A general Purpose Image-Based Electors Smart Card Using an Enhanced Least Significant Bit Steganographic Method for Information Hiding: A case study of the Kenyan Electoral Process.

Gabriel Macharia Kamau[1], Stephen Kimani[2] and Waweru Mwangi[3]

[1] School of Computer Science and Information Technology, Dedan Kimathi University of Technology, P.O. Box 657-10100, Nyeri, Kenya

[2,3] Institute of Computer Science and Information Technology, Jomo Kenyatta University of Agriculture and Technology, P.O. Box 62000-00200, Nairobi, Kenya

Abstract

To legally participate in any political elections in Kenya, one must of necessity be in possession of at least two documents i.e. the elector's card and the national identity card. Both of these documents make use of the bearer's face photograph and some textual identification information.

The unprecedented growth in digital imaging and processing technologies have meant that, with little or no knowledge of the technology involved one is conveniently able to tamper with photographs and identity information on such documents resulting in cases of fraud e.g double registration and voting which have consistently marred our electoral process.

In this paper we propose the use of a general purpose smart card based identification document during elections. The textual identification information of the card bearer and his or her biometric data are embedded in his or her face photograph using an enhanced, randomized least significant bit steganographic method.

An experimental design was setup to determine the effectiveness of the method by comparing the original and the reconstructed images' statistical characteristics to ensure that no notable differences exist.

The experimental results show improved security of the biometric data and personal identification details.

Keywords: *Biometrics, Steganography, Smart Card, Carrier Image*

.

1. Introduction

The voting process in Kenya is becoming increasingly complex by the day. During the next and subsequent general elections for example, there will be more ballots cast at the same time than in any other voting time in the history of the country. Past experiences have shown that malpractices ranging from double registration, double voting and vote buying do occur. The duration taken during verification of a voter's documents creates unnecessary delays and in some instances illegitimate documents can be used in a voting process. Fast and reliable identification of a voter is therefore very important and necessary to make the elections smooth, effective and fraud free.

Biometric authentication systems are a viable and reliable method to address the issue of use of multiple documents during voting. This is the automatic verification of a person's claimed identity through the use of physiological traits of the individual eg fingerprints, face, hand veins, iris, retina, palm print etc.

A biometric system helps in authenticating one's identity by use of a smart card and a password or identity number. The biometric template is compared with the details of the individual stored elsewhere e.g. in a central database. Fig. 1 shows a sample picture of a smart card.

Fig. 1 A sample picture of a smart Card

2. Smart Card

A smart card is a portable credit-card-sized document with a memory and a microprocessor. It can manipulate and store data. To retrieve its contents, it is normally swiped in a reader. A smart card contains a digital certificate and makes use of a knowledge factor such as a password.
There are various types of smart cards in literature [1]. The template - on - card (TOC) smart card is used in this paper. In a TOC, the original individual's verification biometric template is stored inside the smart card's memory. To authenticate the person's identity, the reader retrieves the identifying template from the smart card's memory and tries to match it with that stored for a registered individual in a database stored elsewhere.
A number of smart card biometric based identification systems exist and are used the world over. Examples include the UK's asylum seekers' card and the US defense department access card [1]. Biometric based smart cards can be used as a common identity document for a variety of requirements within business or government environments. This is because they can support multiple authentication and authorization methods e.g. use of biometric password keys, digital certificates etc. Smart cards can also be applied in controlling physical access to buildings and social places, authenticating computer logon, accessing remote networks and email etc.
In Kenya and indeed most African countries, smart cards can promote convenience to voters and the voting process administrators bringing about huge saving in terms of costs through consolidation of authentication process by use of a single identity document.
.

