World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:9, No:10, 2015

Study of Qualitative and Quantitative Metric for Pixel Factor Mapping and Extended Pixel Mapping Method

Indradip Banerjee, Souvik Bhattacharyya, Gautam Sanyal

Abstract**—**In this paper, an approach is presented to investigate the performance of Pixel Factor Mapping (PFM) and Extended PMM (Pixel Mapping Method) through the qualitative and quantitative approach. These methods are tested against a number of well-known image similarity metrics and statistical distribution techniques. The PFM has been performed in spatial domain as well as frequency domain and the Extended PMM has also been performed in spatial domain through large set of images available in the internet.

Keywords**—**Qualitative, quantitative, PFM, EXTENDED PMM.

I. Introduction

SECURITY of information is one of the aspect of information technology and communication system. Computer, Laptop, etc. are the most important communication media that can join different parts of the world through internet. In contradiction, safety and security of communication systems are the challenging issue in day to day life. Consequently, the concept of "Information Hiding" [1] has been introduced by the researchers. After that, the theory of cryptography [2] and the concept of watermarking [3] have been established. But nowadays, by rising computational supremacy, regular cryptographic and watermarking algorithms have been established to be evidence for weak point against mathematical and statistical methods. The research on reverse engineering techniques has increased the processing power. So, the most important race between researches in cryptanalysis [4] and watermarking detection [5] has been developed. To solve the problems, the concept of steganography [6] has been established by the researchers. Steganography diverges from cryptography. Cryptography is a secure communication which changes the information into a specific form and so an eavesdropper cannot recognize it. Steganography techniques attempt to hide the existence of the message itself, so that an observer does not know that the data are even there or not.

Steganography gives the certification over the data using some tag or labelling on some objects like text, audio, video, image. Fig. 1 illustrates the types of steganography used in modern days.

Hiding the existence of the message and creating the covert channel is the objective of the Steganography. The message is hidden in another object. Consequently, the transmitted item will be identical, considering to every individual's eye. Steganalysis is the art of detecting any hidden data on the communication channel. Steganography is rumpled when the presence of the hidden message is uncovered. Fig. 2 shows the framework of the steganography.

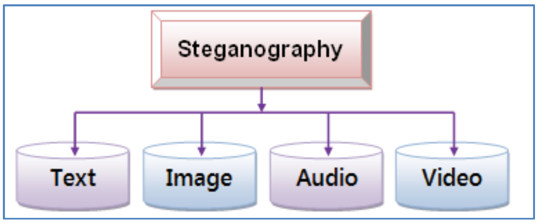

Fig. 1 Types of Steganography

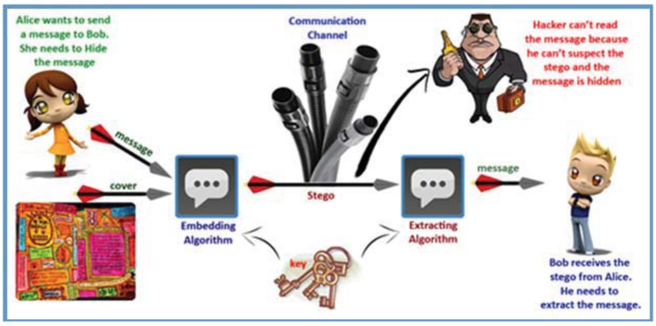

Fig. 2. Frame Work of Steganography

PFM [7], [8] is based on both spatial and frequency domains with the help of Gray scale images (512x512). More than 1000 images have been used, which have been collected from image databases that are used by most researchers. In this system, the input message (MSG_{BIT}) can be adopted as digital character format and cope with bit stream. Image converts from spatial domain ($f_{m,n}$) to frequency domain ($f_{x,y}$), in case of frequency domain method. Four-bit pair €(x, y) has been used for embedding in a separate DCT coefficient of frequency domain method or pixel intensity of spatial domain method. The selection of €(x, y) depends on one of the author's mathematical function. The selection technique is based on €(x, y) and its position on the image. Mapping of four bit from the secret message is used for embedding in a pixel's €(x, y) position, which is based on the maximum prime factor value (PFV) of image €(x, y) value. PFM Table (Table

Indradip Banerjee and Gautam Sanyal are with the National Institute of Technology, Durgapur, India (e-mail: indradip.banerjee@yahoo.com, nitgsanyal@gmail.com).
Souvik Bhattacharyya is with the University Institute of Technology, The University of Burdwan, Burdwan (e-mail: souvik.bha@gmail.com).

I) shows the PFV value of the corresponding 4-bit embedding character. After embedding in €(*x, y*), it is generating the stego €(x_1, y_1) and reconstruct the image value $f'_{x,y}$ converting the image ($f'_{x,y}$) from frequency to spatial domain image ($f'_{m,n}$) as stego in frequency domain method. Most of the existing methods can embed one bit per coefficient or pixel intensity, whereas this method can embed four bit per coefficient. Embedding four bit per coefficient can extend the embedding capacity four times than previous methods in spatial or frequency domain. At the receiver side, reverse operation has been carried out to get back the original information.

TABLE I
PFM TABLE

Prime Factor Value (PFV)	Embed 4 Bit	Prime Factor Value (PFV)	Embed 4 Bit
2	0000	23	1000
3	0001	29	1001
5	0010	31	1010
7	0011	37	1011
11	0100	41	1100
13	0101	43	1101
17	0110	47	1110
19	0111	53	1111

Extended PMM: This method is an extension of PMM method [9] which is based on special domain with the help of gray scale images (512x512). The input message can be adopted as digital character format. Two bit pair has been used for embedding in a separate pixel and the selection of embedding pixel also depends on mathematical function. The selection technique is based on pixel intensity value and pixel position on image. Mapping of two bit from the secret message is used for embedding in a pixel which is based on the intensity value, previous pixel value and number of the ones (in binary) present in that pixel. The existing PMM two-bit tables have been used here for mapping and the previous embedded pixel intensity value used for mapping. The concept of previous pixel intensity value is added in this work along with the embedding pixel selection method. Extraction process starts again by selecting the same reverse process required during embedding. At the receiver side, the reverse operation has been carried out to get back the original information.

