Future Generation Computer Systems 90 (2019) 222–239

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

One-pass lossless data hiding and compression of remote sensing data[☆]

Bruno Carpentieri, Arcangelo Castiglione, Alfredo De Santis, Francesco Palmieri, Raffaele Pizzolante[*]

Dipartimento di Informatica, Università degli Studi di Salerno, Via Giovanni Paolo II, 132, I-84084, Fisciano (SA), Italy

HIGHLIGHTS

- A one-pass framework for lossless compression and simultaneously reversible data hiding for remote sensing data.
- In a single pass a marked and compressed image is produced by the proposed framework.
- The resulting compressed marked image can be decompressed and restored (the exact input image is obtained) or decompressed only (a marked version of the input image is obtained).
- The marked image can be used for several purposes, in which it is not necessary to extract the original data and an acceptable grade of degradation is tolerated.

ARTICLE INFO

Article history:
Received 24 March 2018
Received in revised form 11 June 2018
Accepted 25 July 2018
Available online xxxx

Keywords:
Reversible data hiding
Reversible invisible watermark
Hyperspectral data hiding
Hyperspectral images protection
Data integrity
Authentication

ABSTRACT

The information obtained by means of spectral remote sensing (i.e., the hyperspectral images) are involved in several real-life scenarios and applications. Historical research, monitoring of environmental hazards, forensics and counter-terrorism are some examples of contexts in which the hyperspectral data play an important role.

In many contexts, the hyperspectral images could also play sensitive roles (e.g., in military applications, etc.) and are generally exchanged among several entities, in order to carry out different tasks on them. Therefore, it is important to guarantee their protection. A meaningful choice is the protection through *data hiding* techniques.

In fact, by means of reversible data hiding techniques, the imaging data become a sort of *information carrier* and can be used for delivering other important data that can be used, for instance, to check the integrity of the original imaging data.

In this paper, we introduce a one-pass framework that is able to perform the lossless data hiding and the lossless compression of the marked stream, at the same time, by exploiting the capabilities of the predictive paradigm. Substantially, in a single pass, a marked and compressed stego image is obtained, which can be exactly restored by the receiver: by decompressing and reversibly reconstructing the original unaltered image. In addition, our framework also permits to perform only the decompression (without the extraction of the hidden information). In this manner, the resulting stego (marked) hyperspectral image, could be used for several purposes, in which it is not necessary to extract the original data and an acceptable grade of degradation is tolerated. We also implement a *proof-of-concept* of the proposed framework to assess the effectiveness of our contribution. Finally, we report the achieved experimental results, which outperform other similar approaches.

1. Introduction

By means of remote sensing it is possible to acquire information about distant objects, without coming into physical contact with them, by exploiting the fact that an object reflects, absorbs and emits electromagnetic radiation according to its chemical composition [1–3]. Again, by analyzing the energy of the radiation as a function of the wavelength, it is possible to obtain a *spectral signature*. A spectral signature can be considered as a sort of unambiguous fingerprint that can be used to uniquely characterize any given material. For example, different organizations, such as NASA, etc., have cataloged the spectral signatures of various minerals, hence permitting an easier identification of such materials.

[☆] This document is a collaborative effort.

[*] Corresponding author.

E-mail addresses: bcarpentieri@unisa.it (B. Carpentieri), arcastiglione@unisa.it (A. Castiglione), ads@unisa.it (A. De Santis), fpalmieri@unisa.it (F. Palmieri), rpizzolante@unisa.it (R. Pizzolante).

https://doi.org/10.1016/j.future.2018.07.051
0167-739X/© 2018 Elsevier B.V. All rights reserved.

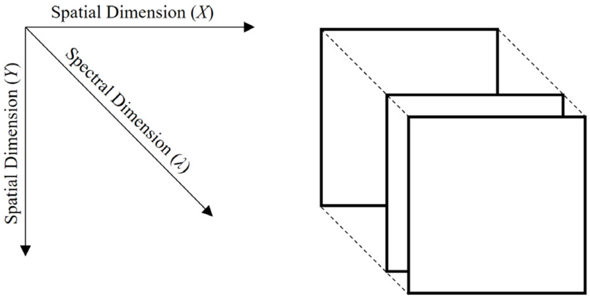

Fig. 1. Example of a hyperspectral datacube.

The information acquired through spectral remote sensing (i.e., the *hyperspectral images*) are used for several purposes and applications, varying from surveillance to historical research, archeology, environmental sciences, monitoring of environmental hazards, assessment of food quality, and several other real-life contexts, such as laboratory research, biomedical and Earth imaging, forensics, counter-terrorism, skin health checking, etc. Notice that such images can be air-borne and space-borne acquired.

In detail, a hyperspectral sensor acquires information by considering the portion of the electromagnetic spectrum that varies from the visible part (400–760 nanometers — nm) to near-infrared (about 2400 nm). More precisely, we can see a hyperspectral image as a *datacube*, since it is composed by a sequence of narrow contiguous *spectral bands*, composing a three-dimensional data, as shown in Fig. 1. In general, each band covers a bandwidth of about 10 nm and a hyperspectral image is composed by a few hundreds of bands [1,4,5].

