Chapter 3
A Recent Study on High Dimensional Features Used in Stego Image Anomaly Detection

Hemalatha J
Thiagarajar College of Engineering, India

KavithaDevi M.K.
Thiagarajar College of Engineering, India

Geetha S.
Vellore Institute of Technology Chennai, India

ABSTRACT

This chapter describes how ample feature extraction techniques are available for detecting hidden messages in digital images. In the recent years, higher dimensional features are extracted to detect the complex and advanced steganographic algorithms. To improve the precision of steganalysis, many combinations of high dimension feature spaces are used by recent steganalyzers. In this chapter, we present a summary of several methods existing in literature. The aim is to provide a broad introduction to high dimensional features space used so for and to state which the most accurate and best feature extraction methods is.

INTRODUCTION

The tremendous communication technology growth and unrestricted practice of internet have significantly smoothed the data transfer. In spite of this practice, it makes the communication channels more vulnerability to data security terrorizations and initiating the unauthorized information access. To provide a solution to this problem, data hiding concepts such as steganography, watermarking are emerged. Steganography is an art of hiding the data in an innocuous cover medium such as image, audio, video and text, firewall, protocols, etc. This technology can be misused by terrorists and criminals for scheduling and synchronizing the felonious activities. The idea of a way to do this is, the secret messages are em-

DOI: 10.4018/978-1-5225-4044-1.ch003

bedded in digital images and post it in public spots so that the others are not known about the message existence. Later this chapter will discuss about the types of steganography, tools existing for hiding the secret messages, applications of steganography, misuse of steganography, etc. In contrast, steganalysis is an art of sensing the data hidden in the digital media. The aim is to gather adequate proof since a cover image is hidden by a secret message. The purpose of using the digital image as a carrier file is wide availability of high-resolution pixels. Fundamentally three common image formats are used for the hiding purpose; they are JPEG, BMP, and GIF. Each format will perform differently when it is embedded by a hidden message. Consequently, various steganalysis algorithms are there for each image format.

For the GIF (Graphics Interchange Formats) image format, palette based image steganalysis is used predominantly. It encourages only 8 bpp (bits per pixel), pixel colors are indicated from the color palette table. It contains 256 distinct colors and it will be mapped to the 24-bit RGB color image. The hiding algorithm strength lies in lessening the probability of the color palette change and also in lessening the visual distortion about the occurrence of the hidden message.

On the other hand, JPEG image format is the most popular cover choice for hiding the stego content. With the background of JPEG images, some standard steganographic algorithms are available such as JSTEG, F5 (Westfeld et al., 2001), Outguess (outguess), etc. In the JPEG image format, each image is divided into 8 by 8 blocks; in each block first component is the DC component remaining is the AC components. In F5 algorithm matrix embedding is used to embed the message bits in the DC coefficients. Likewise, in the starting age of steganography and steganalysis Fridrich et al. (Fridrich et al., 2003) proposed a practice for appraising the unaltered histogram to calculate the number of changed bits and the hidden message length. To calculate this initially the image has been cropped by four columns and then the image has been recompressed by the quantization table. The preceding histogram of DCT coefficients will very close to the original image. Also in (Fridrich et al., 2003) technique has been proposed for attacking the outguess algorithm.

In the ancient day's steganalysis done by visual analysis: detection has been done with the naked eye or analyzing the bit planes of an image separately for any scarce appearance in an image. Then the steganographers are very clever in designing the steganography tools and algorithms since it cannot be distinguished with visual attacks. Later steganalysis did by statistical analysis: examining the statistical properties of an image, whether the properties are changed due to steganographic embedding. This statistical steganalysis can be categorized into specific and universal/blind. In the case of specific steganalysis, the detection can be done only if the embedding steganographic algorithms are previously known. While in universal/blind steganalysis, the stego objects can be detected when it is embedded with any steganographic algorithms.

Most of the recent techniques for blind steganalysis are depending on two phases. In the first phase, the statistical features are extracted with reducing the dimensionality. In the second phase train the classifier with the set of clean and stego images, the decision has to be carried out for detecting the stego images from the extracting features. In the earlier days, the statistical features space is in fewer dimensions, the existing classifier has been achieved good performance. To increase the recognition rate of modern steganographic systems, the feature dimension has to be extensively increased.

Steganalysis consistency is strongly influenced by the cover source. In the prisoner's problem, choosing a cover source should be a better mask for hiding the secret message. In Eve's problem, even a single-bit message hidden also eves should successfully detect the message.