3. Biometric Authentication Using Finger Prints

Biometric based authentication system makes use of personal attributes data or body data e.g. iris or finger prints [2]. This is because this data is not only unique to the individual concerned but it also remains throughout the individual's life span. The growth in the Information Technology industry and the continued embracing of the same in every facet of life makes it very necessary to have reliable and secure identification.
Every individual possesses unique finger print properties which remain so throughout his life time. Accordingly, the finger prints matching is one of the most secure and reliable technique of proving identity. They consist of a unique pattern formed by ridges of raised skin and furrows of lowered skin. Making an inked imprint of a thumb leaves an impression of ridges while the un inked areas of the fingers represents the furrows between the ridges [3]. Terminations or bifurcation of finger print ridges known as minutiae are most unique and therefore important in the identity of an individual. Fig. 2 shows the parts of a finger print.

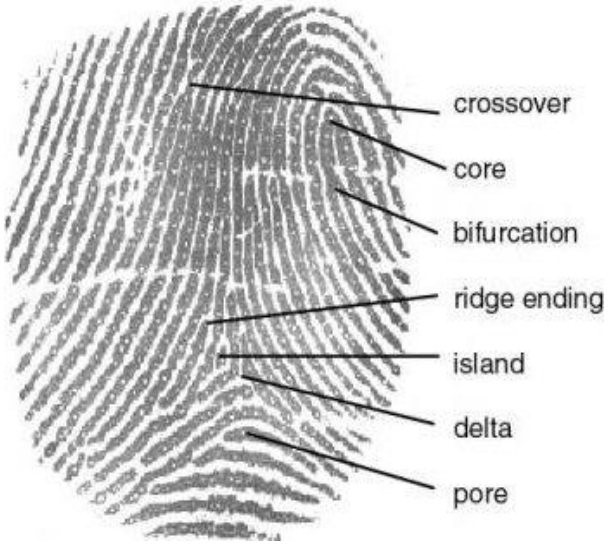

Fig. 2 Parts of a Finger print

For this paper, use of finger print biometric is chosen due to the smart card's memory size and also time factor for verification. Commonly available smart cards have approximately between 8-16 KB of memory while the minimum size of a finger print template usable for matching and analysis is approximately a few hundreds of bytes. For the sake of time, the inbuilt processor in the smart card should complete the matching process as fast as possible (real time).

4. The Enhanced Least Significant Bit Steganographic Method for Information Hiding

Steganography is the science of hiding a confidential message in an innocent looking container file e.g. an image in a way that does not introduce perceptible distortions to the carrier file [4]. It helps to conceal the existence of data. By use of steganography, confidential details can be embedded in a cover file and transmitted to the intended person without creating suspicion. A successful steganographic system embeds information imperceptibly in a cover file making sure that the exact information is extractable at the other end.
This paper proposes a biometric – based all purpose voters' card which stores the voter's details including identity card details and uses his finger prints template in

the smart card to verify his identity. The finger print template and the voter's details are embedded in the voter's face image stored in the smart card while the same details are stored in a centralized database elsewhere. After extracting the voter's details from his image in the smart card, they are matched with the details in the database. This helps in arresting the issues of tampering with the voter's documents and also curbs double registration and voting. It also makes the process of voting faster and saves on costs.

In a steganographic system, there exists an embedding algorithm and an extraction algorithm. Using the embedding algorithm, the cover image's bits are slightly modified to accommodate the hidden information.

The most commonly used method in the spatial domain is the least significant bit (LSB) insertion method which is a simple steganographic algorithm that takes the least significant bit in some bytes of the cover medium and swaps them with a sequence of bytes containing the secret data in order to conceal the information. However its imperceptibility and hiding capacity are relatively low. This is as revealed by the statistical characteristics of its resultant stego images compared to the original cover images [5].

The use of an enhanced LSB method which utilizes varied and random bits in a true color image to embed the confidential message is proposed in this paper. The Linear Congruential Generator (LCG) method proposed by D.H. Lehmer is used to generate the pseudo random numbers used to match the specific bits in the cover image where the secret data bits are hid. LCG method is one of the most successful random number generators. It is also fast and saves on computer memory.

