algorithms

MDPI

Article

A Novel Perceptual Hash Algorithm for Multispectral Image Authentication

Kaimeng Ding [1,2], Shiping Chen [1,3,*] and Fan Meng [2]

1 School of Networks and Tele-Communications Engineering, Jinling Institute of Technology, Nanjing 211169, China; dkm@jit.edu.cn
2 State Key Laboratory of Resource and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Science, Beijing 100101, China; mengf@lreis.ac.cn
3 Commonwealth Scientific and Industrial Research Organization (CSIRO), Data61, Sydney, NSW 1710, Australia
* Correspondence: Shiping.Chen@data61.csiro.au; Tel.: +61-2-9372-4663

Received: 21 December 2017; Accepted: 8 January 2018; Published: 14 January 2018

Abstract: The perceptual hash algorithm is a technique to authenticate the integrity of images. While a few scholars have worked on mono-spectral image perceptual hashing, there is limited research on multispectral image perceptual hashing. In this paper, we propose a perceptual hash algorithm for the content authentication of a multispectral remote sensing image based on the synthetic characteristics of each band: firstly, the multispectral remote sensing image is preprocessed with band clustering and grid partition; secondly, the edge feature of the band subsets is extracted by band fusion-based edge feature extraction; thirdly, the perceptual feature of the same region of the band subsets is compressed and normalized to generate the perceptual hash value. The authentication procedure is achieved via the normalized Hamming distance between the perceptual hash value of the recomputed perceptual hash value and the original hash value. The experiments indicated that our proposed algorithm is robust compared to content-preserved operations and it efficiently authenticates the integrity of multispectral remote sensing images.

Keywords: multispectral remote sensing image; perceptual hash; integrity authentication; affinity propagation; feature fusion

1. Introduction

Due to the rapid growth of remote sensing, multispectral (MS) remote sensing images have exhibited increasing potential for more and more applications ranging from independent land mapping services to government and military activities. However, with the development of image processing and network transmission techniques, it has become easier to tamper with or forge multispectral remote sensing images during the process of processing and transmission. For example, the widespread use of sophisticated image editing tools can make the authenticity and integrity of MS images suffer from a serious threat, even rendering the multispectral images useless. Therefore, ensuring the content integrity of the MS image is a major issue before the multispectral image can be used. A perceptual hash algorithm, also known as a robustness hash algorithm, is able to solve the problems of MS image content authentication, while the classic cryptographic authentication algorithms, such as MD5 and SHA1, are not suitable for this purpose since they are sensitive to each bit of the input image.

A perceptual hash algorithm maps an input image into a compactible feature vector called perceptual hash value, which is a short summary of an image's perceptual content. Perceptual hash algorithms have been developed as a frontier research topic in the field of digital media content security, and they can be applied for image content authentication, image retrieval, image registration, and digital watermarking. Similar to cryptographic hash functions, the perceptual hash algorithm

compresses the representation of the perceptual features of an image to generate a fixed-length sequence, which ensures that perceptually similar images produce similar sequences [1].

In recent years, a number of perceptual hash algorithms have been proposed to meet the requirements of different types of data [2–15]. Ahmed et al. [2] propose a perceptual hash algorithm for image authentication, which uses a secret key to randomly modulate image pixels to create a transformed feature space. It offers good robustness and it can detect minute tampering with localization of the tampered area, but it is not able be applied to an MS image with many more bands. Hadmi et al. [3] propose a novel perceptual image hash algorithm based on block truncation coding. Sun et al. [4] develop a perceptual hash based on compressive sensing and Fourier–Mellin transformation. The proposed method is robust to a wide range of distortions and attacks, and it yields better identification performances under geometric attacks such as rotation attacks and brightness changes. Yan et al. [5] propose a multi-scale image hashing method by using the location-context information of the features generated by adaptive and local feature extraction techniques. Cui et al. [6] propose a hash algorithm for 3D images by selecting suitable Dual-tree complex wavelet transform coefficients to form the final hash sequence. Qin et al. [7] propose a perceptual hash algorithm for images based on salient structure features, which can be applied in image authentication and retrieval. Chen et al. [8] propose a perceptual audio hash algorithm based on maximum-likelihood watermarking detection, which can be applied in a content-based search. Yang et al. [9] propose a wave atom transform (WAT) based image hash algorithm using distributed source coding to reduce the size of hash code, providing a better performance than existing WAT. Tang et al. [10] propose a perceptual hash algorithm with innovative use of discrete cosine transform and local linear embedding, which can be used in image authentication, image retrieval, copy detection and image forensics. Sun et al. [11] propose a video hash model based on a deep belief network, which generates the video hash from visual attention features. Chen et al. [13] propose a Discrete Cosine Transform (DCT) based perceptual hash scheme to track vehicle candidates and achieve a high level of robustness. Qin et al. [14] exploit the circle-based and the block-based strategies to extract the structural features. Fang et al. [15] adopt a gradient-based perceptual hash algorithm to encode invariant macro structures of person images to make the representation robust to both illumination and viewpoint changes.

However, only a few researches on the perceptual hashing for multispectral remote sensing image have been carried out. The existing perceptual hash algorithms do not take the characteristics of MS images into consideration and they are not suitable for MS images. Therefore, new perceptual hash algorithms need to be introduced to solve the problems of MS image authentication.

Different from MS remote sensing images, panchromatic (PAN) remote sensing images with the characteristics of high spatial resolution and low spectral resolution can be authenticated by the existing perceptual hash algorithms for normal images, as the digital form of PAN images is the same as normal images. In contrast, an MS remote sensing image is characterized by lower spatial resolution than PAN, but higher spectral resolution, which makes the existing perceptual hash algorithms for imaging unsuitable for multispectral remote sensing image authentication.