II. APPLIED SIMILARITY METRICS

We explain the experimental results as well as the hiding performance of presented methods based on some similarity metrics, which are discussed below. Capacity of hiding data and imperceptibility of the stego image are tested here using various performance metrics. Imperceptibility of the image is called the quality of the image.

Qualitative Metrics

There is an inclusive range of approaches to qualitative metrics. A qualitative attitude is a general way of thinking about directing qualitative research. It defines, either explicitly or implicitly, the role of the researchers, stages of research and the method of data analysis are the purpose of the qualitative metrics. Qualitative research is designed to expose a target audience's range of behaviour and the observations that drive it with reference to specific matters or issues. It uses comprehensive studies of small groups of data to monitor and sustenance the construction of theories. The results of qualitative research are descriptive, predictive, collection of word data and question answers.

The quality of the stego image produced by the proposed methods and the stego image has been tested thorough statistical parameters like Mean, Standard Error Mean, Trimmed Mean, Standard Deviation, Variance, Coefficient of Variation, Sum, First quartile (Q_1), Median (Q_2), Third quartile (Q_3), Range, Interquartile (IQR), Mode, N for Mode, Kurtosis, MSSD and Covariance etc. The hypothesis tests i.e. the goodness of fit tests have also been used for better distribution with the help of Kolmogorov Smirnov Test. The rank and statistic value can measure the goodness of fit [10] and it also declares the best fitted distribution technique for the image. The details of the tests are discussed below:

Mean: The mean [11] is used to denote one measure of the central tendency either of a probability distribution or random variable which is characterized by distribution technique in statistics or probability. The mathematical expectation, arithmetic mean and average are used to refer a central value of a discrete data set. The sum of the values divided by the number of values is the mean. In case of finite population and by considering every member of that population, the population mean of a property is equal to the arithmetic mean of the given property. For example, the population mean weight is equal to the sum of the weights of every individual divided by the total number of individuals. For the small sample data set, the sample mean may diverge from the population mean.

Standard Error Mean: The standard error (SE) [12] of the mean is the standard deviation of a sampling distribution of a statistic, most commonly it also called the standard deviation of the mean. The term may also be used to refer to an estimate of that standard deviation, consequent from a particular sample used to compute the estimate.

Trimmed Mean: A truncated mean or trimmed mean [13] is a statistical degree of central tendency and this is like mean and median. It involves the calculation of the mean after disposal of given parts of a probability distribution or sample at the high and low end. This number of points to be cast-off is usually given as a percentage of the total number of points, but may also be given as a fixed number of points.

Standard Deviation: In the statistics, the standard deviation [14] is a measure that is used to quantify the amount of variation or diffusion of a data set. The standard deviation is represented by the Greek letter sigma (σ). A standard deviation close to 0 indicates that the data points incline to be very close to the mean, which is also called the expected value of the set. When it is indicating a high standard deviation then the data points are spread out over a wider range of values.

Variance: In probability theory and statistics, the variance [15] measures how extreme a set of numbers is spread out. When a variance is 0, then it indicates that all the values are identical. Variance is always positive values. A small variance indicates that the data points tend to be very close to the mean i.e. the expected value and hence it is adjacent to each other. While the variance is high then it indicates that the data points are very spread out around the mean and expansive from each other.

Coefficient of Variation: In probability theory and statistics, the coefficient of variation (CV) [16] is a standardized measure of spreading of a probability distribution. It is defined as the ratio of the mean and standard deviation. It is also known as the variation coefficient. The absolute value of the CV is sometimes known as relative standard deviation (RSD), which is expressed as a percentage. The coefficient of variation (CV) is defined as the ratio of the standard deviation to the mean.

Quartile: In descriptive statistics, the quartiles [17] of a ranked set of data values are the three points that divide the data set into four equal groups, each group comprising a quarter of the data. A quartile is a type of quantile. The first quartile (Q_1) is defined as the middle number between the smallest number and the median of the data set. The second quartile (Q_2) is the median of the data. The third quartile (Q_3) is the middle value between the median and the highest value of the data set.

In applications of statistics, the quartiles of a ranked set of data values are the four subsets whose boundaries are the three quartile points. Thus, an individual item might be described as being "in the upper quartile".

- **First quartile** (Labelled as Q_1) also called the lower quartile or the 25th percentile. It splits off the lowest 25% of data from the highest 75%.
- **Second quartile** (Labelled as Q_2) also called the median or the 50th percentile. This is the cuts of the data set in half.
- **Third quartile** (Labelled as Q_3) also called the upper quartile or the 75th percentile. This is splits off the highest 25% of data from the lowest 75%.
- **Interquartile range** (Labelled as IQR) is the difference between the upper and lower quartiles. So, the IQR = Q_3 - Q_1.

Range: In arithmetic, the range [18] of a set of data is the difference between the largest and smallest values. However, in imaginative statistics, this concept of range has a more difficult meaning. The range is the size of the smallest interval which contains all the data and provides an indication of statistical distribution. It is measured in the same units as the data. Since it only depends on two of the observations, it is most useful in representing the dispersion of small data sets.

Mode: The mode [19] is the value that appears most often in a set of data. The mode of a discrete probability distribution is the value *x* at which its probability mass function takes its maximum value. In other words, it is the value that is most likely to be sampled. The mode of a continuous probability distribution is the value *x* at which its probability density function has its maximum value. So, informally speaking, the mode is at the peak.

Kurtosis: In probability theory and statistics, kurtosis [20] comes from the Greek word i.e. κυρτός, kyrtos or kurtos, meaning "curved, arching". It is calculating any measure of the "peakedness" of the probability distribution of a real-valued random variable. In a similar way to the concept of skewness, kurtosisis is a descriptor of the shape of a probability distribution. For the skewness, there are different ways of quantifying it for a theoretical distribution and corresponding ways of estimating it from a sample from a population.

MSSD: The mean of the squared successive differences (MSSD) [21] is used as an approximation of variance. It is calculated by taking the sum of the differences between consecutive observations aligned, then taking the mean of that sum and dividing by two.