The formula is explained below.

$$X_{n+1} = (aX_n + c) \bmod m \qquad (1)$$

Where:

X_0 is the starting value , the seed ; $0 <= X_0 < m$

a is the multiplier; $a \geq 0$

c is the increment; $c \geq 0$

m is the modulus; $m > X_0$, $m > a$, $m > c$

The desired sequence of random numbers $< Xn >$ is then obtained by setting

$X_{n+1} = (aX_n + c) \bmod m$, $n \geq 0$

X_n is chosen to be in $[0, m-1]$, $n \geq 0$

Given that the previous random number was X_i, the next random number X_{i+1} can be generated as follows.

$$X_{i+1} = f(X_i, X_{i-1,\ldots,} X_{i-n+1}) \bmod m = (a_1 x_i + a_2 x_{i-1} + \ldots + a_n x_{i-n+1} + c) \bmod m \qquad (2)$$

A stego key (k) is used during extraction which in this case is the message digest of the user supplied password.

According to Hull & Dobell [6], a linear congruential sequence defined by m,a,c and X_0 has full period if and only if the following three conditions hold:

- The only positive integer that exactly divides m and c is 1
- If q is a prime number that divides m, then q divides a -1
- If 4 divides m, then 4 divides a -1

Additionally, the value of m should be rather large since the period cannot have more than m elements. The value of m should also necessitate a fast computation of (aXn+c) i.e speed the generation of random numbers. Observing all these requirements, the parameters for the LCG were picked as follows:

4.1 Modulus (m)

The 48-bit computer word length was picked as the value of m. Any Pentium IV and above computer should have this word length or larger. This in essence provides the size of m to be 2^{48} which is equivalent to **281,474,976,710,656.** For the sake of this experiment and bearing in mind that the digitals images used are a few kilobytes in size, this period was sufficient enough to set up the experiment.

To ensure faster generation, m is recommended to be a power of 2 or close to a power of 2 and hence the choice of the word length. Using the AND operation also enhance speed instead of the normal division operation which is considered slower.

4.2 The seed (X_0)

The first value of the seed (X_0) is supplied by the message digest of the user supplied password. This is done using a special form of encryption that uses a one-way algorithm which when provided with a variable length unique input (message) will always provide a unique fixed length output called hash, or message digest.

4.3 Multiplier (a) and Increment *(c)*

To ensure full period and in following with requirements identified above, the values of the multiplier and the increment are picked as follows.

a (Multiplier) = 25214903917

c (Increment) = 11

The values are used to initialize the random number generator used for the experiment.
The numbers generated by this PRNG determines the specific bits in the pixel bytes of the cover image where data bits of the secret data file are to be embedded.

For example considering storing the 200, which binary representation is 11001000 in a grid of 3 pixels of a 24-bit image, utilizing a single LSB of each color channel the enhanced LSB algorithm will store the significant bits of the message randomly into the cover image bits as shown in Fig. 3 and 4 below assuming the LCG parameters (X_0 = 7, a = 7, c = 7, m = 9)

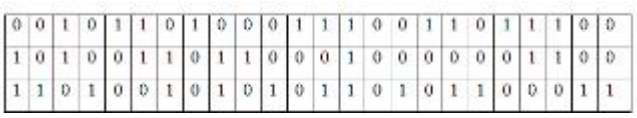

Fig. 3 Original Bits

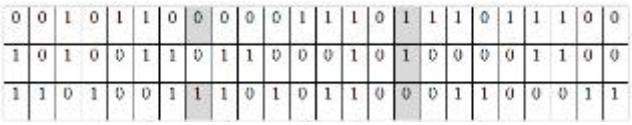

Fig. 4 Modified Image Bits

5. Research Method

The experimental research method was used to measure the effect of using selected varied and random pixels during the embedding process on imperceptibility. This method represents the standard practice applied in manipulating independent variables in order to analyze the generated data to test the research hypotheses. A notable advantage of experimental research is the fact that it enables other researchers to easily replicate the experiment and be able to validate the results. It is therefore considered an accurate method of research [7], as the researcher can effectively establish a causal relationship between variables by manipulating independent variable(s) to assess the effect upon dependent variable(s).

