Steganalysis Using Image Quality Metrics (Normalized Correlation Metric)

Adil Ismaeel Kahdum
Department of physics, Ibn Al-Haitham College, University of Baghdad.

Abstract

A new steganalysis method presented in this work to test the existence of hidden data in an mage and detect the data. The method depending on the comparison of the stego image with the host image using Normalized Correlation Function Metric, this method must be applied on the images that have the same scene (human eyes cannot distinguish one from the other), and by calculating the Normalized Correlation Function Metric value the hidden data detected, then the amount of the data calculated too with the percentage of changing pixels in the stego image.

Introduction

Steganography is the art of hiding information in an innocuous cover. Its basic purpose is to make communication unintelligible to those who do not possess the right keys. The message can be hidden inside of images or, other digital objects, which remains imperceptible to a casual observer. By embedding a secret message into a cover image, a stego image is obtained. As the stego-image does not contain any easily detectable visual artifacts due to message embedding, techniques.

The goal of steganography is to communicate securely in a completely undetectable manner, in such a way that an adversary should not be able to differentiate in any sense between cover image

(image not containing any secret message) and stego image (image containing a secret message)(1).

Steganography is a game between the hider and the seeker. The hider wants to hide as much information as they can without being discovered. This maximum amount is defined by the tools a seeker is able to analyze data with.

Steganography deals with hiding messages in cover images. To embed a message, the cover image is slightly modified by embedding techniques (2).

The process of detecting steganographic messages is known as steganalysis and a particular steganalysis technique is called an attack, if the image is carefully chosen then visual detection is difficult (3).

The goal of steganalysis is to break steganography. That is, a steganalysis detector attempts to detect the presence/absence of an embedded message. Steganalysis can be classified into two general categories: (a) passive and (b) active.

The major goals of passive steganalysis are the following:

• Detect presence/absence of hidden message in a stego image.
• Identify the stego embedding algorithm.

While active steganalysis deals with the following:

• Estimate the embedded message length.
• Estimate location(s) of the hidden message.
• Estimate the secret key used in embedding.
• Estimate some parameters of the stego embedding algorithm.
• Extract the hidden message (1, 3, 4).

In the present work Image quality Metrics method will be used as a steganalysis method to stegdetect the hidden data (security message in images.

Image quality Metrecs

Image quality metric (IQM) is essential for most image processing applications. Any image and video acquisition system can use the quality metric to adjust itself automatically for obtaining

improved quality images. It can be used to compare and evaluate image processing systems and algorithms (5).

Image quality metrics are paramount to provide quantitative data on the fidelity of rendered images. Typically the quality of an image synthesis method is evaluated using numerical techniques which attempt to quantify fidelity using image to image comparisons, several image quality metrics have been developed to predict the visible differences between a pair of images (6).

Present technique for steganalysis of images that have been potentially subjected to Steganographic algorithms, both within the passive warden and active warden frameworks. Present hypothesis is that steganographic schemes leave statistical evidence that can be exploited for detection with the aid of image quality features and multivariate regression analysis. To this effect image quality metrics have been identified based on the analysis of variance (ANOVA) technique as feature sets to distinguish between cover images and stego images. The classifier between cover and stego images is built using multivariate regression on the selected quality metrics and is trained based on an estimate of the original image.

This work is based on the fact that hiding information in digital media requires alterations of the signal properties that introduce some form of degradation, no matter how small; these degradations can act as signatures that could be used to reveal the existence of a hidden message.

Image quality metrics are categorized into six groups according to the type of information they are using. The categories used are:

1- Pixel Difference-based measures.
2- Correlation-based measures.
3- Edge-based measures.
4- Spectral Distance-based measures.
5- Context-based measures.
6- Human Visual System-based measures (7, 8).

Normalized Correlation Function Metric

Normalized Correlation Function Metric is one of the best known methods that evaluate the degree of closeness between two functions. This measure can be used to determine the extent to which the original image and the stego image are close to each other, even after embedding data. Localization that is detection of the presence of the hidden data relies on the use of Correlation Function of two images (9).

This measure measures the similarity between two images, hence in this sense it's complementary to the difference-based measures.

This work based on the following equation:

$$S = \frac{\sum_{i,j=0}^{N-1} (C(i,j) \times \hat{C}(i,j))}{\sum_{i,j=0}^{N-1} (C(i,j))^2}$$

Where:

S Normalized Correlation Metric value.

N the size of the images under test.

$C(i,j)$ $(i,j)^{th}$ pixel value of original image.

