Diala , Jour , Volume , 42 , 2010

Hidden Data Detection In An Image Using The , Histogram And Entropy

Asst.Adil Ismaeel Kahdum
Department of Physics
College of Education lbn Al-l-laitham
University of Baghdad

Abstract :
The aim of present work is to detect the hidden data in an image by comparing any two identical images (in the external scene) using the Histogram of each image under test and computing the Entropy of each image, so we can after these process break the hidden data detected in the images under test.

Introduction :
Steganography is the art of passing information through apparently innocent tiles in a manner that the very existence of the message is unknown. The term Steganography literally means "Covered Writing"ll). Secret can be hidden inside all sorts of cover information: text image, audio video... (21.

The goal of Steganography is to communicate securely in a completely undetectable manner, such 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 messagele 3|.
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.
Steganalysis is the process of detecting the existence of the Steganography in a cover medium and rendering its useless 1. .

In present work image histogram and entropy are used as ' Steganalysis methods to detect the hidden data (secret message) in the images.

in the present work The Histogram and the Entropy of the images will be used as a Steganalysis method to detect the hidden data (security message) in images.

Histogram :

A histogram is a graph which is used to determine the distribution of a set of data. It represents the number of pixels with the generic grayscale level. Therefore, the histogram gives the frequency of occurrence of the random event ||grayscale intensity" appearing in the sampling space (RANGE) defined by all the pixels in the digital image (4) The histogram is an approximation of the probability density function of a random variable whose realization is the particular set of pixel values found in the image. The image histogram, a first-order statistical measure of the distribution of image pixel amplitudes, is a widely used source of histogram- based features. These play an important role in region-based image segmentation. The shape of an image histogram provides many clues as to the character of the sample values. Commonly used histogram-based features include the mode (i.e., the most frequently occurring amplitude) (51.

The histogram of an image is a function that provides the frequency of occurrence for each intensity level in the image.

Image segmentation schemes often incorporate histogram information into their strategies. The histogram can also serve as a basis for measuring certain textural properties, and is the major statistical tool for normalizing and requantizing an image.

For a given gray level of an image, the value of the cumulative histogram is equal to the number of elements in the image's domain that have gray level value less than or equal to the given gray level. The cumulative histogram can provide useful information about an image (6). '

Entropy :

Since the concept of entropy in physics was introduced into information theory by Shannon, it has been applied in image processing tasks, such as image segmentation, compression, retrieval, restoration, reconstruction, enhancement, correction, fusion and matching (7)

The entropy .definition of a single random variable can be extended to a pair of random variables. The image entropy, Eq.
(1), is usually estimated using an image histogram.

$$H(X) = -\sum_{i=1}^{n} P_i \log P_i$$

Where $P_i = \Pr[X = x_i]$

Let : F= f(x,y)]mxn be an image of size MxN, where f(m,y) is the gray value of pixel (m,y) l < x < m,1 < y < n, and f(x,y) EGl =l0,1,......,L -1) the set of grey levels. The entropy of a pair of discrete random variables with a distribution Pij is defined as:

$$H(f)=-\sum_{i=1}^{n}\sum_{j=1}^{m}P_{ij}\log P_{ij}$$

Where pijdenotes the probability of f(x,y) If F is a sub image, h(f) is called local entropy. The more similar two images are the lower is their variance entropy.

The entropy could be used as a similarity measure between two images (8) Thus the present method of detecting the hidden data depends on the dissimilarity between two images by using the image Histogram and Entropy .But the two images under test must be identical in the external scene that means the human eye cannot distinguish between them.

Results :

As shown in Figure (1), Figure (2) and Table (1), the following results can be concluded : 1- In spit of that the each group of images under test (A1, A2, A3) and (B 1, B2, B3) and (C1, C2, C3) have the same scene; in fact we can see that they have different histograms (human eye can distinguish and discriminate

between their histograms), that means they have different structures (the gray level of any pixel in the first image is not equals with the gray level of the same pixel in the second image).

2- The Entropy value of each image is different (in spit of that they have the same scene).

3- The PSNR value of any two images (comparing any image with the original image ,A2 with A1 and A3 with A1) and so on is not equals to zero and more than (33 dB) that means the human eyes cannot distinguish between them.

That means every image is different from the other in the same group.

Conclusions : .

In this work a method depending on the Histogram and Entropy of the image presented to detect any changing in the structure of any (unknown image) by comparing it with an original image (known image) under a condition that the two images must be identical in the scene (human eye cannot distinguish between them) and that means detect the hidden data in an known image and then destroy this data.