An MS image obtains information from the visible and the near-infrared spectrum of reflected light; it is composed of a set of (more than three) monochrome images of the same scene, each of which is referred to as a band and is taken at a specific wavelength. Whereas the normal color image is composed of only three monochrome images, the PAN image has only one band. The bands of the MS remote sensing image represent the earth's surface in different spectral sampling intervals and have clear physical meanings, while the existing perceptual hash algorithm does not take this into account and cannot perceive the content of each band. Moreover, multispectral images are generally of large sizes (some may be over several GB), while most existing perceptual hash algorithms compute the hash value from an image's global features and are generally not sensitive to local modification in the multispectral remote sensing images.

This paper addresses the above problems by presenting a novel perceptual hash algorithm for multispectral remote sensing image authentication. In this paper, we made the following contributions:

(a) In the proposed hash algorithm, we adopt affinity propagation (AP) clustering to separate the bands into several band subsets to reduce redundancy, in which mutual information (MI) is chosen to measure the similarity of the bands. Therefore, the input image of our algorithm can be of an arbitrary band number to avoid setting the number of band clusters.

(b) Based on the analysis of data characteristics of the MS image and the basic principles of perceptual hash, we adopt the band fusion technique to extract the principle features and obtain compact hash values for each band subset, which overcomes the deficiencies in existing perceptual hash algorithms for mono-spectral images.

(c) We introduce grid entropy-based adaptive weighted fusion rules to obtain comprehensive features of the grid in the same geographic region. This helps improve the preservation of detailed information as much as possible.

The remainder of this paper is organized as follows. Section 2 gives a brief introduction to perceptual hash algorithms and discusses the related work. Section 3 describes our proposed algorithm in detail. Section 4 presents our experimental results and analysis. Finally, we draw conclusions in Section 5.

2. Preliminaries

2.1. Overview of Perceptual Hash

As shown in Figure 1, perceptual hash algorithms generally consist of the following stages: preprocessing, feature extraction, feature quantification, hash generation.

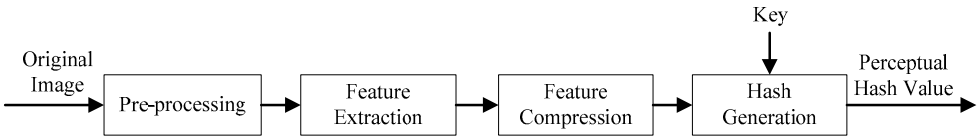


Figure 1. Framework of perceptual hash algorithms.

Pre-processing generally removes redundant information in an image, making it easier to extract features from the image later on. Feature extraction is to extract the principle features of the image using a specific extraction method. Feature compression is to fuzzy up the extracted features in order to enhance robustness. Hash generation is to make the quantitative characteristics more abstract, because the quantitative features may lead to a large amount of data being suitable as an output sequence. For image authentication, a perceptual hash algorithm should possess the following properties:

1. Compactness: The hash value of the image should be as compact as possible, so that it is easier to transport, store and use.
2. Sensitivity to tampering: Visually distinct images should have significantly different hash values.
3. Perceptual robustness: The hash generation should be insensitive to a certain level of distortion with respect to the input image.
4. Security: Calculation of image hashing depends on a secret key, that is, different secret keys can generate significantly different hash values for the same image.

For the authentication of an MS image, the perceptual hash algorithm has to be sensitive to malicious tampering operations and to be robust to content-preserving ones. Compared with normal images, MS images have higher requirements for measurement accuracy, and their pixels with coordinates can be used for measuring geometrical locations after the image has been corrected and processed, which means the authentication algorithm should have high authentication precision and be able to detect micro-tampering of the image.

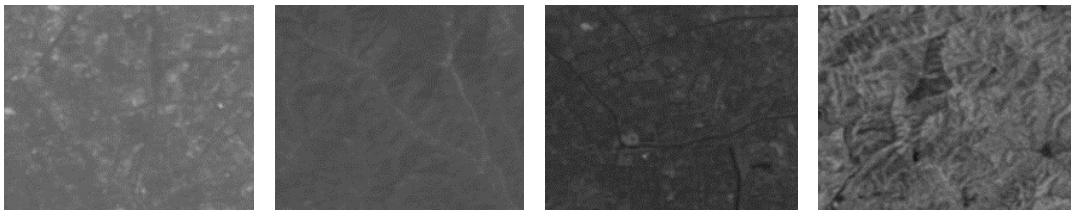
The simplest way to generate multispectral image perceptual hashing is to generate the hash value for each band. However, this would lead to a huge data volume of hash values, while the perceptual hash values should be as compact as possible to be convenient for data authentication. In addition, there are some correlations between each band, which result in high redundancy and a great amount of computation time wasted in hash generation.

In this paper, we separate the bands into several band subsets with a clustering algorithm to reduce the content redundancy, and then we adopt a band fusion-based feature extraction technique to obtain the compact feature of each band subset, which could be suitable for hashing computation. Furthermore, grid division is adopted to divide each band into grids and make the hash value more sensitive to local modification in the MS images.

On the other hand, feature extraction is a key stage of the perceptual hash. For remote sensing image authentication, edge characteristics based on perceptual hash can achieve higher precision [16]. The sensing images would have little value if the edge characteristics had been greatly changed. Additionally, edge characteristics contain effective information for applications such as object extraction, image segmentation and target recognition. Therefore, we adopt edge features as the perceptual feature to generate perceptual hash value.

2.2. *Affinity Propagation and Mutual Information-Based Band-Clustering*

Different bands of a multispectral image have different spectral responses and can be distinguished from each other based on their grayscales. Even in the same image, there are big grayscale differences between the bands, especially between the visible band and infrared band. Figure 2 shows several comparison instances of different type regions (mountain and urban) of the Landsat thematic mapper (TM) image data of Band 1 (InR_1) and Band 4 (InR_4). Obviously, there are some differences between mountainous areas and the urban areas. Furthermore, even for the same surface feature, the difference between different bands is obvious, as shown in Figure 2b,d.