Covariance: In probability theory and statistics, how much two random variables change together is measured by covariance [22]. If the greater values of one variable mainly correspond with the greater values of the other variable, the same holds for the smaller values, i.e., the variables tend to show similar behaviour, the covariance is positive. In the opposite case, when the greater values of one variable mainly correspond to the smaller values of the other, i.e., the variables tend to show opposite behaviour, the covariance is negative. The sign of the covariance therefore shows the tendency in the linear relationship between the variables. The magnitude of the covariance is not easy to interpret. The normalized version of the covariance, the correlation coefficient, however, it is shown by the magnitude and strength of the linear relation.

Kolmogorov Smirnov Test: The Kolmogorov–Smirnov test [23] or KS test is a non-parametric test of the equality of continuous, one-dimensional probability distributions. KS test can be used to compare one or two sample with a reference probability distribution. The Kolmogorov–Smirnov statistic measures a distance between the cumulative distribution function of the reference distribution and the empirical distribution function of any sample. The KS test can also quantify a distance between the empirical distribution functions of two samples. The null distribution of this statistic is calculated under the null hypothesis that the sample is drawn from the reference distribution or that the samples are drawn from the same distribution.

The empirical distribution function F_n for independent and identically distributed random variables observations X_i is defined in (1):

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{[-\infty,x]}(X_i) \tag{1}$$

where $I_{[-\infty,x]}(X_i)$ is the indicator function, equal to 1 if $X_i \leq x$ and equal to 0 otherwise.

The Kolmogorov–Smirnov statistic for a given cumulative distribution function $F(x)$ is (2)

$$D_n = sup_x|F_n(x) - F(x)| \tag{2}$$

where sup_x is the supremum of the set of distances. By the Glivenko-Cantelli theorem, if the sample comes from distribution $F(x)$, then D_n converges to 0 almost surely in the limit when n goes to infinity. Kolmogorov strengthened this result, by effectively providing the rate of this.

Quantitative Metrics

The science and technology investigate the quantitative research by observing the systematic experiential analysis of observable indications via statistical, mathematical and computational techniques [24]. The objective of quantitative metrics is to develop the mathematical models and theories related to exact value of the information. In this observation, the measurement is very accurate because the relationship between data and mathematical expression is quantitative. Quantitative information is any data which can be represented as the numerical form such as the value of the similarity matrices, embedding percentages, etc. We analyse the data with the help of some similarity metrics like MSE, PSNR, Correlations, RMSE, SSIM, KL divergence, Entropy.

Mean Square Error (MSE): It is computed by averaging the squared intensity of the cover and stego image pixels [25]. Equation (3) shows the MSE:

$$MSE = \frac{1}{NM}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} e\left(m,n\right)^2 \quad (3)$$

where *NM* is the image size (*N x M*) and *e(m,n)* is the reconstructed image.

Peak Signal-to-Noise Ratio (PSNR): A mathematical measure of image quality is signal–to-noise ratio (SNR) [25], which is based on the pixel difference between two images [26]. The SNR measure the estimate of Stego image and cover image. PSNR is shown in (4):

$$PSNR = 10\log_{10}\frac{S^2}{MSE} \quad (4)$$

where, *S* stands for maximum possible pixel value of the image. If the PSNR is greater than 36 DB then the visibility looks same in between cover and stego image, so HVS not identified the changes.

Correlations: Pearson's correlation coefficient is widely used in statistical analysis as well as image processing [27]. Here this can be used to Cover and Stego images to find the difference between these two images. The correlation is shown in (5):

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})\cdot(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\cdot\sum_{i=1}^{n}(y_i-\bar{y})^2}} \quad (5)$$

The X_i and Y_i are the cover image and bar of X and Y are stego image positions.

Root Mean Square Error (RMSE): RMSE [28] is a frequently used measure of the difference between Cover and Stego Image values. These individual differences are called residuals and the RMSE aggregates them into a single measure of predictive power. The RMSE is shown in (6):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i}-X_{model,i})^2}{n}} \quad (6)$$

X_{obsi} and X_{modeli} are two image vectors i.e. cover and stego.

Structural Similarity Index (SSIM): Wang et. al [29] proposed Structural Similarity Index concept between original and distorted image [10]. The Stego and Cover images are divided into blocks of 8 x 8 and converted into vectors. Then two means, two standard derivations and one covariance value are computed. After that, the luminance, contrast and structure comparisons based on statistical values are computed. Then the SSIM is computed between Cover and Stego images. SSIM is shown in (7):

$$SSIM = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\mu_x^2 + \mu_y^2 + C_2\right)} \quad (7)$$

KL divergence: With the help of probability density function (PDF) for each Image (cover and stego) we estimate the Kullback-Leibler Divergence [30]. KL divergence is shown in (8):

$$D\left(p\,\|\,q\right) = \sum_x p\left(x\right)\log\frac{p\left(x\right)}{q\left(x\right)} \quad (8)$$

Entropy: Entropy is a measure of the uncertainty associated with a random variable [16]. Here, a 'message' means a specific realization of the random variable. It is shown in (9):

$$\Delta S = .\int\frac{dQ_{rev}}{T} \quad (9)$$

where, S is the entropy and T is the uniform thermodynamic temperature of a closed system divided into an incremental reversible transfer of heat into that system (dQ).

Cross Entropy: The cross entropy [31] between two probability distributions over the image intensity measures the average number of element needed to identify an event drawn from the image matrix. The cross entropy for the distributions p and q over a given image is defined as:

$$H(p,q) = E_p[-\log q] = H(p) + D_{KL}(p||q) \quad (10)$$

where $H(p)$ is the entropy of p, and $D_{KL}(p||q)$ is the Kullback-Leibler divergence of q from p.

III. Experimental Results

In this section, we deliberate the experimental results of the PFM method in spatial and frequency domain with the help of gray scale images, extended PMM method in spatial domain using also gray scale images. First thing is the capacity of hiding data and another is the imperceptibility of the stego image, also called the quality of stego image. This capacity and the quality of image are discussed here with the help of quantitative as well as qualitative nature of similarity metrics. A comparative study with some existing methods like LSB, PVD and the methods proposed by the author are also discussed in this section. Experimental results of stego images are computed based on the images taken form the freely available image database in the internet and used by most of the researchers.