$\hat{C}(i,j)$ $(i,j)^{th}$ pixel value of stego image.

There is one condition to apply this method and to obtain accurate results, this condition is the two images under test must be the same in the external scene (visible properties), that means the human eyes cannot distinguish between them (7, 8).

The Results

As we see and notice from figure (1) and table (1), while every pair of images look identical in the shape (scene) in fact they are not, they are different in the content; and we can estimate and conclude the following results:

1- The result of applying present method on the original image with itself [(image (a) with (a), (b) with (b), (c) with (c) and (d) with (d)] the Normalized Correlation Function Metric value equals to (one), and the hidden data length equals to (zero).That means the minimum value of Normalized Correlation Function Metric value equals to zero , in other words when the Normalized Correlation Function Metric value equals to (zero) that means there is no hidden data in the image .

2- When the value of Normalized Correlation Function Metric value is greater than (one) that means the two images (under test) are not identical ,in other words the stego image is carrying hidden data (secret message) ,and the hidden data length (calculated) is greater than (zero).

Table (1): Images under test, changing pixels rate, Normalized Correlation Metric value and hidden data length.

	Images Under Test	Changing Pixels Rate	Normalized Correlation Metric value	Hidden data length (bit)
1.	Image (a) With Image (a)	zero	1	zero
2.	Image (a) With Image (b)	0.69873	1.02808	900000
3.	Image (c) With Image (c)	zero	1	zero
4.	Image (c) With Image (d)	0.135569	1.00752	146415
5.	Image (e) With Image (e)	zero	1	zero
6.	Image (e) With Image (f)	0.656535	1.01338	393921
7.	Image (g) With Image (g)	zero	1	zero
8.	Image (g) With Image (h)	0.0364457	1.00305	38268