(**a**) (**b**) (**c**) (**d**)

Figure 2. Ground comparison in thematic mapper (TM) imagery: (**a**) City in band 1; (**b**) Coteau area in band 1; (**c**) City in band 4; (**d**) Coteau area in band 4.

Taking this into account, we adopt band clustering to separate the bands into several groups (band subsets). Band-clustering is used to identify the subset of bands that are as independent as possible, and it is adopted to obtain compactable hash values in this algorithm.

In this paper, we use AP clustering [17,18] with MI to divide the bands into several band subsets. The AP algorithm is an exemplar-based clustering algorithm that uses max-sum belief propagation to obtain an optimal exemplar which can determine the number of clusters automatically, with a message exchange approach. The final clustering centers will be generated depending on the given dataset. Compared with traditional clustering algorithms, such as K-means clustering, AP clustering does not need to set the clustering number. Therefore, it is an efficient clustering technique to deal with datasets of many instances because of its better clustering performance over traditional methods.

The AP clustering algorithm starts with the construction of a similarity matrix $S \in R^{L\times L}$, in which the element $s(i,j)(i \neq k)$ measures the similarity between band k (InR_k) and band i (InR_i). In our perceptual hash algorithm, MI is chosen to construct the similarity matrix, i.e., each $s(i, j)$ denotes the mutual information between the i-th and the j-th while L is determined by the number of bands.

MI measures the statistical dependence between two random variables and can therefore be used to evaluate the relative utility of each band to classification [19]. Given two random variables X and Y with marginal probability distributions $p(X)$ and $p(X)$ and joint probability distribution $p(X, Y)$, their MI is defined as:

$$I(X,Y) = \sum_X \sum_Y p(X,Y) \log \frac{p(X,Y)}{P(X)p(Y)} \quad (1)$$

It follows that MI is related to entropy as:

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \quad (2)$$

where $H(X)$ and $H(Y)$ are respectively the entropies of X and Y, and $H(X, Y)$ is their joint entropy.

Treating the band's multispectral images as random variables, MI can be used to estimate the dependency between them, and was introduced for band selection in [20,21]. Using Equation (2), the mutual information between each band of the multispectral image can be calculated, which can be used to evaluate the relative utility of each band to classification.

2.3. Band Fusion-Based Edge Feature Extraction

As mentioned above, edge characteristics-based perceptual hash can achieve higher precision for MS image integrity authentication, so we adopt the band feature fusion technique in order to obtain the robust edge feature of the band sets separated by band-clustering.

So far, many fusion algorithms have been developed for multispectral band fusion. One of the most important fusion techniques is the wavelet-based method, which usually uses the discrete wavelet transform (DWT) in the fusion [22]. Since the DWT of image signals produces a non-redundant image representation, it can provide better spatial and spectral localization of image information as compared to other multiresolution representations. Therefore, we adopt DWT-based fusion techniques to obtain the robust fusion features of the band subset.

The key step of DWT-based fusion techniques is to define a fusion rule to create a new composite multiresolution representation, which uses different fusion rules to deal with various frequency bands. The process of applying the DWT can be represented as a set of filters. After first-level decomposition, the image is decomposed into a low frequency sub-image (LL_1) which is the approximation of the original image and a group of high-frequency sub-images (LH_1, HH_1, and HL_1) which contain abundant edge information, as shown in Figure 3a. The low frequency sub-image is the approximation of the original signals for this level and can be further decomposed until the desired resolution is reached. Figure 3b illustrates wavelet decomposition of an image at level 2. In Figure 3b, LL_2, HL_2, LH_2, and HH_2 are the sub-images produced after the LL_1 sub-image that are being further decomposed. Since the second-level high-pass sub-images (also called middle frequency sub-images) LH_2, HH_2, and HL_2 also contain abundant edge information, and are more robust than high frequency ones, but more fragile than low frequency ones, we use middle-frequency sub-images to extract the robust bits.

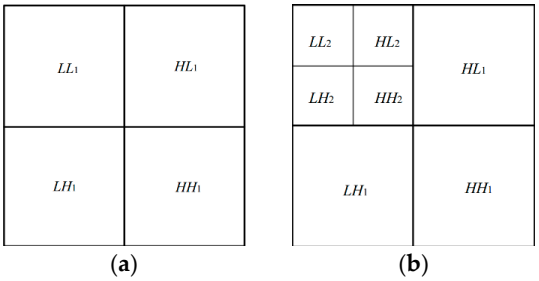


Figure 3. Two-dimensional discrete wavelet transform of 2D signals: (**a**) One-level discrete wavelet transform (DWT); (**b**) Two-level DWT.

3. Our Perceptual Hash Algorithm Design

The proposed perceptual hash algorithm for MS sensing image authentication consists of three main stages: (a) bands clustering, (b) feature extraction, and (c) hash value generation. The schematic diagram of the proposed algorithm is given in Figure 4.

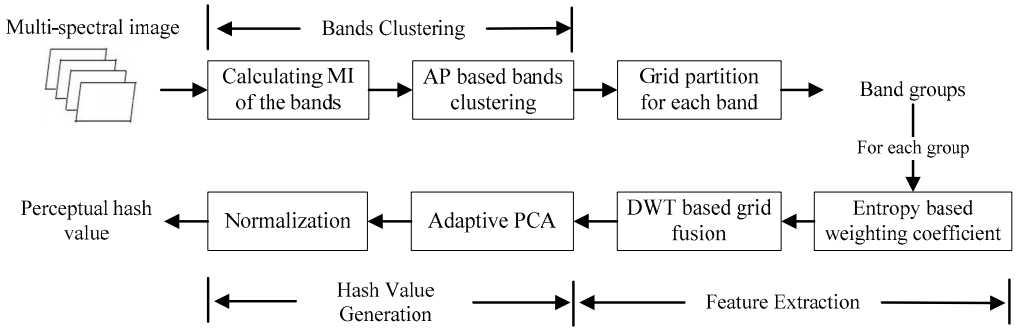


Figure 4. Framework of the proposed perceptual hash algorithm.