Qualitative Metrics

The qualitative approach has been discussed by means of the relative error of statistical parameters with respect to the covering values. Relative errors are plotted in a graph and the qualitative phases i.e. Best, Good, Moderate, etc. are categorized. The transition occurring in the graph helps in dividing to the phase. The quality of the steganography approach is depending upon the corresponding relative error phase. The relative error is in between 0 and 1. When, the relative error is near to 0, the performance is better. It has been observed that the Extended PMM and PFM methods are robust and secure, which have been tested by various statistical parameters. So, Extended PMM and PFM methods both of them work altogether than LSB, PVD mechanisms.

Mathematical Schemes:

1. Estimate the relative error

$$e_{i,j} = \frac{|Cover_i - X_{i,j}|}{Cover_i} \quad (11)$$

where, $i = 1,2,3, \ldots\ldots\ldots, 17$; $j = 1,2,3$. All these $e_{i,j} \in [0,1]$. Here i denotes the statistical parameters and j denotes the steganography methods. X is denoting the statistical parameter values of steganography methods.

2. Demolish the matrix to form a sequence $\{e_l\}_{l=1}^{51}$, where $l = 3(i-1) = j$
3. Next sort these e_l in increasing order.
4. Plot the graph with the help of e_l and identify the phases $\varepsilon_{Best}, \varepsilon_{Good}, \varepsilon_{Moderate}, \varepsilon_{Not\ Acceptable}$. The phases are established in the graph by assuming the transition occurrence.

PFM Method: PFM method has been implemented in image spatial and frequency domain. The qualitative approach is described with the help of spatial domain technique.

Table II illustrates the relative error, where i is the statistical parameters and j is the steganography methods. In (11), the X is denoting the statistical parameter values of LSB, PVD and PFM. Fig. 3 shows the graph by assuming the transition occurrence.

The phases are identified by the help of mathematical function (11). If the relative error $e_{i,j}$ is less than or equal to ε_{Best} then it is considered as 'Best'. While if $e_{i,j}$ is greater than or equal to ε_{Good} and less than ε_{Best}, it is considered as 'Good'. If $e_{i,j}$ is greater than or equal to $\varepsilon_{Moderate}$ and less than ε_{Good} then it is considered as 'Moderate'. When $e_{i,j}$ is greater than or equal to $\varepsilon_{Not\ Acceptable}$ then it is considered as 'Not acceptable'. Table III describes the relative values of statistical parameters for PFM, LSB & PVD.

TABLE II
RELATIVE ERROR $e_{i,j}$ OF STATISTICAL PARAMETERS FOR LSB, PVD & PFM

$j \downarrow$ $i \rightarrow$ where $i = 1,2,\ldots,17$ $j = 1,2,3$

STATISTICAL PARAMETERS	LSB	PVD	PFM
Mean	0.5943	0.5943	0.0001
Standard Error Mean	0.2826	0.2826	0.0014
Trimmed Mean	0.5950	0.5950	0.0001
Standard Deviation	0.2816	0.2816	0.0012
Variance	0.6424	0.6422	0.0023
Coefficient of Variation	0.1963	0.1963	0.0000
Sum	0.5943	0.5943	0.0001
First quartile (Q_1)	0.7451	0.7451	0.0000
Median (Q_2)	0.5926	0.5926	0.0000
Third quartile (Q_3)	0.5096	0.5096	0.0000
Range	0.3631	0.3095	0.0238
Interquartile (IQR)	0.2830	0.2830	0.0000
Mode	0.5347	0.5347	0.0099
N for Mode	0.2111	0.2090	0.0852
Kurtosis	0.0000	0.0000	0.0119
MSSD	0.6413	0.6385	0.0623
Covariance	0.2814	0.2814	0.0002

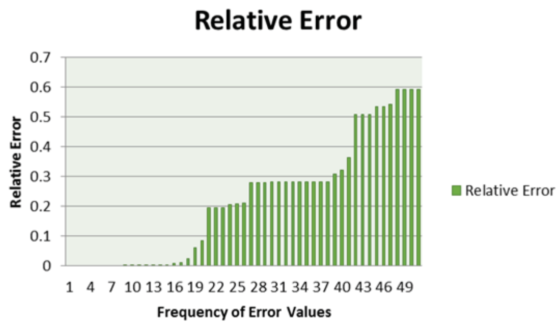

Fig. 3 Relative error graph for LSB, PVD & PFM

TABLE III
RELATIVE VALUES OF STATISTICAL PARAMETERS FOR PFM, LSB & PVD

STATISTICAL PARAMETERS	PFM	LSB	PVD
Mean	Best	Moderate	Moderate
Standard Error Mean	Best	Good	Good
Trimmed Mean	Best	Moderate	Moderate
Standard Deviation	Best	Good	Good
Variance	Best	Moderate	Moderate
Coefficient of Variation	Best	Best	Best
Sum	Best	Moderate	Moderate
First quartile (Q_1)	Best	Moderate	Moderate
Median (Q_2)	Best	Moderate	Moderate
Third quartile (Q_3)	Best	Moderate	Moderate
Range	Best	Moderate	Moderate
Interquartile (IQR)	Best	Good	Good
Mode	Best	Moderate	Moderate
N for Mode	Best	Good	Good
Kurtosis	Best	Best	Best
MSSD	Best	Moderate	Moderate
Covariance	Best	Good	Good

Kolmogorov Smirnov Test: By this hypothesis test, it can be observed that the performance of both PFM and the LSB are better by the help of the relative error of their Rank and Test parameter value for Burr, Log-Gamma, Lognormal, Beta, Gamma, Cauchy and Normal distribution. Also, it has been noted that performance of both PVD and PFM are same for Burr, Gamma, Cauchy and Normal distribution. The table of relative test results is furnished in Table IV. Hence, it can be proven that the PFM is robust and secure like LSB, PVD and the PFM is also acceptable for the hiding concept.

