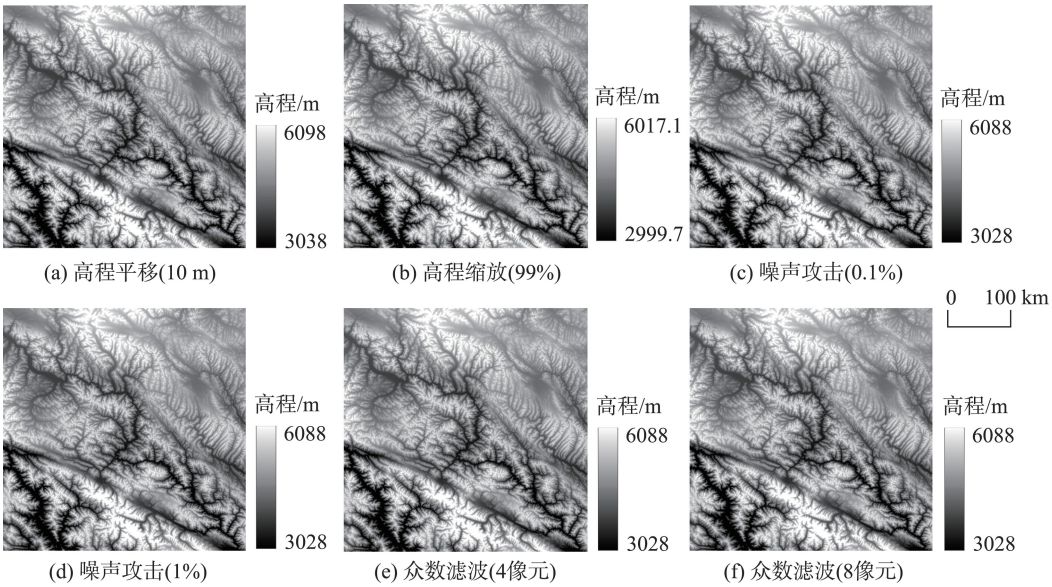
3.3 篡改检测能力测试

本算法能检测出对DEM数据内容的修改或篡改。在认证过程中，只要有一个格网单元的高程中误差检测或哈希度量未通过认证，即可认定数据已丧失完整性，数据即不能通过认证。通过对DEM应用场景的分析，本文定义的改变内容的操作为：① 高程平移；② 高程缩放；③ 噪声攻击；④ 众值滤波；⑤ 部分替换等。本文以数据B为例展示各类攻击方式的渲染图，结果如图4所示。

表1 鲁棒性测试统计结果

Tab. 1 Statistical results of robustness tests

攻击类型	数据A				数据B			
	相对中误差检测	哈希度量	认证结果	准确性	相对中误差检测	哈希度量	认证结果	准确性
水印嵌入	通过	通过	通过认证	准确	通过	通过	通过认证	准确
格式转为img	通过	通过	通过认证	准确	通过	通过	通过认证	准确
格式转为dat	通过	通过	通过认证	准确	通过	通过	通过认证	准确



(a) 高程平移(10 m)　(b) 高程缩放(99%)　(c) 噪声攻击(0.1%)

(d) 噪声攻击(1%)　(e) 众数滤波(4像元)　(f) 众数滤波(8像元)

图4　DEM数据不同攻击类型渲染图

Fig. 4　Rendering graphs of different DEM data attack types

其中，高程平移为将数据的高程值全部增加10 m；高程缩放为将高程值全部缩放至99%；噪声攻击为随机修改部分高程值，本文中测试了随机修改0.1%、1%比例的高程值，噪声范围为(−20 m，20 m)；众值滤波是一种常用的DEM数据平滑处理方法[32]，当半数及以上像元具有相同值时，根据相邻像元数据值的众数替换栅格中的像元，本文测试了4像元(滤波器内核为与当前像元正交的4个像元)与8像元(滤波器内核为与当前像元相邻的8个像元)滤波处理；部分替换即将某块内容替换为其他内容。

表2为对数据A与数据B进行各类篡改攻击后的认证结果。可以看出本算法对各种攻击的识别均具有有效性。其中，针对高程平移、高程缩放、噪声攻击(1%)、众值滤波处理，在高程中误差检测阶段即没有通过认证，故直接认定为数据不通过认证。针对数据A的噪声攻击、针对数据A与数据B的众值滤波攻击虽然通过了中误差检测，然而在哈希度量阶段，部分格网单元的内容感知被检测到发生了变化，故亦没有通过认证。由此可见，该算法对改变内容的攻击具有广泛有效性。