3.1. Band-Clustering

We employ AP clustering to divide the original MS image I into several band groups, and we do not need to set the clustering number. Firstly, the similarity matrix S is constructed, where each matrix element is the mutual information between InR_k and InR_i and the matrix size is determined by the number of bands. Then, the messages of similarity matrix S are updated repeatedly until some stop after a fixed number of iterations, at which point we need to set the damping factor λ and the maximum number of iterations in advance.

After the step of band-clustering, the bands are divided into band clusters (band subsets). That is to say, the L bands are divided into N clusters, and each cluster is denoted as IC_n ($n < N$). For each band cluster IC_n:

$$IC_n = \left\{ IC_n^1, IC_n^2, ...IC_n^i \right\} \tag{3}$$

where $IC_n^i \in I$; therefore, the original MS image can also be expressed as:

$$I = \{IC_1, IC_2, ..., IC_N\} \tag{4}$$

3.2. Band Fusion-Based Edge Feature Extraction

The band fusion-based feature extraction is intended for encoding the perceptual information from source bands into a single one containing the best aspects of the original bands, which includes two steps: grid division and feature extraction. The details of band fusion-based edge feature extraction on the band subsets are described as follows.

3.2.1. Grid Division

To make the hash value more sensitive to local modification in the MS images, each band is partitioned into $M \times N$ grids, and the grid is denoted as G_{wh}^k in which $w = 1, 2, \ldots, M$, $h = 1, 2, ..., N$, and k denote the band number. Thus, the grids in different bands of the same position composed a grid set which is denoted G_{wh}, as follows.

$$G_{wh} = \{G_{wh}^1, G_{wh}^2, \ldots, G_{wh}^n\} \tag{5}$$

As the tamper location ability is based on the resolution of the grid division, the higher the resolution of the grid division, the more fine-grained the authentication granularity can be. While the computational cost would be raised at higher resolutions, we need to segment the remote sensing

image into more grids in order to increase both the time for computing perceptual hash values, as well as comparing the values. The choice of the grid division resolution thus presents a trade-off between the cost and tamper location ability. Our work aims at designing such an authentication model with good balance between cost and performance.

3.2.2. Feature Extraction

Selecting the fusion rules is of great importance; it directly affects the fusion quality and the sensitivity of the perceptual hash algorithm. For each grid set G_{wh}, the features are extracted and fused based on discrete wavelet transform (DWT) with adaptive weighted rules.

The whole process can be divided into the following steps:

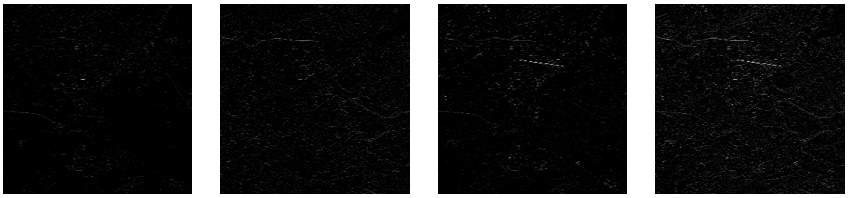
1. To obtain the fixed length of hash values, each grid G_{wh}^k is firstly conducted with the normalization of bilinear interpolation to resize with the size of $m \times m$.
2. Two-level DWT is applied to each resized grid, which is decomposed into different kinds of coefficients at different scales. To extract more robust detailed information for generating hash values, we choose the two-level high-pass coefficients LH_2, HH_2, and HL_2 as the basic perceptual feature.
3. For each grid, the three sub-bands LH_2, HH_2, and HL_2 are fused into one matrix though the fusion rule of maximum, and the fusion result is expressed as a matrix denoted by M_{wh}^k.
4. For each band cluster IC_n, the adaptive weighted fusion is made on the matrix M_{wh}^k of the grid in different bands, and the fusion matrix is denoted FM_{wh}. The fusion process should satisfy the followed conditions, i.e., the preservation of as much detailed information as possible. To do this, the fusion coefficient of each matrix is as follows:

$$\alpha_k = \frac{E_{wh}^k}{E_{wh}^{Total}} = \frac{E_{wh}^k}{\sum_{i=1}^{n} E_{wh}^i} \tag{6}$$

where α_k is the weighted coefficient of the matrix M_{wh}^k of the grid in the kth band, and E_{wh}^k is the entropy of the corresponding grid. Obviously, α_k depends on the entropy of the grid and is different from other areas. Then, the pixel of the matrix FM_{wh} can be computed as follows:

$$FM_{wh}(i,j) = \sum_{k=1}^{n} \alpha_k M_{wh}^k(i,j) \tag{7}$$

A fusion example of sub-band HL_2 is given in Figure 5, where Figure 5a–c show the intermediate frequency components of the grids in InR_3, InR_4 and InR_7 respectively, and Figure 5d shows the fusion result. It is observed that the fusion result retains the obvious edge features of the grid in several bands.



(**a**) (**b**) (**c**) (**d**)

Figure 5. A fusion example of the sub-band: (**a**) the intermediate frequency components of the grid in InR_3; (**b**) the intermediate frequency components of the grid in InR_4; (**c**) the intermediate frequency components of the grid in InR_7; (**d**) the fusion result.