TABLE IV
RELATIVE VALUES OF HYPOTHESIS TEST RESULT FOR PFM, LSB & PVD

Kolmogorov Smirnov Test:			
Distribution	**PFM**	**LSB**	**PVD**
Burr	Good	Good	Good
Log-Gamma	Moderate	Moderate	Best
Lognormal	Moderate	Moderate	Best
Beta	Moderate	Moderate	Best
Gamma	Good	Good	Good
Cauchy	Good	Good	Good
Normal	Good	Good	Good

TABLE V
RELATIVE ERROR $e_{i,j}$ OF STATISTICAL PARAMETERS FOR LSB, PVD & PFM

$j \downarrow$ $i \rightarrow$ where $i = 1, 2, \ldots, 17$ $j = 1, 2, 3$

STATISTICAL PARAMETERS	**LSB**	**PVD**	**Ex PMM**
Mean	0.5943	0.5943	0.5928
Standard Error Mean	0.2826	0.2826	0.2826
Trimmed Mean	0.5950	0.5950	0.5936
Standard Deviation	0.2816	0.2816	0.2816
Variance	0.6424	0.6422	0.6422
Coefficient of Variation	0.1963	0.1963	0.1956
Sum	0.5943	0.5943	0.5929
First quartile (Q_1)	0.7451	0.7451	0.7451
Median (Q_2)	0.5926	0.5926	0.5926
Third quartile (Q_3)	0.5096	0.5096	0.5096
Range	0.3631	0.3095	0.3214
Interquartile (IQR)	0.2830	0.2830	0.2830
Mode	0.5347	0.5347	0.5446
N for Mode	0.2111	0.2090	0.2062
Kurtosis	0.0000	0.0000	0.0000
MSSD	0.6413	0.6385	0.6533
Covariance	0.2814	0.2814	0.2810

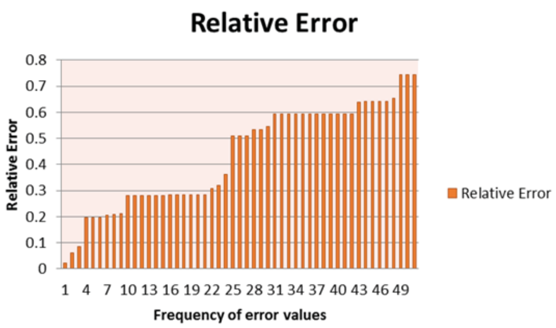

Fig. 4 Relative error graph for LSB, PVD & Extended PMM

Extended PMM Method: This method has been implemented in image spatial domain. The qualitative approach is described with the help of relative error.

Table V illustrates the relative error, where i denotes the statistical parameters and j denotes the steganography methods. In (11), X denote the statistical parameter values of LSB, PVD and Extended PMM. Fig. 4 shows the graph by assuming the occurrence of qualitative transitions.

The phases are identified by the help of mathematical function in (11). If $e_{i,j}$ is less than or equal to ε_{Best} then it is considered as 'Best'. If $e_{i,j}$ is greater than or equal to ε_{Good} and less than ε_{Best}, it is considered as 'Good'. If the $e_{i,j}$ is greater than or equal to $\varepsilon_{Moderate}$ and less than ε_{Good} then it is considered as 'Moderate'. When $e_{i,j}$ is greater than or equal to $\varepsilon_{Not\ Acceptable}$ then it considered as 'Not acceptable'. Table VI describes the relative values of statistical parameters for Extended PMM, LSB & PVD.

TABLE VI
RELATIVE VALUES OF STATISTICAL PARAMETERS FOR LSB, PVD & EX PMM

STATISTICAL PARAMETERS	**LSB**	**PVD**	**Ex PMM**
Mean	Good	Good	Good
Standard Error Mean	Good	Good	Good
Trimmed Mean	Good	Good	Good
Standard Deviation	Good	Good	Good
Variance	Moderate	Moderate	Moderate
Coefficient of Variation	Best	Best	Best
Sum	Good	Good	Good
First quartile (Q_1)	Moderate	Moderate	Moderate
Median (Q_2)	Good	Good	Good
Third quartile (Q_3)	Good	Good	Good
Range	Good	Good	Good
Interquartile (IQR)	Good	Good	Good
Mode	Good	Good	Good
N for Mode	Good	Good	Good
Kurtosis	Best	Best	Best
MSSD	Moderate	Moderate	Moderate
Covariance	Good	Good	Good

TABLE VII
RELATIVE VALUES OF HYPOTHESIS TEST RESULT FOR EX PMM, LSB & PVD

Kolmogorov Smirnov Test:			
Distribution	**LSB**	**PVD**	**Ex PMM**
Burr	Good	Good	Best
Log-Gamma	Moderate	Best	Best
Lognormal	Moderate	Best	Best
Beta	Moderate	Best	Best
Gamma	Good	Good	Best
Cauchy	Good	Good	Best
Normal	Good	Good	Best

Kolmogorov Smirnov Test: In this hypothesis test, it has been observed that the Extended PMM performs better than the LSB and PVD by means of relative error of their Rank and Test parameter value for Burr, Log-Gamma, Lognormal, Beta, Gamma, Cauchy and Normal distribution. The tabular form of relative test results is furnished in Table VII. Hence, from that point of view, it can be proved that the Extended PMM is robust and secure than LSB, PVD and the PFM is also acceptable for the hiding concept.

Quantitative Metrics

In qualitative approach, the methods have been tested through some techniques and the assessment is marked. But in the quantitative approach, the metrics are proving the robustness of methods with the help of quantitative values. The comparison with other methods through quantitative metrics are also substantiate that the developed techniques are superior to others methods.

Experimental Results of PFM Method: This section describes the results and analysis of PFM method through the quantitative nature. This method has been developed in spatial as well as frequency domain. Both are described below separately.

Spatial Domain: The similarity metrics of spatial domain technique with gray scale images are described in Tables VIII-XI.

Frequency Domain: The similarity metrics of frequency domain technique with gray scale images are described in Tables XII-XV.