An experiment was carried out to test the relationship between the specific embedding process (i.e. proposed method) and the outcome (i.e. imperceptibility level). Essentially the output of traditional least significant bit steganography method was used to evaluate the performance and effectiveness of the proposed method's output by comparing the stego images generated by the proposed method with those generated by the traditional least significant bit steganography method. This is commonly referred to as comparative experiment [8].

6. Design

The design model framework is based on the generic model presented by Birgit Pfitzmann as shown Fig. 5

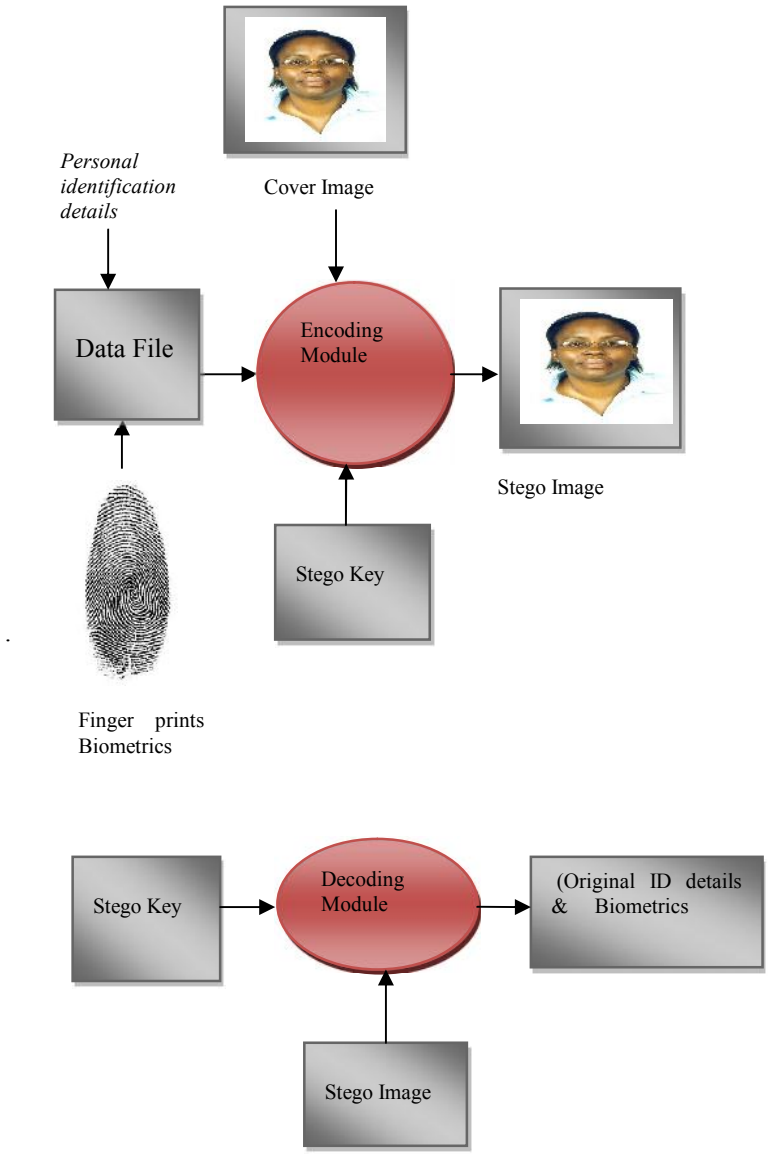

Fig. 5 Proposed Design framework [9]

The embedding algorithm is depicted in Fig. 6 below while the extraction algorithm is shown in Fig. 7

Input : Cover Image, Secret file (Payload)
Output : Stego image (image containing hidden file)
16. Use LCG to
Select a random pixel
Select a random pixel color channel
Select a random color channel bit
2. Let bitToWrite [x][y][channel][bit] denote the selected bit in a specific color channel for writing
3. Let m_i denote the message bit embedded in a color channel bit,
bitToWrite[x][y][channel][bit]
4. For all image color channels do the following
5. If LSB (bitToWrite[x][y][channel][bit]) = m_i then
6. do nothing
7. If LSB(bitToWrite[x][y][channel][bit]) not equal to m_i then
8. bitToWrite[x][y][channel][bit] = m_i
9. while secret file length; Repeat step 16 to 23 to embed the entire message
10. Close stream.