3.3. Data Compaction Based on Self-Adaptive PCA

Since the hash value has to be as compact as possible in order to keep the content-preserving ones robust, we apply principal component analysis (PCA) on the fused feature fusion matrix FM_{wh} for noise reduction and data compaction. The PCA algorithm reduces the large dimensionality of image data in order to reduce the dimensionality of independent feature space. It is widely used in data reduction and data compression, as it is able to discover the relationships among the variables [23,24]. By using PCA on matrix FM_{wh}, the linear correlation of the matrix element can be removed. This means that the noise can be effectively removed and the extracted feature achieves data compression.

The grid's principal components are then standardized in order to obtain the fixed-length string which is then encrypted by using a cryptographic encryption algorithm that takes RC4 as an example to enhance the security. The encrypted string is the hash value of the grid denoted as $PH^k_{w,j}$.

All of the grid's perceptual hash values $PH^k_{w,j}$ are put together as the hash value of the clustering, denoted as PH^k:

$$PH^k = PH_{0,0}||PH_{0,1}||\ldots||PH_{w,h} \tag{8}$$

Finally, the hash value of the original multispectral is denoted PH as follows.

$$PH = \left\{PH^1, PH^2, ..., PH^N\right\} \tag{9}$$

where N is the number of band clusters. Clearly, the hash length depends on band clusters and grid division.

3.4. Integrity Authentication

At the receiver, the authentication process is implemented via the comparison between the reconstructed perceptual hash value and the original one: the higher the perceptual hash values' difference, the greater the corresponding images' difference. Although the hamming distance is frequently used to evaluate the difference between two sequences, it is not suitable for this purpose, because the length of the hash value may vary along with the change of algorithm parameters. We have adopted the followed "Normalized Hamming Distance" [25] to evaluate the difference between two hash values:

$$Dis = \left(\sum_{i=1}^{L}|h1(i) - h2(i)|\right)/L \tag{10}$$

where h_1 and h_2 are perceptual hash values with L length. It is observed that the normalized hamming distance Dis is a float between 0 and 1. If the Dis of two perceptual hash values of the same area is lower than the threshold T_h, it means that the corresponding area is content-preserving; otherwise, it means that the content of the corresponding area has been tampered with.

Furthermore, the tampering can be located in the corresponding geographic regions by comparing each hash value of each grid. The higher the resolution of the grid division, the more fine-grained the authentication granularity can be. To obtain higher resolutions, we need to divide the image into more grids, compute more unit perceptual hash values, and compare more hash values.

4. Experiments and Discussion

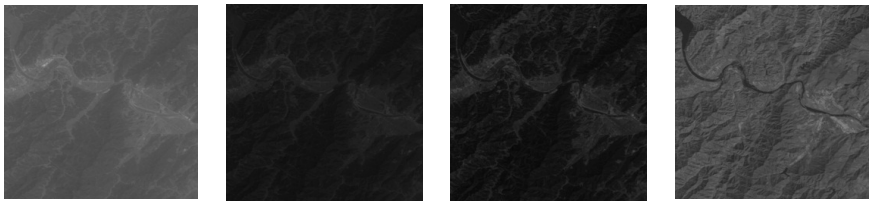
In this section, we evaluate the robustness of our proposed perceptual hash algorithms from two aspects. The first one is the robustness against content-preserving manipulations, wherein perceptually identical images under distortions would have similar hashes, which is important for content-based image identification. The other is the sensitivity to tampering of the image, whereby the image that has been tampered with would have different hashes to the corresponding original image.

All experiments were implemented on a computer with a 2.40 GHz Intel i7 processor and 4.00 GB memory running Windows 10 operating system. The test software was developed using Microsoft Visual Studio 2013 in C++.

4.1. Perceptual Hash Values Generation

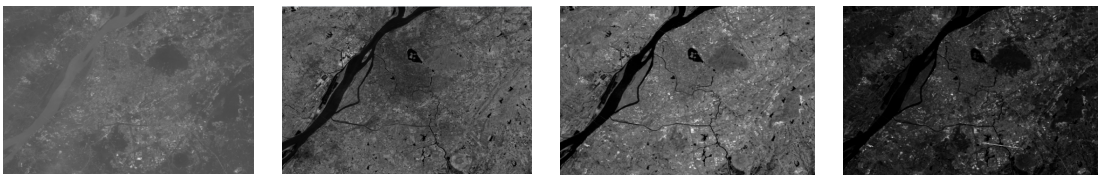
There are several parameters in the perceptual hash algorithm, and we describe the parameter settings used in our experimental results in the following. Referring to the existing research [17,18], we set the damping factor λ as 0.5 and the maximum number of iterations as 100 during the band-clustering. The clustering process need not set the clustering number. During the grid division process, the size of non-overlapping grids is 128 $\times$ 128 pixels.

We opted for the Landsat TM image as an example to validate robustness performances and tamper sensitivity. Figures 6 and 7 present several bands of the typical TM images to show the comparison between the bands of the same area, in which the content of the same geographical location in each band is quite different. The details of the mutual information between the bands of the above two images are given in Tables 1 and 2.



(**a**) (**b**) (**c**) (**d**)

Figure 6. Different bands of Image A for testing: (**a**) band 1; (**b**) band 4; (**c**) band 5; (**d**) band 7.



(**a**) (**b**) (**c**) (**d**)

Figure 7. Different bands of Image B for testing: (**a**) band 1; (**b**) band 4; (**c**) band 5; (**d**) band 7.

Table 1. The mutual information between the bands of the TM image A.

Band	InR_1	InR_2	InR_3	InR_4	InR_5	InR_6	InR_7
InR_1	3.2901	1.4594	1.0248	0.3750	0.4729	0.3153	0.5264
InR_2	1.4594	2.8759	1.3769	0.5142	0.6558	0.4224	0.6977
InR_3	1.0248	1.3769	3.3143	0.4844	0.8609	0.5246	1.0225
InR_4	1.3750	0.5142	0.4844	3.8516	0.8623	0.4968	0.5814
InR_5	0.4729	0.6558	0.8609	0.8623	4.5497	0.6991	1.5869
InR_6	0.3153	0.4224	0.5246	0.4968	0.6991	2.8549	0.6597
InR_7	0.5264	0.6977	1.0225	0.5814	1.5869	0.6597	3.9207

Table 2. The mutual information between the bands of the TM image B.