TABLE VIII
MSE AND PSNR OF PFM, LSB & PVD

	MSE			PSNR		
Message Length	PFM	LSB	PVD	PFM	LSB	PVD
500	0.0285	0.018	0.0138	58.588	65.584	66.717
1000	0.0559	0.0375	0.0277	56.66	62.392	63.712
2000	0.1098	0.0707	0.0551	54.724	59.636	60.722
5000	0.2714	0.1737	0.1391	53.794	55.732	56.699
10000	0.4379	0.3401	0.2917	52.012	52.815	53.482
15000	0.6206	0.514	0.474	50.989	51.021	51.373
20000	0.8101	0.6844	0.6633	49.712	49.778	49.914
30000	1.0962	1.0388	1.1542	46.048	47.965	47.508

TABLE XI
CORRELATION AND RMSE OF PFM, LSB & PVD

	Correlations			RMSE		
Message Length	PFM	LSB	PVD	PFM	LSB	PVD
500	1	1	0.999997	0.033	0.097	0.096
1000	0.99999	0.99999	0.999994	0.046	0.13	0.135
2000	0.99998	0.99998	0.999988	0.125	0.188	0.19
5000	0.99997	0.99996	0.99997	0.297	0.29	0.298
10000	0.99991	0.99993	0.999937	0.526	0.392	0.433
15000	0.99987	0.99989	0.999897	0.727	0.487	0.546
20000	0.99977	0.99985	0.999856	0.871	0.564	0.644
30000	0.99964	0.99977	0.999751	0.967	0.691	0.838

TABLE X
SSIM AND CROSS ENTROPY OF PFM, LSB & PVD

	SSIM			Cross Entropy		
Message Length	PFM	LSB	PVD	PFM	LSB	PVD
500	1	1	1	4.35E+01	17.52882	13.55991
1000	1	1	1	85.01501	43.0382	27.3302
2000	1	1	1	166.9952	126.4012	55.85684
5000	1	1	0.999	425.8677	281.8571	147.6679
10000	0.999	1	0.999	884.1487	582.9952	331.6544
15000	0.997	1	0.998	1403.106	855.6058	554.9409
20000	0.997	1	0.998	2572.633	1111.579	753.9296
30000	0.996	0.999	0.997	3630.424	1594.039	1453.505

TABLE XI
ENTROPY AND KL DIVERGENCE OF PFM, LSB & PVD

	Entropy			KL Divergence		
Message Length	PFM	LSB	PVD	PFM	LSB	PVD
500	7.4405	7.4451	7.445	7.18E-07	5.41E-07	4.17E-07
1000	7.4409	7.4452	7.4451	1.40E-06	1.32E-06	8.41E-07
2000	7.4415	7.4455	7.4452	2.75E-06	3.95E-06	1.72E-06
5000	7.4419	7.4458	7.4453	6.98E-06	8.76E-06	4.53E-06
10000	7.4424	7.4461	7.4457	1.44E-05	1.81E-05	1.02E-05
15000	7.4438	7.4465	7.4458	2.26E-05	2.65E-05	1.70E-05
20000	7.4457	7.4467	7.4462	3.17E-05	3.44E-05	2.30E-05
30000	7.4469	7.4474	7.447	4.83E-05	4.92E-05	4.44E-05

TABLE XII
MSE AND PSNR OF PFM, LSB & PVD

Message Length	MSE			PSNR		
	PFM	LSB	PVD	PFM	LSB	PVD
100	0.188564	0.017975	0.013847	55.37621	65.58415	66.71714
200	0.370255	0.037483	0.027664	52.4458	62.39244	63.71162
500	0.993813	0.070717	0.055069	48.15776	59.63557	60.72173
1000	2.106766	0.173714	0.139065	44.89464	55.73246	56.69863
2000	4.327614	0.340054	0.291687	41.76832	52.81533	53.48163
3000	6.621696	0.514008	0.47398	39.92111	51.02111	51.3732
4000	8.925777	0.684372	0.663311	38.62434	49.77788	49.91363
5000	11.20823	1.038841	1.154232	37.63543	47.96531	47.50787

TABLE XIII
CORRELATION AND RMSE OF PFM, LSB & PVD

Message Length	Correlations			RMSE		
	PFM	LSB	PVD	PFM	LSB	PVD
100	0.999954	0.999996	0.999997	0.246063	0.097284	0.096339
200	0.999909	0.999992	0.999994	0.341054	0.129864	0.135359
500	0.999756	0.999985	0.999988	0.540367	0.188201	0.189765
1000	0.999483	0.999962	0.99997	0.76879	0.289952	0.297812
2000	0.998939	0.999926	0.999937	1.094649	0.391805	0.432775
3000	0.998379	0.999888	0.999897	1.346844	0.486556	0.546174
4000	0.997817	0.999851	0.999856	1.55667	0.563909	0.643951
5000	0.997262	0.999773	0.999751	1.743785	0.691398	0.838396

TABLE XIV
SSIM AND CROSS ENTROPY OF PFM, LSB & PVD

Message Length	SSIM			Cross Entropy		
	PFM	LSB	PVD	PFM	LSB	PVD
100	1	1	0.999999	3.00E-06	17.52882	13.55991
200	1	1	0.999985	5.96E-06	43.0382	27.3302
500	1	1	0.999858	1.65E-05	126.4012	55.85684
1000	1	0.999999	0.999453	3.53E-05	281.8571	147.6679
2000	0.999998	0.999988	0.998927	7.35E-05	582.9952	331.6544
3000	0.999992	0.999928	0.998459	0.000112	855.6058	554.9409
4000	0.999985	0.999778	0.998056	0.000151	1111.579	753.9296
5000	0.999976	0.999086	0.997261	0.000188	1594.039	1453.505

Experimental Results of PMM Method: This section describes the results and analysis of Extended PMM through the quantitative nature. This is a spatial domain technique using gray scale images. Tables XVI-XIX show the quantitative metrics value.