Fig. 6 The Proposed enhanced LSB method embedding

Input : Stego Image, Password message digest
Output : Secret file

1. Use LCG to
Select a random pixel
Select a random pixel color channel
Select a random color channel bit
2. Let bitToRead([x][y][channel][bit])denote the selected bit in a specific color channel for reading
3. Let m_i denote the message bit read in a color channel bit bitToRead([x][y][channel][bit])
4. For all image color channels do the following
5. If LSB(bitToRead([x][y][channel][bit]) not equal to m_i then
6. do nothing
7. If LSB (bitToRead([x][y][channel][bit]) = m_i then
8. bitToRead ([x][y][channel][bit])= m_i
9. Pack bit in bitSet
10. While secret file length; Repeat step 11 to 20 to read the entire file
11. Close stream.
12. Obtain the entire message stream and convert it back into ASCII format

Fig. 7 The Proposed enhanced LSB algorithm general extracting algorithm

7. Experimental Design and Testing

An analysis to examine the statistical properties of the stego images produced by the proposed method and the traditional LSB method was carried out. Statistical attacks are more powerful than visual attacks as they are able to reveal the tiniest modifications in the statistical properties of an image [10].
The following image quality metrics were employed for this purpose:

7.1 Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE)

Both of these metrics are the most common and widely used full reference metrics for objective image quality evaluation. In particular, PSNR is used in many image processing applications and considered as a reference model to evaluate the efficiency of other objective image quality evaluation methods [11]. PSNR as a metric computes the peak signal-to-noise ratio, in decibels, between two images. It is the ratio between the maximum possible power of an image signal and the power of corrupting noise that affects the fidelity of its representation. In the literature, PSNR has shown the best advantage almost over all other objective image quality metrics under different image distortion environments and strict testing conditions [12].
On the other hand, MSE measures the statistical difference in the pixel values between the original and the reconstructed image [13] [14]. The mean square error represents the cumulative squared error between the original image and the stego-image. PSNR and MSE are defined in equation (1) and equation (2) below [13] [14].

$$MSE = \left(\frac{1}{MN}\right) \sum_{i=1}^{M} \sum_{j=1}^{N} \left(X_{ij} - \overline{X_{ij}}\right)^2 \quad (1)$$

$$PSNR = 10.\log_{10} \frac{I^2}{MSE} db \quad (2)$$

Where:
X_{ij} is the i^{th} row and the j^{th} column pixel in the original (cover) image,
$\overline{X_{ij}}$ is the i^{th} row and the j^{th} column pixel in the reconstructed (stego) image,
M and *N* are the height and the width of the image,
I is the dynamic range of pixel values, or the maximum value that a pixel can take, for 8-bit images: I=255.

A lower MSE value means a better image quality ie lesser distortion in the cover image while the higher the PSNR value the better the quality of the image [15].

In order to evaluate the performance of the proposed method an experimental design was set up. Stego images from both the traditional LSB method and the proposed method were compared using the testing metrics discussed above. All the experiments were implemented and run on a PC Pentium IV Duo core, 2.1 GHz with 2GB of RAM under the Windows 7 Home Edition operating system. The following constants were ensured.

- Same images were used on both the methods
- Same information was embedded in each image ie equal payload
- Same evaluation metrics were used for each image
- Five digital pass port size face images were used as test data files (cover images). Table 1 shows the list of these digital images.