Band	InR_1	InR_2	InR_3	InR_4	InR_5	InR_6	InR_7
InR_1	3.8032	1.3787	1.2635	0.1976	0.2785	0.2564	0.7383
InR_2	1.3787	3.6052	1.6015	0.1938	0.3104	0.2085	0.7552
InR_3	1.2635	1.6015	4.1446	0.2561	0.3137	0.2488	0.8458
InR_4	0.1976	0.1938	0.2561	4.5152	0.5029	0.1878	0.4335
InR_5	0.2785	0.3104	0.3137	0.5029	4.6314	0.1555	0.9323
InR_6	0.2564	0.2085	0.2488	0.1878	0.1555	3.1426	0.3227
InR_7	0.7383	0.7552	0.8458	0.4335	0.9323	0.3227	4.3498

The results of band clustering for each TM image are the same as below: $IC_1 = \{I^1, I^2, I^3\}$, $IC_3 = \{I^6\}$, $IC_4 = \{I^5, I^7\}$. Thus, InR_1, InR_2 and InR_3 would be divided into a band group and be generated as the perceptual hash value. Similarly, InR_5 and InR_7 would be generated as the perceptual hash value as a group.

4.2. Performance of Perceptual Robustness

Perceptual robustness is the most significant difference between perceptual hash and cryptography hash. An ideal perceptual hash algorithm should be resistant to commonly-used remote sensing image content-preserving manipulations, which means that the normalized Hamming distance between the two hash values of the original image and the processed one by the content-preserving manipulation should be under the pre-determined threshold T. In this paper, the threshold T is set as 0.05.

In order to evaluate the performance of perceptual robustness for the algorithm, we utilized data compaction and digital watermark embedding as examples of content-preserving manipulations for testing, in which data compaction involves lossy compression (90% JPEG compression) and lossless compression, and digital watermark embedding adopts least significant bit (LSB) embedding. In this paper, to describe the perceptual robustness, we adopt the percentage of the grid's hash values in which no changes occurred that exceed the threshold T; the results are shown in Table 3. It can be seen from Table 3 that this algorithm can maintain its robustness with respect to the lossless compression of multi-spectral images and LSB watermark embedding, and can maintain near-robustness to lossy compression.

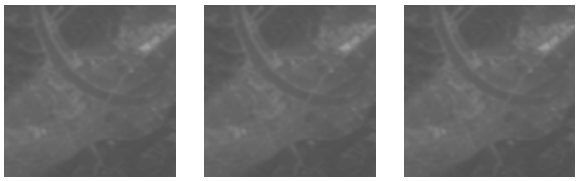
Table 3. The results of the robustness test.

Manipulation	Lossless Compression	Digital Watermarking	JPEG Compression (90%)
Image A (12 × 8 partition)	100%	100%	87.5%
Image B (4 × 4 partition)	100%	100%	93.75%

The robustness of the algorithm can be adjusted by the pre-determined threshold T, i.e., the greater the threshold T is, the stronger the robustness of the algorithm. However, the relationship between robustness and sensitivity to tampering is contradictory, and may directly affect sensitivity to tampering if the robustness is overemphasized.

In contrast, cryptographic authentication methods cannot achieve better certification, since they treat the above manipulation as illegal operations and their hash value would be changed dramatically after content-preserving manipulations. We utilized SHA-256 (Secure Hash Algorithm 256) as an example of cryptographic algorithms to compare with our proposed algorithm. Figure 8 shows examples of content-preserving manipulations; Figure 8a shows the original grid and Figure 8b–c show the results of content-preserving manipulations. Obviously, there are almost no differences between each of them, while the hash values of SHA-256 are very different, as shown in Table 4. By contrast, our proposal keeps the content-preserving manipulations robust, and the perceptual hash values remain unchanged. On the other hand, as the conventional perceptual hash algorithms for images do not take

the multiband characteristic of MS images into account and cannot be applied directly to MS images, we do not compare them with our proposed algorithm.



(**a**) (**b**) (**c**)

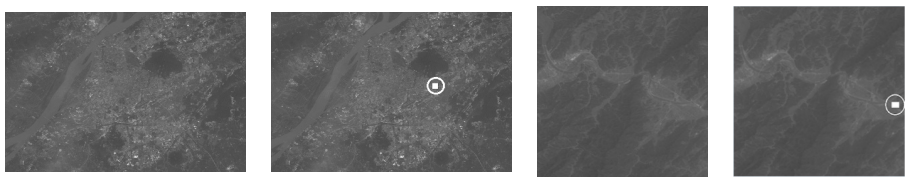
Figure 8. Examples of content-preserving manipulations: (**a**) Original grid; (**b**) After lossy compression; (**c**) After watermark embedded.

Table 4. Comparison of the hash values before and after content-preserving manipulations.