TABLE XV
ENTROPY AND KL DIVERGENCE OF PFM, LSB & PVD

Message Length	Entropy			KL Divergence		
	PFM	LSB	PVD	PFM	LSB	PVD
100	6.677885	7.44512	7.445017	32096	5.41E-07	4.17E-07
200	6.677695	7.445197	7.445057	32096	1.32E-06	8.41E-07
500	6.678876	7.445477	7.445153	3.21E+04	3.95E-06	1.72E-06
1000	6.681039	7.445803	7.445341	3.21E+04	8.76E-06	4.53E-06
2000	6.68535	7.446146	7.445717	3.21E+04	1.81E-05	1.02E-05
3000	6.689584	7.446484	7.445837	32096	2.65E-05	1.70E-05
4000	6.693826	7.446715	7.44616	32096	3.44E-05	2.30E-05
5000	6.699424	7.447361	7.44702	3.21E+04	4.92E-05	4.44E-05

TABLE XVI
MSE AND PSNR OF EXTENDED PMM, LSB & PVD

Message Length	MSE			PSNR		
	Extended PMM	LSB	PVD	Extended PMM	LSB	PVD
500	0.019165	0.017975	0.013847	65.30571	65.58415	66.71714
1000	0.038563	0.037483	0.027664	62.26912	62.39244	63.71162
2000	0.076477	0.070717	0.055069	59.29549	59.63557	60.72173
5000	0.188751	0.173714	0.139065	55.37191	55.73246	56.69863
10000	0.381187	0.340054	0.291687	52.31942	52.81533	53.48163
15000	0.568932	0.514008	0.47398	50.5802	51.02111	51.3732
20000	0.684212	0.684372	0.663311	49.7789	49.77788	49.91363

TABLE XVII
CORRELATION AND RMSE OF EXTENDED PMM, LSB & PVD

	Correlations			RMSE		
Message Length	Extended PMM	LSB	PVD	Extended PMM	LSB	PVD
500	0.999993	0.999996	0.999997	0.10298	0.097284	0.096339
1000	0.999986	0.999992	0.999994	0.145491	0.129864	0.135359
2000	0.999973	0.999985	0.999988	0.206792	0.188201	0.189765
5000	0.999932	0.999962	0.99997	0.323471	0.289952	0.297812
10000	0.999863	0.999926	0.999937	0.45846	0.391805	0.432775
15000	0.999796	0.999888	0.999897	0.560833	0.486556	0.546174
20000	0.999755	0.999851	0.999856	0.61435	0.563909	0.643951

TABLE XVIII
SSIM AND CROSS ENTROPY OF EXTENDED PMM, LSB & PVD

Message Length	SSIM			Cross Entropy		
	Extended PMM	LSB	PVD	Extended PMM	LSB	PVD
500	0.999989	1	0.999999	64.49798	17.52882	13.55991
1000	0.999953	1	0.999985	111.3092	43.0382	27.3302
2000	0.999874	1	0.999858	171.3853	126.4012	55.85684
5000	0.999619	0.999999	0.999453	415.9149	281.8571	147.6679
10000	0.999318	0.999988	0.998927	2762.02	582.9952	331.6544
15000	0.998958	0.999928	0.998459	4956.63	855.6058	554.9409
20000	0.998702	0.999778	0.998056	5243.456	1111.579	753.9296

TABLE XIX
ENTROPY AND KL DIVERGENCE OF EXTENDED PMM, LSB & PVD

	Entropy			KL Divergence		
Message Length	Extended PMM	LSB	PVD	Extended PMM	LSB	PVD
500	7.088114	7.44512	7.445017	3.19E-06	5.41E-07	4.17E-07
1000	7.088101	7.445197	7.445057	5.49E-06	1.32E-06	8.41E-07
2000	7.08801	7.445477	7.445153	8.43E-06	3.95E-06	1.72E-06
5000	7.087567	7.445803	7.445341	2.04E-05	8.76E-06	4.53E-06
10000	7.087073	7.446146	7.445717	NA	1.81E-05	1.02E-05
15000	7.086213	7.446484	7.445837	NA	2.65E-05	1.70E-05
20000	7.085546	7.446715	7.44616	NA	3.44E-05	2.30E-05

IV. CONCLUSION

In this article, we investigated the performance of the PFM and Extended PMM method in various domains using various image similarity measure metrics. The approach of quantitative and qualitative similarity metrics is furnished in this method. A comparative study also has been shown with some other existing methods like LSB & PVD. From the experimental results, it can be seen that the embedding capacity of the presented methods are much better compared to existing methods. With the computation of various images, quantitative and qualitative similarity metrics have been shown through some figures for measuring the similarity between the cover image as well as stego image and these methods give a good result.

REFERENCES

[1] Abbas Cheddad, Joan Condell, Kevin Curran, Paul Mc Kevitt, Digital image steganography: Survey and analysis of current methods Signal Processing 90, 2010, pp. 727– 752.
[2] Eskicioglu, A.M., Litwin, L. "Cryptography", Potentials, IEEE (Volume: 20, Issue: 1), Feb/Mar 2001, Page(s): 36 - 38, ISSN: 0278-6648
[3] Fu-Hau Hsu, Min-Hao Wu, Shiuh-Jeng Wang, "Dual-watermarking by QR-code Applications in Image Processing", 2012 9th International Conference on Ubiquitous Intelligence and Computing and 9th International Conference on Autonomic and Trusted Computing.
[4] Kaminsky, A., Kurdziel, M., Radziszowski, S. "An overview of cryptanalysis research for the advanced encryption standard", Military Communications Conference, 2010 - MILCOM 2010, San Jose, CA, Oct.31 2010-Nov.3 2010, Page(s): 1310 - 1316, ISSN: 2155-7578
[5] Wenjun Zeng; Liu, B."A statistical watermark detection technique without using original images for resolving rightful ownerships of digital images" Image Processing, IEEE Transactions on (Volume:8, Issue: 11), Nov 1999, Page(s):1534 - 1548, ISSN: 1057-7149.
[6] N.N. EL-Emam, "Hiding a large amount of data with high security using steganography algorithm," Journal of Computer Science, vol. 3, no. 4, pp. 223-232, April 2007.
[7] Indradip Banerjee, Souvik Bhattacharyya and Gautam Sanyal, "Study and Analysis of Steganography with Pixel Factor Mapping (PFM) Method", International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 8, August 2013.
[8] Indradip Banerjee, Souvik Bhattacharyya and Gautam Sanyal, "Robust image steganography with pixel factor mapping (PFM) technique", Computing for Sustainable Global Development (INDIACom), 2014 International Conference on 5-7 March 2014. Page(s): 692 - 698, Print ISBN: 978-93-80544-10-6. Publisher: IEEE Xplore Digital Library.
[9] Indradip Banerjee, Souvik Bhattacharyya and Gautam Sanyal. "Hiding & Analyzing Data in Image Using Extended PMM". In Proceedings of Computational Intelligence: Modeling, Techniques and Applications (CIMTA- 2013), Kalyani University, Kalyani, West Bengal, India, September 27-28, 2013. Proceedings published in "Procedia Technology, Elsevier".
[10] Hossein Mehrannia & Alireza Pakgohar. "Using Easy Fit Software for Goodness-Of-Fit Test and Data Generation" International Journal of Mathematical Archive-5(1), 2014, Page: 118-124.
[11] Feller, William (1950). Introduction to Probability Theory and its Applications, Vol I. Wiley. p. 221. ISBN 0471257087
[12] Everitt, B.S. (2003) The Cambridge Dictionary of Statistics, CUP. ISBN 0-521-81099-X.
[13] Arulmozhi, G.; Statistics for Management, 2nd Edition, Tata McGraw-Hill Education, 2009, p. 458.
[14] Bland, J.M.; Altman, D.G. (1996). "Statistics notes: measurement error." (PDF). Bmj, 312(7047), 1654. Retrieved22 November 2013.
[15] Yuli Zhang, Huaiyu Wu, Lei Cheng (June 2012). Some new deformation formulas about variance and covariance. Proceedings of 4th International Conference on Modelling, Identification and Control (ICMIC2012). pp. 987–992.
[16] Broverman, Samuel A. (2001). Actex study manual, Course 1, Examination of the Society of Actuaries, Exam 1 of the Casualty

