Signal Processing 90 (2010) 2911–2922

Contents lists available at ScienceDirect

Signal Processing

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

A high capacity reversible data hiding scheme using orthogonal projection and prediction error modification

Wien Hong [a,*], Tung-Shou Chen [b], Yu-Ping Chang [c], Chih-Wei Shiu [d]

[a] Yu Da University, Department of Information Management, Miaoli 361, Taiwan
[b] National Taichung Institute of Technology, Department of Computer Science and Information Engineering, Taichung, 404, Taiwan
[c] Yu Da University, Department of Applied English, Miaoli, 361, Taiwan
[d] National Chung-Hsing University, Department of Computer Science and Engineering, Taichung, 402, Taiwan

ARTICLE INFO

Article history:
Received 8 September 2009
Received in revised form
23 January 2010
Accepted 15 April 2010
Available online 20 April 2010

Keywords:
Reversible data hiding
Histogram-shifting
Orthogonal projection

ABSTRACT

The payload of a histogram-shifting based reversible data hiding technique is primarily determined by the peak height of the corresponding image histogram. It is known that the peak height of the prediction error histogram is usually higher than that of image histogram itself. In 2009, Tsai et al. adopted the idea of histogram-shifting technique and embedded message by modifying the prediction errors to achieve a higher payload. They used a block-based approach and leaved some pixel values unmodified to predict other pixel values in a block. However, the unmodified pixels took up some embeddable space, causing a reduction in payload. In this paper, a new reversible data hiding technique based on prediction error modification is proposed. The orthogonal projection technique is employed to estimate the optimal weights of a linear predictor to increase the prediction accuracy and the embedding capacity. The experimental results revealed that the proposed method outperformed the existing histogram-shifting based reversible data hiding technique in terms of payload and image quality.

1. Introduction

Data hiding is a technique that embeds data into digital media to convey secret messages by slightly altering the content of the media, so that the embedded data is imperceptible [1,2]. Many proposed data hiding techniques [3–5] are non-reversible; namely, the embedded media are distorted and cannot be restored. If the embedded media can be recovered through a specially designed algorithm, then the data hiding technique is termed reversible data hiding. When a digital image is used to embed data, the image used to carry data is called the cover image, and the image with data embedded is called the stego image. Images in many applications often allow no distortion, such as images for military, medical or law enforcement uses. In these applications, reversible data hiding gives a solution to the distortion problem since the original cover image can be completely recovered [6].

Recently, a number of data hiding techniques [6–13] proposed in relevant studies have the capability to recover the original image to meet the increasing needs in this field. In 2006, Ni et al. [7] proposed a very different reversible data hiding technique based on the histogram-shifting technique. They adjusted pixel values between peak and zero points to conceal data and to achieve reversibility. However, the payload is limited by the most frequent pixel values in the cover image, resulting in fewer payloads compared to other reversible data hiding schemes. Besides, their method provides no mechanism to control the distortion–capacity relationship; therefore, equal amount of distortion is presented in both fully

* Corresponding author. Tel.: +886 37 651188/6129; fax: +886 37 652143.
E-mail addresses: wienhong@ydu.edu.tw, hongwien@ms53.hinet.net (W. Hong), tschen@ntit.edu.tw (T.-S. Chen), yptexas@webmail.ydu.edu.tw (Y.-P. Chang), chihwei.shiu@gmail.com (C.-W. Shiu).

0165-1684/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2010.04.012

embedded and partially embedded images, which is undesirable for cover images with small payload embedded.

In 2007, Thodi and Rodriguez [8] proposed a reversible data hiding technique by combining the histogram-shifting and expansion-embedding techniques [6,7]. They used a special designed location map to record the overflow/underflow information and had a better performance over Tian's [6] method. Although their method effectively improves the embedding capacity, the compressibility of the location map is undesirable for some image types. Hu et al. [9] solved their compressibility problem by constructing two types of sparser location maps: one from embedding and another from shifting. This feature significantly reduces the size of compressed location map, and thus, benefits the embedding capacity.