Hash Value	Original Hash Value	After Lossless Compression	After Watermark Embedded
Our proposed algorithm	f8f718effc12faee f21ef6faf005f4f3 eef2001c08fafe0a 5b0902fdf0f4f4fd	f8f718effc12faee f21ef6faf005f4f3 eef2001c08fafe0a 5b0902fdf0f4f4fd	f8f718effc12faee f21ef6faf005f4f3 eef2001c08fafe0a 5b0902fdf0f4f4fd
SHA-256	7ab724d6a78b5b70 1d005c1a13de6cdc 7b8a890c98204380 4df7abff96d38f8c	b45d4e47cbf18b2d 6427e04022c9bc4a 149397fb054156d2 dc2f965424c285ae	acbe3194602df9e2 5d7f1e560f472ab5 dfcb813d6c073fb3 e2fca9f5295e9f58

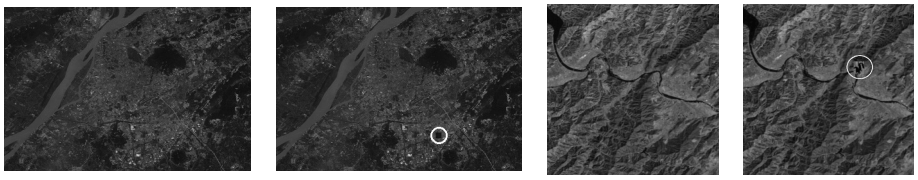
4.3. Performance of Sensitivity to Tampering

The authentication process of a multi-spectral image should be able to detect the local tampering of the bands, which means that the tampered and original images should have significantly different hashes. To test the performance of sensitivity with respect to tampering, we take several kinds of tampering operations, including removing, appending and changing the object, as shown in Figures 9–11. When the band of the image is tampered with, the regenerated perceptual hash values of the grid and the whole image would be changed, and the malicious tampering can be detected. Take the tampering of InR_1 of the original image A as an example, as shown in Figure 9a,b, the perceptual hash values of the image and grid cell would be changed tremendously, as shown in Table 5.



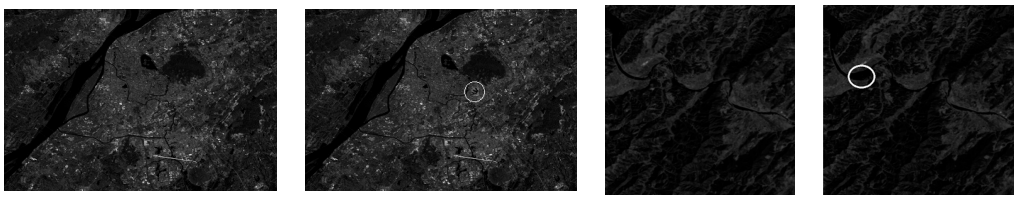
(**a**) (**b**) (**c**) (**d**)

Figure 9. Tampering Test 1 (removing the object): (**a**) InR_1 of the original image A; (**b**) InR_1 of the tampered image A; (**c**) InR_1 of the original image B; (**d**) InR_1 of the tampered image B.



(**a**) (**b**) (**c**) (**d**)

Figure 10. Tampering Test 2 (appending the object): (**a**) InR_3 of the original image A; (**b**) InR_3 of the tampered image A; (**c**) InR_5 of the original image B; (**d**) InR_5 of the tampered image B.



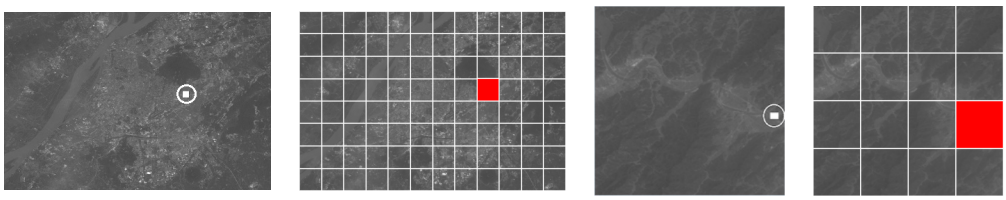
(**a**) (**b**) (**c**) (**d**)

Figure 11. Tampering Test 3 (changing the object): (**a**) InR_7 of the original image A; (**b**) InR_7 of the tampered image A; (**c**) InR_7 of the original image B; (**d**) InR_7 of the tampered image B.

Table 5. The normalized hamming distance of the tampered grid.

Compared Images	Figure 8a,b	Figure 8c,d
The normalized hamming distance	0.1016	0.1484

For the above tampering example, the comparison of the hash values of each grid can be used to locate the tampering with respect to the corresponding geographic region, and the location granularity will depend on the resolution of the grid divisions. Figure 12 shows the tamper location of Test 1 as an example.



(**a**) (**b**) (**c**) (**d**)

Figure 12. Tamper location of Test 1: (**a**) InR_1 of the tampered image A; (**b**) Tamper location of Image A; (**c**) InR_1 of the tempered image B; (**d**) Tamper location of Image B.

4.4. Security Analysis

As described in Section 3, the performance of the security of the hash values is dependent on the security of the cryptographic encryption algorithm. The security of the chosen RC4 in this paper is widely researched and recognized, so that the security of our algorithm is guaranteed.

5. Conclusions

In this paper, we have proposed a perceptual hash algorithm for multispectral remote sensing image authentication. In order to compactly represent the perceptual features of the multispectral image, we have adopted an affinity propagation algorithm to classify the MS images into several clusters based on the mutual information of the bands of these images. Dividing each band into a grid, the features of the grid cell at the same location within the cluster are extracted and fused based on DWT, while PCA-based data compression on the fused feature helps reduce the influence of noise. The final perceptual hash value can be acquired after the compressed feature has been encrypted by the cryptographic encryption algorithm. Experimental results have shown that the proposed algorithm is robust against normal content-preserving manipulations, such as data compaction and digital watermark embedding, and has good sensitivity to detect local detailed tampering of the multispectral image. Thus, the algorithm efficiently authenticates the content integrity of multispectral remote sensing images, and overcomes the defects of the existing perceptual hash algorithms which do not consider multispectral remote sensing images.

The aims of future works are outlined as follows: firstly, to improve the robustness against JPEG compression; secondly, to expand the algorithm to include hyperspectral remote sensing images which contain many more bands than multispectral images.