Actuarial Society (2001 ed. ed.). Winsted, CT: Actex Publications. p. 104. ISBN 9781566983969. Retrieved 7 June 2014.

[17] Hyndman, Rob J; Fan, Yanan (November 1996). "Sample quantiles in statistical packages". American Statistician 50 (4): 361–365. doi:10.2307/2684934.

[18] George Woodbury (2001). An Introduction to Statistics. Cengage Learning. p. 74. ISBN 0534377556.

[19] Hippel, Paul T. von (2005). "Mean, Median, and Skew: Correcting a Textbook Rule". J. of Statistics Education 13 (2).

[20] Joanes, D. N.; Gill, C. A. (1998). "Comparing measures of sample skewness and kurtosis". Journal of the Royal Statistical Society (Series D): The Statistician 47 (1): 183–189.

[21] “The Mean Square Successive Difference” J. von Neumann, R. H. Kent, H. R. Bellinson and B. I. Hart; The Annals of Mathematical Statistics, Vol. 12, No. 2 (Jun., 1941), pp. 153-162.

[22] Yuli Zhang, Huaiyu Wu,Lei Cheng (June 2012). Some new deformation formulas about variance and covariance. Proceedings of 4th International Conference on Modelling, Identification and Control (ICMIC2012). pp. 987–992.

[23] "Academician Andrei Nikolaevich Kolmogorov (obituary)". Russian Mathematical Surveys 43: 1–9. 1988.

[24] Given, Lisa M. (2008). The Sage encyclopedia of qualitative research methods. Los Angeles, Calif.: Sage Publications. ISBN 1-4129-4163-6.

[25] Yusra A. Y. Al-Najjar, Dr. Der Chen Soong, "Comparison of Image Quality Assessment: PSNR, HVS, SSIM, UIQI", International Journal of Scientific & Engineering Research, Volume 3, Issue 8, August-2012 ISSN 2229-5518

[26] a. L. M. Jean-Bernard Martens, "Image dissimilarity", Signal Processing, vol. 70, no. 3, pp. 155-176, 1998.

[27] J.L. Rodgers, J.L. and W.A. Nicewander, "Thirteen Ways to Look at the Correlation Coefficient", American Statistician 42, 59-66 (1995).

[28] Lehmann, E. L.; Casella, George (1998). Theory of Point Estimation (2nd ed.). New York: Springer. ISBN 0-387-98502-6.

[29] A. A. C. B. Zhou Wang, "A Universal Image Quality Index," IEEE Signal Processing Letters, vol. 9, pp. 81-84, 2002.

[30] Pedro J. Moreno, Purdy Ho, Nuno Vasconcelos "A Kullback-Leibler Divergence Based Kernel for SVM Classification in Multimedia Applications" Conference: Neural Information Processing Systems - NIPS, 2003

[31] Pieter-Tjerk De Boer, Dirk P. Kroese, Shie Mannor, Reuven Y. Rubinstein. "A tutorial on the cross-entropy method." Annals of operations research of Springer Science + Business Media, Inc. Manufactured in The Netherlands. 134.1 (2005): 19-67.

Indradip Banerjee is working at University Institute of Technology, Burdwan University, Burdwan, West Bengal, India. He received his MCA degree in 2009, PGDCA in 2008, MMM in 2005 and BCA (Hons.) in 2003. He has submitted his Ph.D. in Engineering at Computer Science and Engineering Department, National Institute of Technology, Durgapur, West Bengal, India in 2016. His areas of interest are Biometric Information Security, Steganography, Cryptography, Text Steganography, Image Steganography, Quantum Steganography and Steganalysis. He has published 30 research papers in International and National Journals / Conferences.

Souvik Bhattacharyya received his B.E. degree in Computer Science and Technology from B.E. College, Shibpur, India and M.Tech degree in Computer Science and Engineering from National Institute of Technology, Durgapur, India. He has received Ph.D (Engg.) from National Institute of Technology, Durgapur, India. Currently he is working as an Assistant Professor and In-Charge in Computer Science and Engineering Department at University Institute of Technology, The University of Burdwan. His areas of interest are Natural Language Processing, Network Security and Image Processing. He has published nearly 56 papers in International and National Journals / Conferences.

Gautam Sanyal has received his B.E and M.Tech degree National Institute of Technology (NIT), Durgapur, India. He has received Ph.D (Engg.) from Jadavpur University, Kolkata, India, in the area of Robot Vision. He possesses an experience of more than 25 years in the field of teaching and research. He has published nearly 150 papers in International and National Journals / Conferences. Two Ph.Ds (Engg) have already been awarded under his guidance. At present he is guiding six Ph.Ds scholars in the field of Steganography, Cellular Network, High Performance Computing and Computer Vision. He has guided over 10 PG and 100 UG thesis. His research interests include Natural Language Processing, Stochastic modeling of network traffic, High Performance Computing, Computer Vision. He is presently working as a Professor in the department of Computer Science and Engineering and also holding the post of Dean (Students' Welfare) at National Institute of Technology, Durgapur, India.