FILE NAME	DIMENSIONS	FILE SIZE
Alice.jpg	259 x 370 Pixels	20.5 Kilo Bytes
John.jpg	259 x 350 Pixels	9.42 Kilo Bytes
Loise.jpg	259 x 326 Pixels	13.1 Kilo Bytes
Mercy.jpg	241 x 370 Pixels	11.7 Kilo Bytes
Morris.jpg	259 x 303 Pixels	7.68 Kilo Bytes
Peter.jpg	259 x 262 Pixels	12.4 Kilo Bytes

Table 1: Test data Images

The specific data hiding steganogrtaphic method used was taken to be the independent variable (in this case the traditional LSB method and the proposed enhanced LSB method). In order to evaluate the efficiency of the proposed steganography method, the evaluation dependent variables which measure the image distortion levels were considered. Accordingly, for each steganography method (the traditional LSB method and the proposed enhanced LSB method) and for each cover image the value of each dependent variable was measured. The values of the dependent variables for both embedding methods were then compared. This is illustrated in table 2.

IMAGE	METRIC	LEAST SIGNIFICANT BIT (DB)	ENHANCED LEAST SIGNIFICANT BIT (DB)
Alice.jpg	PSNR	57.15	58.78
	MSE	1.12	0.77
John.jpg	PSNR	56.91	58.49
	MSE	1.19	0.82
Loise.jpg	PSNR	56.54	58.00
	MSE	1.28	0.91
Mercy.jpg	PSNR	55.96	57.42
	MSE	1.20	0.85
Morris.jpg	PSNR	56.26	57.65
	MSE	1.38	1.00
Peter.jpg	PSNR	55.36	56.52
	MSE	1.59	1.22

Table 2: Comparison of various values from the two algorithms

8. Experimental Results

8.1 Peak Signal to Noise Ratio (PSNR)

Fig. 8 below shows the comparison of the PSNR values of six stego images for both the traditional LSB method and the enhanced LSB method.

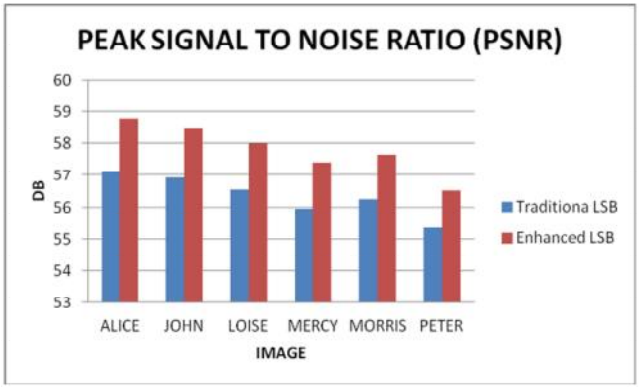

Fig. 8 The PSNR of stego images- Traditional LSB versus Enhanced LSB

Every image tested registered a higher PSNR for enhanced LSB method as compared to the Traditional LSB method showing that the enhanced LSB embedding method distorts the image less improving on imperceptibility of the hidden data since a higher Peak Signal to Noise Ratio (PSNR) indicates less distortion [15].

8.2 Mean Square Error (MSE)

Fig. 9 below shows a summary of the comparison of the MSE of five stego images for both the traditional LSB method and the enhanced LSB method.
For each stego image, a lower MSE was recorded with the enhanced LSB method as compared to the traditional LSB method. A lower MSE value means a better image quality ie lesser distortion in the cover image [15]. This means that stego images generated by the enhanced LSB method have lesser distortions compared to those generated by the traditional LSB method and hence improved imperceptibility.