References :

1- Sanjay Ktunar Jena,G.V.V.K.ris1ma (2007) ,lntemational Journal of Computers,comunications & Controlsvol.ll,No
2,pp149-158,"Blind Steganalysis:Estimation of Hidden Message lengthen.

2- Adil lsmaeel Kahdum,lBN AL-HAITHAM jFOR PURE &
APPL.SC1,VOL.21,(1),2008,"Lossless Data Hiding Using
LSB Method"

3- Adil Ismaeel Kahdum, AL-fatih Journal, No, 32,
2008,"steganalysis Using Image Quality Metrics (Normalized Correlation Metric).

4- Dchen and Xching , PATTERN RECOGMTION AND
STRING MATCHING Kluwer Academic
Publisher,2002,pp399- 452

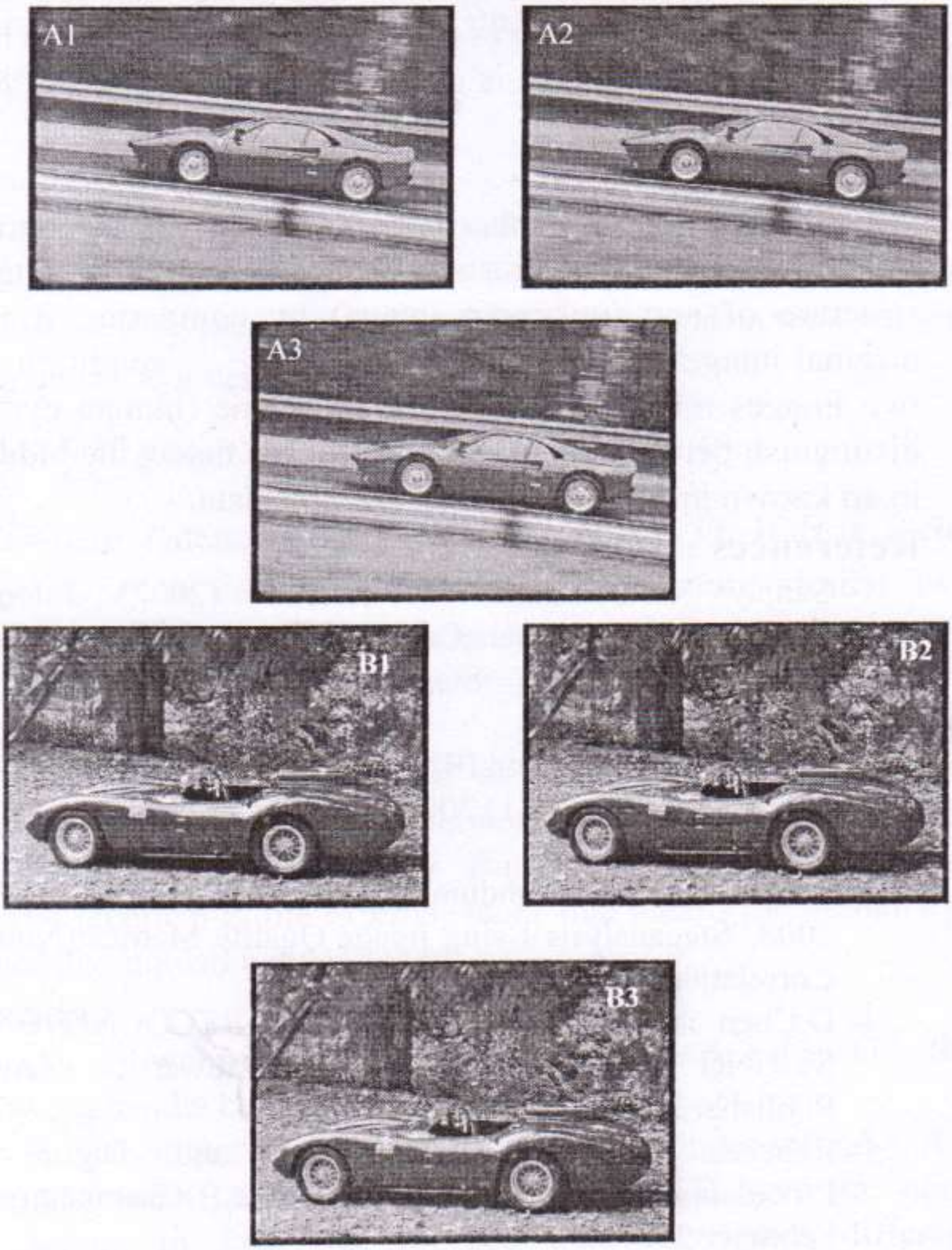
5- Rebecca Bigelow," Mathematics and Digital Image
processing, Wolfram Research, inc., Champaign, USA, february 2007.