Acknowledgments: This study is supported by grants from the State Key Laboratory of Resource and Environment Information System Open Funding Program (Grant No. 201604), the Jiangsu Province Science and Technology Support Program (Grant No. BK20170116), the National Natural Science Foundation of China (Grant Nos. 41601396, 41501431) and the Scientific Research Hatch Fund of Jinling Institute of Technology (Grant Nos. jit-fhxm-201604, jit-fhxm-201605, jit-b-201520).

Author Contributions: The manuscript was written by Kaimeng Ding under the guidance of Shiping Chen. Fan Meng participated in designing the experiments. All authors reviewed the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Niu, X.M.; Jiao, Y.H. An Overview of Perceptual Hashing. *Acta Electron. Sin.* **2008**, *36*, 1405–1411.
2. Ahmed, F.; Siyal, M.Y.; Abbas, V.U. A secure and robust hash-based algorithm for image authentication. *Signal Process.* **2010**, *90*, 1456–1470. [CrossRef]
3. Hadmi, A.; Puech, W.; Said, B.A.E. A robust and secure perceptual hashing system based on a quantization step analysis. *Signal Process. Image Commun.* **2013**, *28*, 929–948. [CrossRef]
4. Sun, R.; Zeng, W. Secure and robust image hashing via compressive sensing. *Multimed. Tools Appl.* **2014**, *70*, 1651–1665. [CrossRef]
5. Yan, C.P.; Pun, C.M.; Yuan, X.C. Multi-scale image hashing using adaptive local feature extraction for robust tampering detection. *Signal Process.* **2016**, *121*, 1–16. [CrossRef]
6. Cui, C.; Mao, H.; Niu, X.; Zhang, L.X.; Hayat, T.; Alsaedi, A. A novel hashing algorithm for Depth-image-based-rendering 3D images. *Neurocomputing* **2016**, *191*, 1–11. [CrossRef]
7. Qin, C.; Chen, X.; Dong, J.; Zhang, X.P. Perceptual Image Hashing with Selective Sampling for Salient Structure Features. *Displays* **2016**, *45*, 26–37. [CrossRef]
8. Chen, N.; Xiao, H.D. Perceptual audio hashing algorithm based on Zernike moment and maximum-likelihood watermark detection. *Digit. Signal Process.* **2013**, *23*, 1216–1227. [CrossRef]
9. Yang, Y.; Zhou, J.; Duan, F.; Liu, F.; Cheng, L.M. Wave atom transform based image hashing using distributed source coding. *J. Inf. Secur. Appl.* **2016**, *31*, 75–82. [CrossRef]
10. Tang, Z.; Lao, H.; Zhang, X.; Liu, K. Robust image hashing via DCT and LLE. *Comput. Secur.* **2016**, *62*, 133–148. [CrossRef]
11. Sun, J.; Liu, X.; Wan, W.; Li, J.; Zhao, D.; Zhang, H.X. Video hashing based on appearance and attention features fusion via DBN. *Neurocomputing* **2016**, *213*, 84–94. [CrossRef]
12. Rivas, A.; Chamoso, P.; Martín-Limorti, J.J.; Bajo, J. Image Matching Algorithm Based on Hashes Extraction. In Proceedings of the Portuguese Conference on Artificial Intelligence, Porto, Portugal, 5–8 September 2017; pp. 87–94.
13. Chen, L.; Hu, X.; Xu, T.; Kuang, H.L. Turn Signal Detection during Nighttime by CNN Detector and Perceptual Hashing Tracking. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 3303–3314. [CrossRef]
14. Qin, C.; Sun, M.H.; Chang, C.C. Perceptual Hashing for Color Images Based on Hybrid Extraction of Structural Features. *Signal Process.* **2018**, *142*, 194–205. [CrossRef]
15. Fang, W.; Hu, H.M.; Hu, Z.; Liao, S.C.; Li, B. Perceptual Hash-based Feature Description for Person Re-identification. *Neurocomputing* **2018**, *272*, 520–531. [CrossRef]
16. Ding, K.M.; Zhu, C.Q. Perceptual hash algorithm for integrity authentication of remote sensing image. *J. Southeast Univ.* **2014**, *44*, 723–727.
17. Kokawa, Y.; Wu, H.; Chen, Q. Improved Affinity Propagation for Gesture Recognition. *Procedia Comput. Sci.* **2013**, *22*, 983–990. [CrossRef]
18. Hang, W.; Chung, F.L.; Wang, S. Transfer affinity propagation-based clustering. *Inf. Sci.* **2016**, *348*, 337–356. [CrossRef]
19. Gong, M.; Zhao, S.; Jiao, L.; Tian, D.; Wang, S. A Novel Coarse-to-Fine Algorithm for Automatic Image Registration Based on SIFT and Mutual Information. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4328–4338. [CrossRef]

20. Guo, B.; Gunn, S.R.; Damper, R.I.; Nelson, J.D.B. Band Selection for Hyperspectral Image Classification Using Mutual Information. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 522–526. [CrossRef]
21. Khelifi, R.; Adel, M.; Bourennane, S. Segmentation of multispectral images based on band selection by including texture and mutual information. *Biomed. Signal Process. Control* **2015**, *20*, 16–23. [CrossRef]
22. Amolins, K.; Zhang, Y.; Dare, P. Wavelet based image fusion techniques—An introduction, review and comparison. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 249–263. [CrossRef]
23. Zhou, F.; Ju, H.P.; Liu, Y. Differential feature based hierarchical PCA fault detection method for dynamic fault. *Neurocomputing* **2016**, *202*, 27–35. [CrossRef]
24. Xu, C.; Gao, S.; Li, M. A novel PCA-based microstructure descriptor for heterogeneous material design. *Comput. Mater. Sci.* **2017**, *130*, 39–49. [CrossRef]
25. Ding, K.; Zhu, C.; Lu, F. An adaptive grid partition based perceptual hash algorithm for remote sensing image authentication. *Wuhan Daxue Xuebao* **2015**, *40*, 716–720.