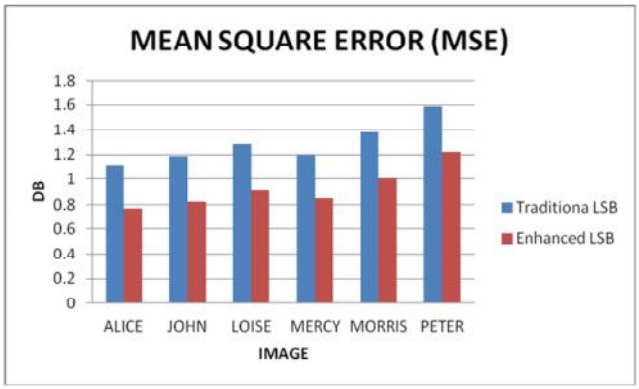

Fig. 9 The MSE of stego images - Traditional LSB versus Enhanced LSB

8.3 Histogram analysis

A histogram is an aggregation method that is used to illustrate data distribution in an image. The histogram is constructed by partitioning the data space into many small ranges, with each range corresponding to a bin. The height of an image histogram bin is then determined by the percentage of data points that fall in the corresponding range. This reveals the data density within each sub-range [16].
Fig. 10 to 12 below shows the comparison of histograms analysis for the image Alice.jpg. Compared to the original image, the standard deviation among pixels is least affected when the enhanced LSB method is used indicating increased imperceptibility.

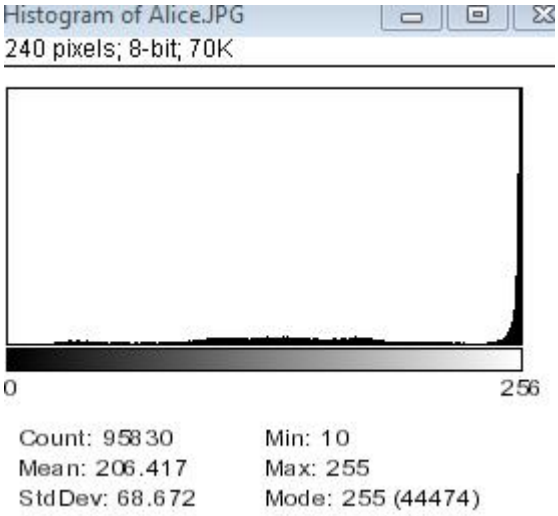

Fig. 10 Histogram for the original image before any data is hidden

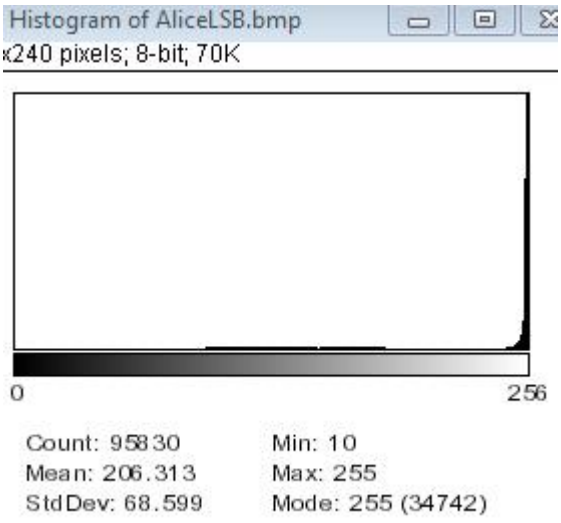

Fig. 11 Histogram for the stego image after embedding data using the LSB steganograpgic method.

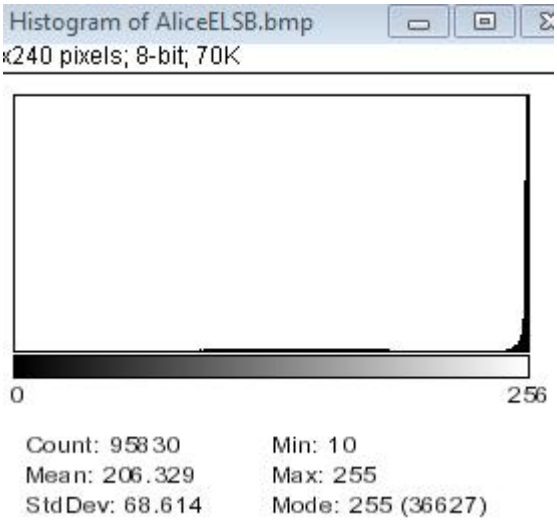

Fig. 12 Histogram for the stego image after embedding data using the proposed enhanced LSB steganograpgic method.