6- Gerhard X. Ritter and Joseph N. Wilson's HANDBOOK Of
Computer Vision Algorithms in Image Algebra's, CRC Press
LLC, USA, 2001.

7- Kai-kum Dong, Le-jun Chi, Li Guo, Proceedings of the) 1th Joint Conference on Information Sciences (2008)," Parallel Image Matching Based on Local Entropy Difference Vector.

8-K. KUCZYRSKI and P. MIKOEAJCZAK, OPTOELECTRONICS

REVIEW 11(3), 253- 259, 2003," Information theory based medical image processing.



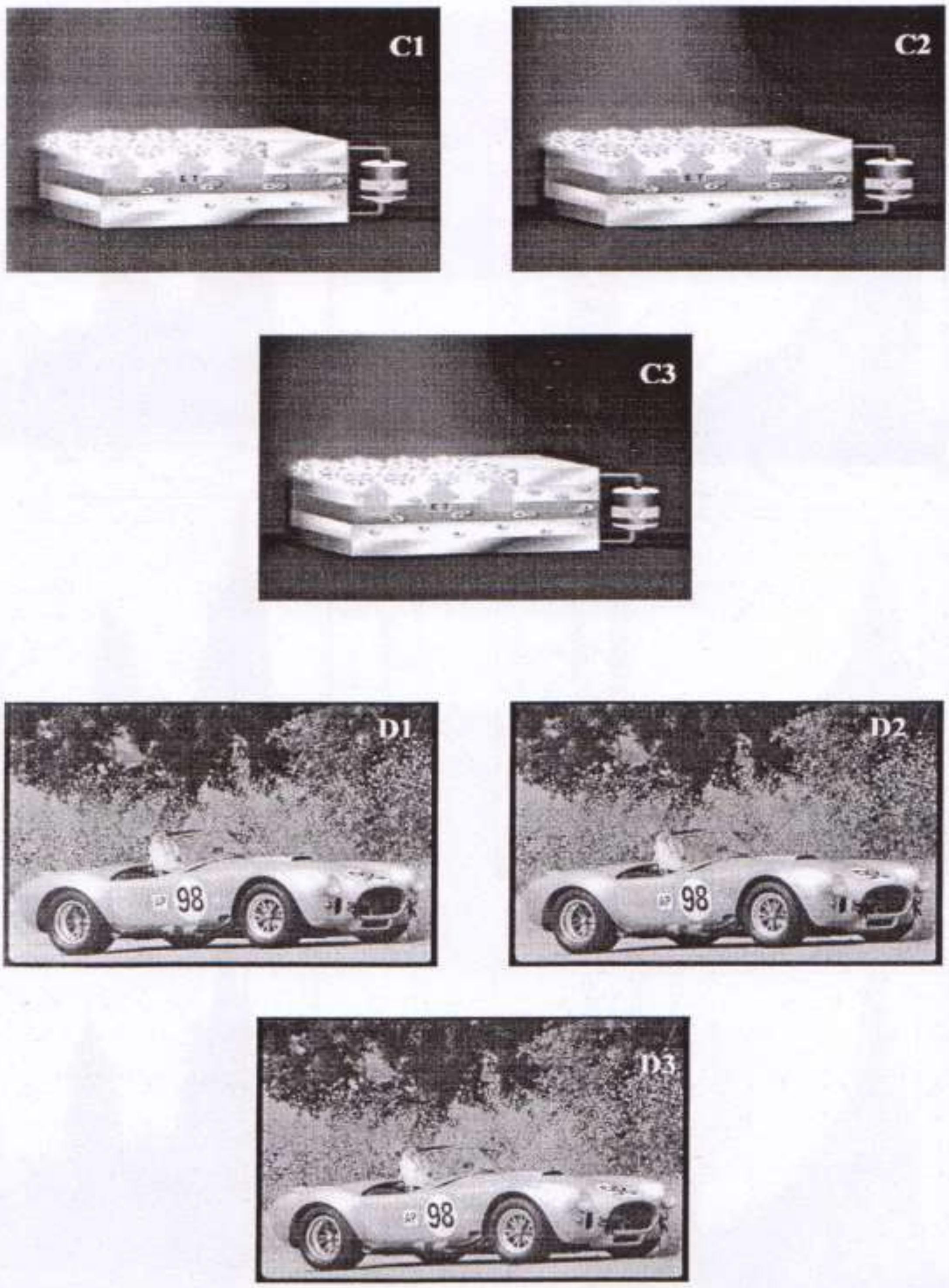
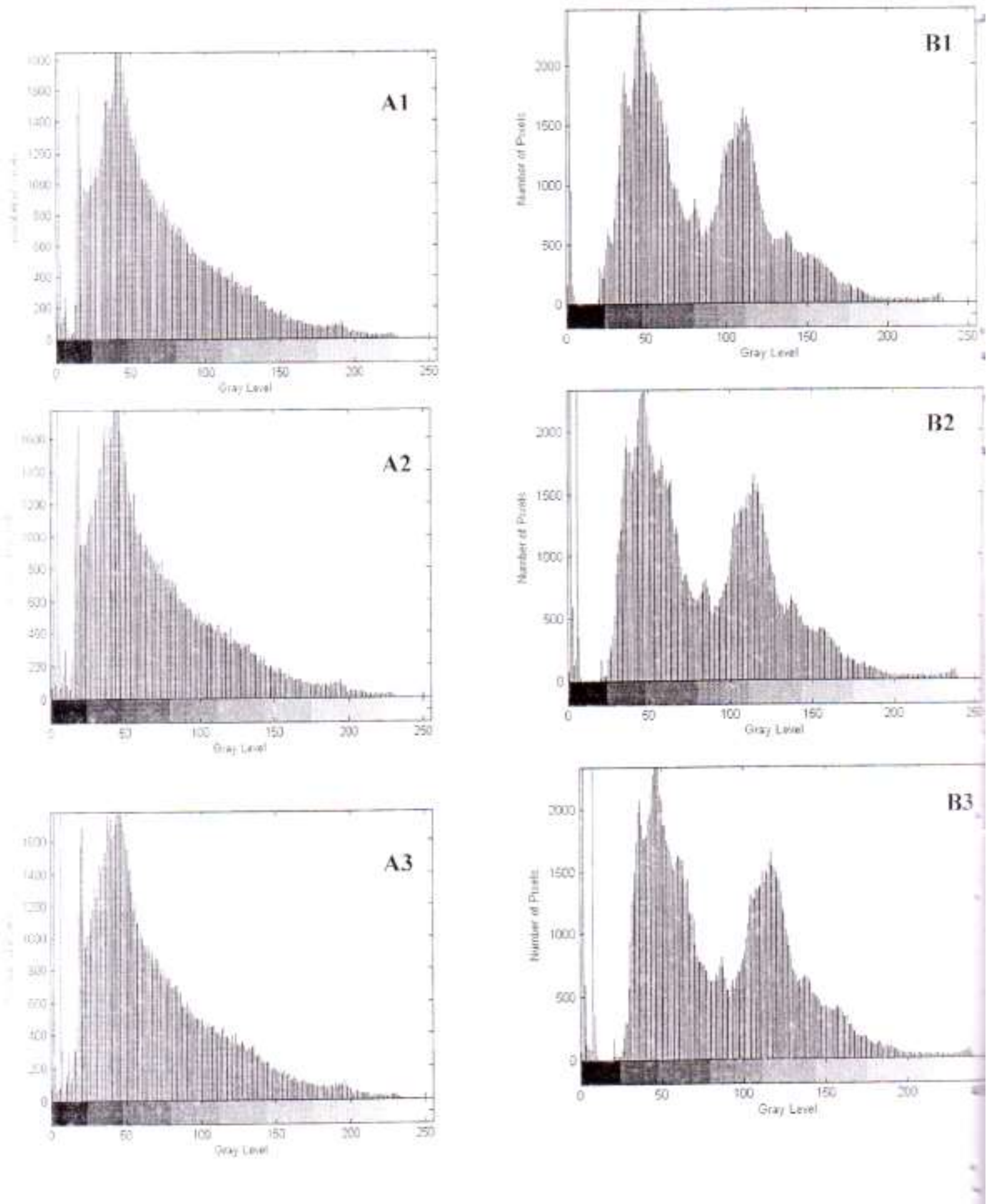