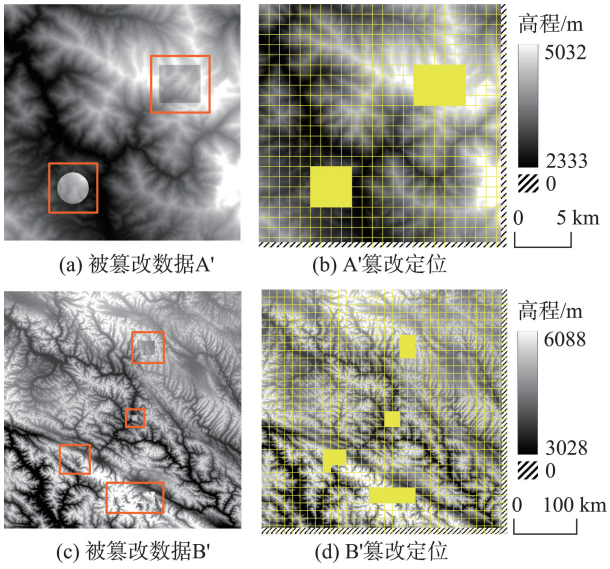
3.4　篡改定位能力测试

本算法在实现数据认证的同时，亦可实现篡改定位。由于篇幅有限，本文仅以“部分替换”攻击来测试篡改定位的能力(图5)，其他攻击方式的定位方法与其相同，且部分替换攻击能更直观地展示被攻击部分与未被攻击部分，亦可对定位结果的准确

表2　篡改检测统计结果

Tab. 2　Statistical results of tamper detection

攻击类型	数据A				数据B			
	中误差检测	哈希度量	认证结果	准确性	中误差检测	哈希度量	认证结果	准确性
不做修改	通过	通过	通过	准确	通过	通过	通过	准确
高程平移(10 m)	不通过	—	不通过	准确	不通过	—	不通过	准确
高程缩放(99%)	不通过	—	不通过	准确	不通过	—	不通过	准确
噪声攻击(0.1%)	通过	不通过	不通过	准确	不通过	—	不通过	准确
噪声攻击(1%)	不通过	—	不通过	准确	不通过	—	不通过	准确
众值滤波(4像元)	通过	不通过	不通过	准确	通过	不通过	不通过	准确
众值滤波(8像元)	通过	不通过	不通过	准确	通过	不通过	不通过	准确
部分替换	不通过	—	不通过	准确	不通过	—	不通过	准确



(a) 被篡改数据A' (b) A'篡改定位

(c) 被篡改数据B' (d) B'篡改定位

图5 篡改定位测试

Fig. 5 Tamper localization tests

性做直观的判断。如图5(a)与图5(b)所示,为对图3(a)与图3(b)所示的数据A与数据B施加替换攻击后的篡改数据A'与B'。图5(b)为对数据B的4处内容进行了替换攻击。在认证过程中,没有同时通过相对中误差检测与哈希度量的格网单元都将被标记为篡改,认证结束也就相应标记了所有被篡改的格网单元。如图5(c)与图5(b)所示为数据A'与B'的篡改定位结果,可以发现被篡改的区域均被有效识别。

为说明哈希检测在篡改检测中的操作中的有效性,表3中给出了篡改区与非篡改区的平均汉明距离。可以发现,篡改区的平均汉明距离远高于算法设置的阈值T_2=0.25,而未篡改区的平均汉明距离远低于该阈值,这也就说明本算法设置的汉明距离阈值具有较强的适用性。

表3 汉明距离统计结果

Tab. 3 Statistical results of the Hamming distance

数据名称	格网类型	平均汉明距离
A'	篡改格网	0.836
	未篡改格网	0.010
B'	篡改格网	0.854
	未篡改格网	0.009

4 结论

本文提出了一种基于感知哈希的DEM数据认证算法,对内容保持的操作具有鲁棒性,对内容改变的操作具有敏感性,本文的主要贡献如下:

(1) 相较于数字签名与脆弱水印等精确性完整性认证方法,本算法将"内容"的一致性作为完整性的唯一判断标准,故能耐受格式转换、水印嵌入等不改变内容的操作,却对改变内容的操作有极强的敏感性,克服了原有认证算法对数据载体过分依赖而较少考虑内容是否一致的缺点;

(2) 引入了格网划分的方法,弥补了传统感知哈希算法高鲁棒性而对微小篡改识别不敏感的问题,一方面顾及了DEM数据细节特征丰富的特点,能实现更为精确的认证,另一方面又可以实现格网单元级别的篡改定位;

(3) 作为感知哈希在DEM数据上的针对性应用,提出了相对高程中误差度量与哈希度量相结合的认证方法,数据只有满足精度标准并感知一致的情况下才能通过认证,更符合DEM数据使用中对精度的实际需求。

然而,针对DEM数据的感知哈希算法是仍需继续挖掘的课题,有许多需要继续研究的问题。一方面,感知哈希感知一致性的判定阈值T_1的确定,仍需要大量的实验进行验证,这也是下一步研究的重点;另一方面,本文没有考虑在几何畸变条件下DEM数据的认证。由于几何畸变将带来更多的失真,对其真实性的判定也有待于进一步的研究。

参考文献(References):

[1] 汤国安.我国数字高程模型与数字地形分析研究进展[J].地理学报,2014,69(9):1305-1325. [Tang G A. Progress of DEM and digital terrain analysis in China[J]. Acta Geographica Sinica, 2014,69(9):1305-1325.]