Later, Lin et al. [10] proposed another scheme and exploited the difference of pixel values to embed data. Although a large amount of data can be embedded in Lin et al.'s method, the key size of their method is considerably large and is not suitable for transmission over a limited bandwidth network.

In 2009, Kim et al. [11] proposed a novel reversible data hiding technique based on sub-sampled image. They sub-sampled the original image into several sub-images and the sub-image that minimizes the spatial correlation among the sub-images is served as a reference image. The difference values are then obtained by subtracting the reference image from other sub-images and secret data are embedded by modifying the histogram of these difference values. Their method greatly improved Lin et al.'s method in that overflow and underflow problem were subtly prevented during embedding and no overhead information was required during the message retrieval.

Hong et al. [12] proposed another reversible data hiding scheme based on the modification of prediction errors. They used a JPEG-LS predictor to predict pixel values, and data are embedded into the prediction errors by modifying the error values using the histogram-shifting technique. Their method modified the pixel values by one grayscale unit at most, and the resulting stego image quality has a lower bound 48.13 dB. However, there are two major drawbacks in their method: (1) The pixel values used to predict another pixel values are previously modified, resulting in a less accurate prediction. (2) The peak height of their method cannot be accurately estimated before data embedding, thus the maximum embedding capacity is unknown unless the cover image is fully embedded.

Tsai et al. [13] proposed a block-based reversible data hiding scheme in 2009. The center points of non-overlapping blocks are selected as basic pixels, and the difference values between the basic pixels and their neighbors are modified to conceal data. They used a location map that is similar to Thodi and Rodriguez's work to record the information required for reversibility. However, the values of basic pixels have to be preserved and will not join the embedding process, resulting in a reduction in payload. Take a 3×3 block as an example; one out of nine pixels cannot be used to embed data. Furthermore, the prediction method used in Tsai et al.'s method simply predicts pixel values by the corresponding basic pixel in the same block, leading to a less accurate prediction. As a result, the peak height of the error histogram is decreased and subsequently decreases the payload.

In this paper, we proposed a novel reversible data hiding scheme based on the orthogonal projection technique (OPT) and prediction error modification. Firstly, the OPT is employed to calculate the weights for a given predictor, and this predictor is then used to predict the pixel values such that the total squared prediction error could be minimized. Secret data are concealed into the error histogram by modifying prediction errors. Higher capacity can be achieved by performing multiple layer embedding. Since the weights for each image layer can be embedded along with the secret data bits, only small amount of side information has to be kept for data extraction and image recovery.

The rest of this paper is organized as follows. In Section 2, the reversible data hiding scheme proposed by Tsai et al. is introduced. In Section 3, the orthogonal projection technique, the embedding, extraction and restoration procedures of the proposed method are presented. Experimental results are given in Section 4, and ending with conclusions and remarks in Section 5.

2. Review of Tsai et al.'s work

In 2009, Tsai et al. proposed a scheme to explore the neighboring similarity of pixels in an image to improve the embedding capacity of the histogram-shifting based reversible data hiding technique. It is well-known that, for this type of data hiding scheme, the sharper the distribution of the histogram, the more the embedding capacity is. Since the natural images preserve stronger correlation among neighboring pixels, Tsai et al. exploited this characteristic and used a simple predictor to construct a residue image, which is composed of prediction errors with sharply distributed histogram. Data bits are embedded into the histogram of prediction errors. Tsai et al.'s scheme consists of two procedures: the data embedding procedure and the data extraction and recovery procedures, which will be briefly described in the following subsections.

2.1. The embedding procedure

In Tsai et al.'s method, prediction errors are modified based on histogram-shifting technique to embed data. The prediction errors are classified into two sets: non-negative and negative sets. Each set has its own peak (the value that occurs most frequently) and zero (the value that occurs least frequently). We denoted the peaks of non-negative and negative prediction errors as p_+ and p_-, and denoted the zeros of non-negative and negative prediction errors as z_+ and z_-, respectively. The steps of embedding procedure are listed below.

Input: The cover image I and data bits S.