For example, hyperspectral images produced by the *NASA AVIRIS (Airborne Visible/Infrared Imaging Spectrometer)* sensors are composed by 224 bands, and are acquired by revealing the frequencies of ultraviolet and infrared rays at wavelengths in the range between 400 and 2500 nm [1].

Motivations. Given the different multidisciplinary application contexts of hyperspectral images, in general, it is very likely that such images are exchanged between different entities, to effectively carry out a wide range of tasks, typically performed in cloud-based environments. Again, many of the application contexts of hyperspectral images are security-sensitive, so particular effort should be devoted in their protection. Although cryptography is commonly used as an effective tool for ensuring security and content authentication [6–9], however, the disguised nature of encrypted data may draw the attention of an adversary. In such a scenario, *data hiding* has been found to be an efficient and effective alternative to cryptography for ensuring security, content authentication and copyright protection [10–14]. In addition, by means of data hiding techniques, imaging data become a sort of *information carriers*, which can be used for delivering other important data [15]. An image for carrying data is called a *cover image*, whereas, the image carrying the embedded information is called a *stego image* or *marked image*. Data hiding on images presents two (virtually conflicting) requirements: the *embedding capacity*, i.e., the number of bits embedded into the host image and the *visual quality* of the embedded image. Notice that if during the transmission an image shows signs of hiding effect, an adversary may notice that the digital media carries secret messages. Therefore, the stego image should have imperceptible differences with respect to the cover image [16,17]. The goal of data hiding is to create schemes with high embedding capacity and good stego image quality. In particular, *reversible data hiding (RDH)* is a technique which not only enables the embedding of hidden data into cover images, but also restores original images from the stego images after the extraction of the embedded data [18–28]. However, the drawbacks of RDH with respect to non-reversible techniques are: lower payloads, larger distortion and higher computational cost.

Again, it is important to highlight that images collected through hyperspectral sensors are generally processed by automatic applications [29]. Thus, it is highly suggested to store, maintain and transmit such data into the original form. However, since hyperspectral sensors generate a significant amount of data, it can be necessary to compress such data. Moreover, though lossy compression algorithms achieve a better compression ratio than lossless ones, they introduce distortion and make the data usable only by applications where a certain level of distortion is allowed. Therefore, it is easy to note that in this scenario only the lossless compression of hyperspectral images permits to have a reduction in the amount of data, without any loss of useful information.

Contribution. In this work we propose a novel *one-pass* framework, which by relying on the capabilities of a predictive paradigm, enables at the same time the data hiding and compression operations on hyperspectral images. More precisely, it achieves, in a single pass, a marked and compressed stego image. On the other side, the receiver can extract the hidden information and then decompress and reversibly reconstruct the original unaltered image. Again, our framework also enables the receiver to perform the decompression without extracting the hidden information. The resulting stego (marked) hyperspectral image could be used for several purposes, e.g., in situations where it is not necessary to extract the original data and an acceptable grade of degradation is tolerated.

In the literature, several lossless compression algorithms for hyperspectral images, based on the predictive paradigm, have been proposed (e.g., the ones addressed in [30–32]). On the other hand, despite various watermarking methods for remote sensing data have been introduced in the state of the art (e.g., [33–35]), the field concerning the high capacity lossless data hiding has not been exhaustively explored so far. Indeed, to the best of our knowledge, our proposal represents the first one-pass framework specifically designed for hyperspectral images, which is able to perform the lossless data hiding and the lossless compression of the marked stream, at the same time. In addition, our framework is also suitable for *on-board* implementations (i.e., for implementations on a hardware-constrained sensor), since it is designed to require limited computational capability and memory space. More precisely, the proposed framework relies on a data hiding approach based on the *modification of prediction errors (MPE)* technique [36] and exploits the fact that, since the histograms in the domain of prediction errors are sharply distributed, the embedding capacity is higher than that of traditional *histogram-shifting methods* for the same image quality. The key idea behind our proposal is that we employ three-dimensional predictors, which are ad-hoc designed to exploit, in the best way possible, the redundancies of hyperspectral images, rather than using bi-dimensional predictors, as for example has been done in [37].

In order to assess the effectiveness of the proposed scheme, we carried out extensive performance evaluation experiments, which show a saving in terms of execution time, ranging from 20.08% to 47.73% with respect to more traditional solutions performing data hiding and then compression. Moreover, our framework enables a gain in terms of embedding capacity ranging from 5.48% to 170.8% with respect to state-of-the-art solutions. Again, by varying the size and the type of the embedded data payload, the two images, namely, the cover image and the produced stego image, are extremely similar and hence difficult to distinguish, thus making the detection of the hidden payload by a malicious user extremely difficult. Finally, when compared to the pure compression, i.e., the compression performed on images which do not carry any hidden payload, our proposal does not significantly affect the compression