[2] 程春泉,黄国满,杨杰. POS与DEM辅助机载SAR多普勒参数估计[J].测绘学报,2015,44(5):510-517. [Feng C Q, Huang G M, Yang J. Doppler centroid estimation for airborne SAR supported by POS and DEM[J]. Acta Geodaetica et Cartographica Sinica, 2015,44(5):510-517.]

[3] 冯增文.基于多期DEM的地质灾害与环境动态监测[D].北京:中国地质大学(北京),2015. [Feng Z W. Multiphase DSM for the dynamic monitoring of geological hazard and environment-a case study of Miyun[D]. Beijing: China University of Geosciences (Beijing), 2015.]

[4] 王雪,李精忠,余斌.基于DEM提取流域特征影响因子的分析[J].测绘与空间地理信息,2019,42(6):38-42. [Wang X, Li J Z, Yu B. Analysis of influence factors for extraction of hydrological basin features from DEM[J]. Geomatics & Spatial Information Technology, 2019,42(6):38-42.]

[5] 陈冬花,邹陈,王苏颖等.基于DEM的伊犁河谷气温空间插值研究[J].光谱学与光谱分析,2011,31(7):1925-1929. [Chen D H, Zou C, Wang S Y, et al. Study on spatial interpolation of the average temperature in the Yili river valley based on DEM[J]. Spectroscopy and Spectral Analysis, 2011,31(7):1925-1929.]

[6] 黄琪.基于DEM与LP的土地平整工程设计优化研究[D].南京:南京农业大学,2011. [Huang Q. Study on optional design for land leveling project based on DEM and LP[D]. Nanjing: Nanjing Agricultural University, 2011.]

[7] 葛韬.某区域虚拟战场环境参谋系统的设计[D].大连:大连理工大学,2017. [Ge T. Design of a regional virtual battlefield environment support system[D]. Dalian: Dalian University of Technology, 2017.]

[8] 丁勇.密码学与信息安全简明教程[M].北京:电子工业出版社,2015. [Ding Y. Cryptography and information security concise tutorial[M]. Beijing: Publishing House of Electronics Industry, 2015.]

[9] 辛运帏,廖大春,卢桂章.单向散列函数的原理、实现和在密码学中的应用[J].计算机应用研究,2002,19(2):25-27. [Xin Y W, Liao D C, Lu G Z. The principle and implement of one-way hash functions and their cryptographic application[J]. Application Research of Computers, 2002, 19(2):25-27.]

[10] 朱长青,杨成松,任娜.论数字水印技术在地理空间数据安全中的应用[J].测绘通报,2010(10):1-3. [Zhu C Q, Yang C S, Ren N. Application of digital watermarking to geospatial data security[J]. Bulletin of Surveying and Mapping, 2010(10):1-3.]

[11] Feistel H. Cryptography and computer privacy[J]. Scientific American, 1973,228(5):15-23.

[12] 许惠.基于数字水印技术的地理数据篡改责任认定方法研究——以DEM数据为例[D].南京:南京师范大学,2015. [Xu H. Tampering responsibility identification method based on digital watermarking for geographical data: A case study of DEM data[D]. Nanjing: Nanjing Normal University, 2015.]

[13] 王刚,任娜,朱长青,景旻.倾斜摄影三维模型数字水印算法[J].地球信息科学学报,2018,20(6):738-743. [Wang G, Ren N, Zhu C Q, et al. The digital watermarking algorithm for 3D models of oblique photography[J]. Journal of Geo-information Science, 2018,20(6):738-743.]

[14] 魏征,闫浩文,张黎明.抗高程平移和裁剪的格网DEM盲水印算法[J].测绘科学,2016,41(8):170-173. [Wei Z, Yan H W, Zhang L M. A blind watermarking algorithm for grid DEM resisting elevation translation and cropping attacks[J]. Science of Surveying and Mapping, 2016,41(8): 170-173.]

[15] 刘爱利,丁浒,田丹,等.基于坡度和坡向分析的DCT域DEM数字水印算法[J].武汉大学学报·信息科学版,2016,41(7):903-910. [Liu A L, Ding H, Tian D, et al. A DCT DEM digital watermarking algorithm based on the analysis of slope and aspect[J]. Geomatics and Information Science of Wuhan University, 2016,41(7):903-910.]

