Modified Matrix Encoding Technique for Minimal Distortion Steganography

Younhee Kim, Zoran Duric, and Dana Richards

George Mason University, Fairfax, VA 22030, USA
{ykim9,zduric,richards}@cs.gmu.edu

Abstract. It is well known that all information hiding methods that modify the least significant bits introduce distortions into the cover objects. Those distortions have been utilized by steganalysis algorithms to detect that the objects had been modified. It has been proposed that only coefficients whose modification does not introduce large distortions should be used for embedding. In this paper we propose an efficient algorithm for information hiding in the LSBs of JPEG coefficients. Our algorithm uses modified matrix encoding to choose the coefficients whose modifications introduce minimal embedding distortion. We derive the expected value of the embedding distortion as a function of the message length and the probability distribution of the JPEG quantization errors of cover images. Our experiments show close agreement between the theoretical prediction and the actual embedding distortion. Our algorithm can be used for both steganography and fragile watermarking as well as in other applications in which it is necessary to keep the distortion as low as possible.

1 Introduction

The goal of digital steganography is to modify a digital object (cover) to encode and conceal a sequence of bits (message) to facilitate covert communication. The goal of steganalysis is to detect (and possibly prevent) such communication. Often, the cover media correspond to graphics files. Graphics files are the typical choice because of their ubiquitous presence in digital society, but any medium that contains a substantial amount of perceptually insignificant data can be used.

Most steganographic methods operate in two steps. First, a cover object is analyzed and the perceptually insignificant bits are identified. It is assumed that changing these bits will not make observable changes to the cover. Second, the identified bits are replaced by the message bits to create an altered cover object. In this paper, cover object is an image in either bitmap or compressed JPEG [13] formats. The perceptually insignificant bits usually correspond to the LSBs in the image representation: in bitmap images these bits correspond to a subset of the LSBs of the image pixels or the LSBs of the color palette entries, in JPEG images they correspond to a subset of LSBs of the JPEG coefficients. Our work applies to both image representations, but our empirical studies have only

J. Camenisch et al. (Eds.): IH 2006, LNCS 4437, pp. 314–327, 2007.

used the JPEG coefficients. Although, the LSBs of JPEG coefficients are usually considered perceptually insignificant modifying some of these bits can produce significant (but imperceptible) distortions of the original image. In this paper we propose an algorithm that embeds a message into the LSBs of a JPEG image. Our algorithm uses modified matrix-coding technique to minimize the distortion of the stego image relative to the clean (non-stego) image.

The paper is organized as follows. In Sec. 2 we briefly review the relevant prior work in the field. In Sec. 3 we provide technical background for our work including the basic facts about JPEG compression and the matrix coding. In Sec. 4 we describe our method and sketch the theoretical analysis of our method. In Sec. 5 we present some experimental results. Finally, in Sec. 6 we present the concluding remarks.

2 Literature Survey

Digital steganography is a relatively new research field [12]. Detailed survey of early algorithms and software for steganography and steganalysis can be found in [12,11,19].

The first quantitative technique for steganalysis was designed by Westfeld and Pfitzmann [17]. They exploited the fact that many steganographic techniques change the frequencies of pairs of values (pairs of colors, gray levels, or JPEG coefficients) during a message embedding process. Their method was shown to be effective in detecting messages hidden by several steganographic techniques. This research prompted interest in both improving statistical detection techniques [5,8] as well as building new steganographic methods that would be difficult to detect by statistical methods [15,18,16,9].

Various attempts have been made to make steganographic content difficult to detect including reducing their capacity or payload and spreading the message across the whole carrier. Anderson and Petitcolas [1] suggested using the parity of bit groups to encode zeroes and ones; large groups of pixels could be used to encode a single bit, the bits that need to be changed could be chosen in a way that would make detection hard.

Provos [15] designed a steganographic method *OutGuess* that spreads a message over a JPEG file; the unused coefficients are adjusted to make the coefficient histogram of the modified file as similar as possible to the histogram of the original image. Fridrich [6] recently developed method for successful breaking of this algorithm. The method exploits the fact that *blockiness* is strongly correlated with the embedding rate. Outguess increases the number of changed bits, which increases *blockiness* between DCT blocks.

Fridrich et al. [5,7,8] reported several techniques for detecting steganographic content in images. If a message is inserted into the LSBs of an image, some features of the image change in a manner that depends on the message size. A possible shortcoming of these methods is that they depend on empirical observations