9. CONCLUSION AND FUTURE WORK

In this paper, a biometric electors' smart card based on an improved LSB steganographic method has been presented. The purpose is to embed the electors details and biometric data in his/her face image in order to provide a platform of authenticating a voter through a single document. These details are embedded in a more imperceptible manner as compared to the way a conventional LSB method would do. There is a demonstration of increased imperceptibility to statistical steganalysis attacks on the cover image as proved through the perceptibility metrics used.

Future research can explore on more permanent embedding methods that are stronger against steganalysis which cannot be destroyed through image manipulations.

References

[1] M.R.M.; Yahaya, Y.H.; Halip, M.H.M.; Khairuddin, M.A.; Maskat, K.;, "The design of fingerprint biometric authentication on smart card for PULAPOT main entrance system," Information Technology (ITSim), 2010 International Symposium in Vol.3, No., 2010, pp. 1-4, 15-17.

[2] K. I. Chang, K.. W. Bowyer, and P. J. Flynn. An Evaluation of Multimodal 2D+3D Face Biometrics. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.27, No.4, 2005, pp. 619-624.

[3] A.K. Jain, and Jianjiang Feng, "Lattent Fingerprint Matching", Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol.33, No.1, 2011 pp.88-100.

[4] F. Mohammad, and M. Abdallah, A Steganographic Data Security Algorithm with Reduced Steganalysis Threat, Birzeit University: Birzeit, (2008).

[5] Gabriel Macharia Kamau, Stephen Kimani, and Waweru Mwangi, "An enhanced Least Significant Bit Steganographic Method for Information Hiding ", Journal of Information Engineering and Applications, Vol. 2, No.9, 2012, pp. 1-12.

[6] T.E.Hull and A.R. Dobell, "Random Number Generators.", SIAM Review, vol.4, No.3,1962,pp.230-254.

[7] M. Shuttleworth, Experiment Resources. (2008), URL:http://www.experiment-resources.com. Accessed on 7th January 2013.

[8] K. Hinkelmann, and O. Kempthorne, Design and Analysis of Experiments: Introduction to Experimental Design, John Wiley & Sons, Inc: Hoboken,New Jersey, 2008.

[9] B. Pfitzmann, (1996). "Information Hiding Terminology", in Information Hiding: First International Workshop (R Anderson, ed), Lecture Notes in Computer Science, 1996, Vol. 1174, pp 347-350.

[10] D. Artz, (2001) "Digital steganography: hiding data within data", Internet Computing, IEEE, Vol. 5,No. 3, 2001, pp. 75-80

[11] Z. Wang, H.R. Sheikh, and A.C Bovik, "No-reference perceptual quality assessment of JPEG compressed images", Proceedings of the International Conference on Image Processing, Vol.1, 2002, pp 477-480.

[12] Z. Wang, A.C. Bovik, and L. Lu, "Why is image quality assessment so difficult?", IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '02),Vol. 4, 2002, pp.3313-3316.

[13] A.Stoica, C. Vertan, and C. Fernandez-Maloigne, "Objective and subjective color image quality evaluation for JPEG 2000 compressed images", International Symposium on Signals, Circuits and Systems, Vol. 1, 2003,pp.137-140.

[14] Z. Wang, H.R. Sheikh, and A.C. Bovik, Objective Video Quality Assessment. The Handbook of Video Databases: Design and Applications.CRC Press, 2003.

[15] J. Mei, S. Li, and X.Tan, "A Digital Watermarking Algorithm Based on DCT and DWT", in Proceedings of the 2009 International Symposium on Web Information Systems and Applications (WISA'09) Nanchang, P. R.

China, 2009,pp. 104- 107.

[16] N. Jacobsen, K. Solanki, U. Madhow, B.S Manjunath and S. Chandrasekaran, "Image-Adaptive high –volume data hiding based on scalar quantization," In proceedings of IEEE military communication conference (MILCOM), Anaheim CA, USA, 2002.