[16] 朱长青,许惠,任娜.顾及地形特征的DEM脆弱水印完整性认证算法[J].地球信息科学学报,2016,18(3):369-375. [Zhu C Q, Xu H, Ren N. A fragile watermarking algorithm for integrity authentication of DEM based on the terrain feature[J]. Journal of Geo- information Science, 2016,18(3):369-375.]

[17] Haitsma J, Kalker T, Oostveen J. Robust audio hashing for content identification[C]. International Workshop on Content-Based Multimedia Indexing, University of Brescia. 2001, 4:117-124.

[18] 牛夏牧,焦玉华.感知哈希综述[J].电子学报,2008,36(7): 1405-1411. [Niu X M, Jiao Y H. An overview of perceptual hashing[J]. Acta Electronica Sinica, 2008,36(7):1405-1411.]

[19] Ruchay A, Kober V, Evtushenko E. Fast perceptual image hash based on cascade algorithm[C]. Applications of Digital Image Processing XL. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2017, 10396.

[20] Wang H, Wang H X. Perceptual hashing- based image Copy-Move forgery detection[J]. Security and Communication Networks, 2018,1-11.

[21] 张春艳,李京兵,王双双.基于离散小波变换和感知哈希的加密医学图像检索算法[J].计算机应用,2018,38(2): 539-544. [Zhang C Y, Li J B, Wang S S. Encrypted image retrieval algorithm based on discrete wavelet transform and perceptual hash[J]. Journal of Computer Applications, 2018,38(2):539-544.]

[22] Wang X, Pang K, Zhou X, et al. A visual model-based perceptual image hash for content authentication[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(7):1336-1349.

[23] Ding K, Meng F, Liu Y, et al. Perceptual hashing based forensics scheme for the integrity authentication of high resolution remote sensing image[J]. Information, 2018,9(9): 1-17.

[24] 南希,李爱农,景金城.中国山地起伏度计算中地形自适应滑动窗口获取与验证[J].地理与地理信息科学,2017,33(4): 34-39. [Nan X, Li A N, Jing J C. Calculation and verification of topography adaptive slide windows for the relief amplitude solution in mountain areas of China[J]. Geogra-

phy and Geo-information Science, 2017,33(4):34-39.]

[25] 陈艳丽,李少梅,刘岱岳.基于规则格网DEM的坡度坡向分析研究[J].测绘与空间地理信息,2009,32(5):36-39. [Chen Y L, Li S M, Liu D Y. Analysis of slope and aspect based on regular grid DEM[J]. Geomatics & Spatial Information Technology, 2009,32(5):36-39.]

[26] 王银堂,李伶杰,胡庆芳等.考虑局部趋势的非一致性水文频率分析方法[J].水科学进展,2017,28(3):406-414. [Wang Y T, Li L J, Hu Q F, et al. Nonstationary hydrologic frequency analysis method considering local trends[J]. Advances in Water Science, 2017,28(3):406-414.]

[28] 晋蓓,刘学军,王彦芳.格网DEM坡度计算模型的相似性研究[J].测绘科学,2009,34(6):131-134. [Jin B, Liu X J, Wang Y F. Similarity assessment on slope models used in grid-based DEM[J]. Science of Surveying and Mapping, 2009,34(6):131-134.]

[29] 李赵红,侯建军.基于Logistic混沌映射的DCT域脆弱数字水印算法[J].电子学报, 2006,34(12):2134-2137. [Li Z H, Hou J J. DCT-Domain fragile watermarking algorithm based on logistic maps[J]. Acta Electronica Sinica, 2006, 34(12):2134-2137.]

[30] 武汉大学测绘学院测量平差学科组.误差理论与测量平差基础[M].武汉:武汉大学出版社,2009. [Subject Group of Surveying Adjustment, School of Surveying and Mapping, Wuhan University. The base of errors theory & surveying adjustment[M]. Wuhan: Wuhan University Press, 2009.]

[31] 国家测绘局.中华人民共和国测绘行业标准(基础地理信息数字产品 1:10 000、1:50 000 数字高程模型)[Z].2001. [State Bureau of Surveying and Mapping. Standards for surveying and mapping industry of the People's Republic of China (basic geographic information digital products 1:10 000, 1:50 000 DEM)[Z]. 2001.]

[32] Jordan G, Meijninger B M L, Hinsbergen D J J V , et al. Extraction of morphotectonic features from DEMs: Development and applications for study areas in hungary and NW Greece[J]. International Journal of Applied Earth Observations & Geoinformation, 2005,7(3):163-182.