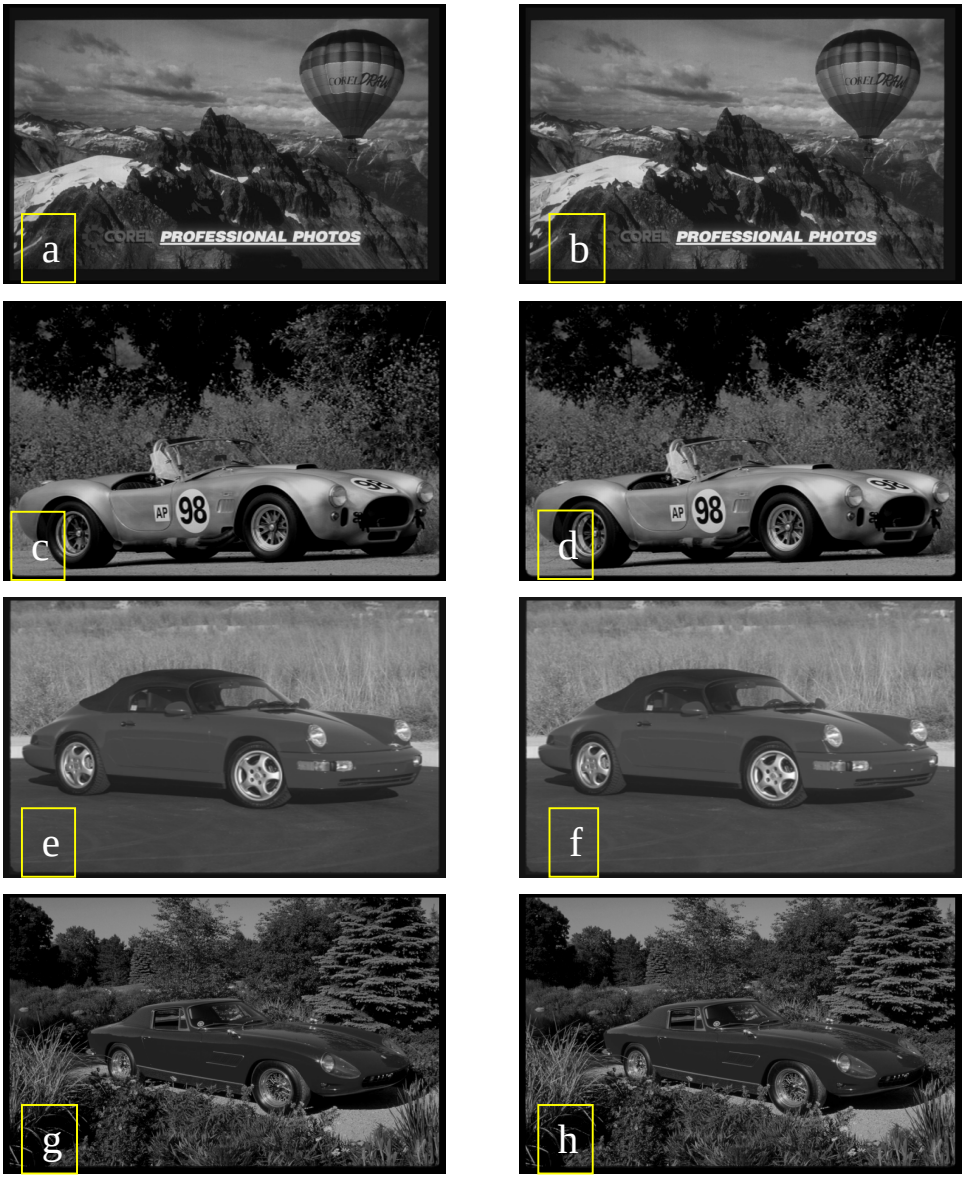

Figure (1): (a,c,e,g) Host images(original), (b,d,f,h) Stego images

Conclusions

In present work a new method applied on several images to detect hidden data (secret message) in stego image by comparing the host image (original image) with the stego image using the calculating of Normalized Correlation Function Metric value, then calculate (estimate) the hidden data length (amount of hidden data) and calculate the percentage ratio of changing pixels of the stego image.

This method can be used to compare any two images ,and when the result of Normalized Correlation Function Metric value between these images is not equals to (one) that means the two images is not identical , there is some difference between them depending on the amount of the hidden data.

References

1-Pierre Moulin, Ralf Ketter, 2005, Proceedings of The IEEE, Vol. 93, No. 12,"Data-Hiding Codes".

2-Jeremiah Joseph Harmsen, April 2003, MSc Thesis, Rensselaer Polytechnic Institute, Troy, New York, "Steganalysis Of Additive Noise Modelable Information Hiding".

3-Po-Chyi Su and C-C Jay Kuo, 2003, IEEE Transactions on Consumer Electronics, Vol. 49, No. 4, pp 824 - 832, "Steganography in JPEG2000 Compressed Images.

4-R.Chandramouli, 2002, Department of Electrical and Computer Engineering, Stevens Institute of Technology, California, "A Mathematical Approach to Steganalysis".

5- D.Venkata Rao, N.Sudhakar, B.Ravindra Babu and L .Pratap Reddy, 2006, GVIP Journal, Vol. 6, Issue. 2, "An Image Quality Assessment Technique Based on Visual Regions of Interest Weighted Structural Similarity".

6-Alan Chalmers, Scott Daly, Ann Mc Namara,Karol Myszkowski and Tom Troscianko, 2000, 23-28 July, Sigraph ,New Orleans , USA " Image Quality Metrics".

7- Ismail Avcıbas, Nasir Memon and Bülent Sankur, 2003, IEEE Transactions on Image Processing,Vol 12, No 2, PP 221-229, “Steganalysis Using Image Quality Metrics”.
8- Ismail Avcıbaş, 2001, PhD Thesis, Bogaziçi University,”Image Quality Statistics And Their Use In Steganalysis And Compression”.
9-Venkatraman .S, Ajith Abraham, Marcin Paprzycki , 2004, Proceedings of the International Conference on Information Technology, Coding and Computing ,” Significance of Steganography on Data Security”.

كشف الصور المجفرة باستعمال مقاييس نوعية الصورة (مقياس دالة الارتباط المعاير)

عادل اسماعيل كاظ
قسم الفيزياء ، كلية التربية ـ ابن الهيثم ، جامعة بغداد

المستخلص

طُبقت في العمل الحالي طريقة جديدة للكشف عن وجود بيانات مخفية (رسائل سرية) في صورة وذلك اعتماداً على مقاييس نوعية الصورة (مقياس دالة الارتباط المعاير) بين صورتين، إن تطبيق هذه الطريقة يجب أن يتم على الصور التي لا يبدو بينها أي اختلاف في المنظر الخارجي (الصور المتطابقة التي لا يمكن لعين الانسان التمييز بينها) ، ومن حساب قيمة دالة الارتباط المعاير يمكن الكشف عن وجود أو عدم وجود بيانات مخفية .
تم تطبيق هذه الطريقة على أربع صور وتم حساب قيمة دالة الارتباط المعاير وحجم البيانات المخفية والنسبة المئوية لعدد عناصر الصورة (البكسل) المغيرة في الصورة المجفرة.