Figure (1): Each of similar group images under test.


A1
A2
A3
B1
B2
B3
Number of Pixels
Gray Level

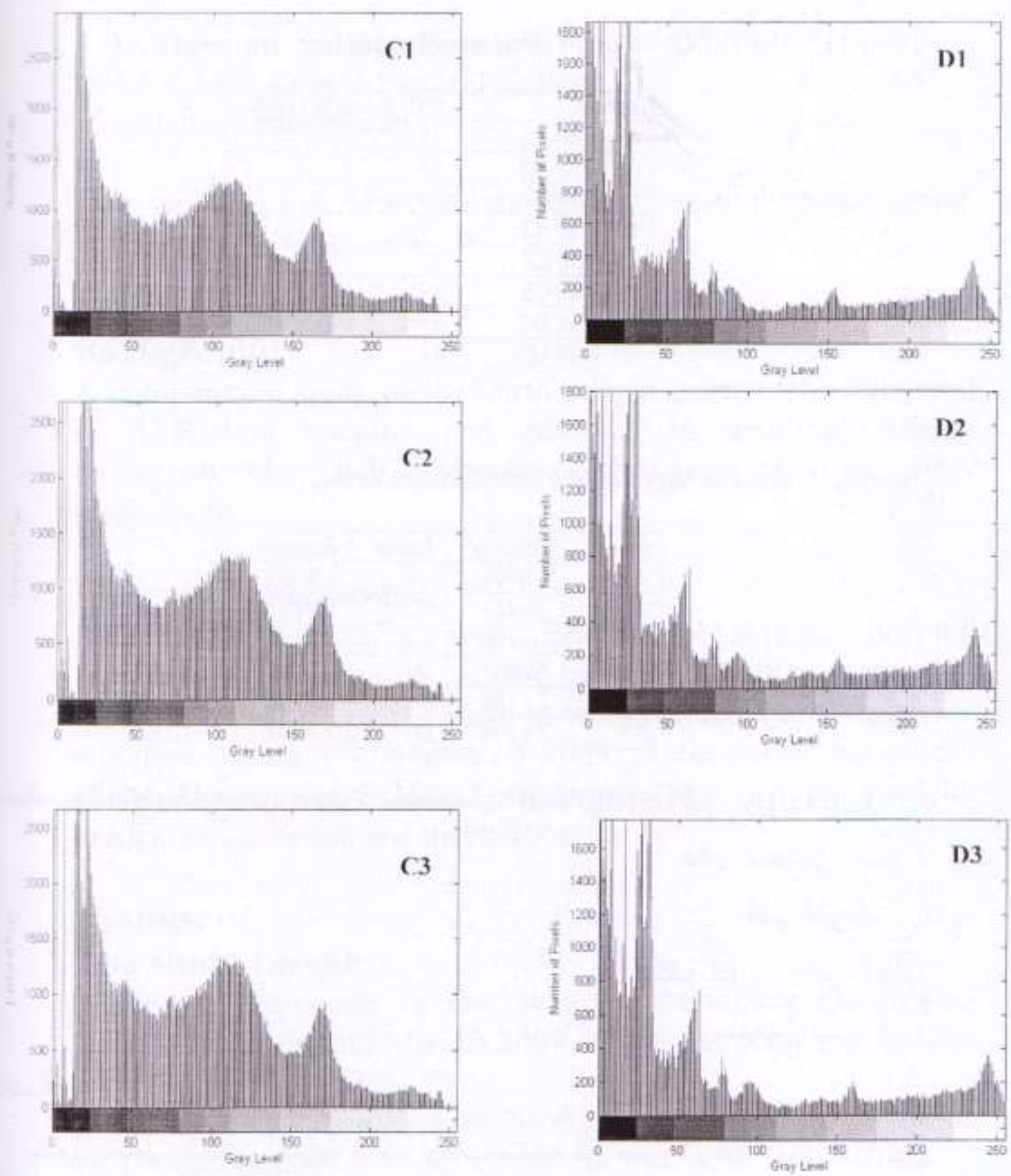


Figure (1): The Histogram of each image in figure (1) respectively.

Table (1): The PSNR value of each two images

Image	PSNR Value (Db)
A1 with A2	41
A1 with A3	37
B1 with B2	39
B1 with B3	37
C1 with C2	45
C1 with C3	39
D1 with D2	51
D1 with D3	41

Table (2): The Entropy value of each image alone.

Image	Entropy	Image	Entropy
A1	7.0990	C1	6.8657
A2	7.1411	C2	6.9162
A3	7.1486	C3	6.9446
B1	7.0887	D1	7.3586
B2	7.1386	D2	7.4499
B3	7.1394	D3	7.4704

كشف بيانات مخفية في صورة باستعمال ENTROPY .HISTOGRAM

م.عادل اسماعيل كاظم قسم الفيزياء كلية التربية- ابن الهيثم- جامعة بغداد

الخلاصة

يتناول البحث الحالي عملية كشف بيانات مخفية في صورة وذلك بأخذ صورتين متطابقتين في المنظر الخارجي ومقارنة مخطط توزيع التواتر HISTOGRAMلكل من الصورتين وكذلك من حساب مقدار entropyلكل من صورة من الصور تحت الاختبار وبعدها ونتيجة لهذا العمل يمكن ازالة او تحطيم البيانات المكتشفة.