International Journal of Engineering Technology and Applied Science

(ISSN: 2395 3853), Vol. 2 Issue 9 September 2016

"Spatial vector quantization in steganography technique"

Bhupendra Prajapati[1], Pinkeshwar Mishra[2] & Keshav Tiwari[3]

M. TECH Department of CSE, Swmivivekanand College of Science & Technology, Bhopal [1]
Head of the Department, Computer Science, Swamivivekanand College of Science & Technology, Bhopal [2]
Assistant Professor, Computer Science, Swamivivekanand College of Science & Technology, Bhopal [3]

Abstract— **Steganography is the science that involves communicating secret data in an appropriate multimedia carrier, e.g., image, audio and video files. It comes under the assumption that if the feature is visible, the point of attack is evident, thus the goal here is always to conceal the very existence of the embedded data. It does not replace cryptography but rather boosts the security using its obscurity features. Steganography has various useful applications. However, like any other science it can be used for ill intentions. It has been propelled to the forefront of current security techniques by the remarkable growth in computational power, the increase in security awareness.**

This paper investigates current state-of-the-art methods and provides a new and efficient approach to digital image steganography. It also establishes a robust steganographic system.

Keywords: steganography, vector quantization, Data hiding, Encryption.

1. INTRODUCTION

Steganography is the art and science of invisible communication. This is consummate through a hiding information in other information, thus hiding the existence of the communicated information. The word steganography is obtained from the Greek words "stegos" meaning "cover" and "grafia" meaning "writing" defining it as "covered writing". In image steganography the information is hidden only in images.

Steganography and cryptography are cousins in the secret activity family. Cryptography clamber a message by using definite cryptographic algorithms for converting the secret data into cryptic form. On the other hand, Steganography secret the message so that it cannot be seen. A message in cipher text might cause impression on the part of the recipient while an "invisible" message created with steganographic methods will not. Anyone attractive in secret communication can always apply a cryptographic algorithm to the data before embedding it to achieve further security. In any case, once the existence of hidden information is disclosed or even conjecture, the purpose of steganography is overcome, even if the message content is not extracted or deciphered.

"Steganography's niche in security is to additive cryptography, not replace it. If a hidden message is encrypted, it must also be decrypted if locate, which provides another surface of protection." Another form of data hiding in digital images is Watermarking. Digital watermarking is the process of inserting supplementary information into a digital cover signal with the aim of providing authentication information. A watermark is called durable with respect to a class of transformations if the embedded information can authentic be detected from the marked signal even if degraded by any transformation within that class. Typical image mortification is JPEG compression, rotation, cropping, additive noise and quantization.

Steganography and watermarking vary in a number of ways including purpose, specification and detection/extraction methods. The most basic difference is that the object of communication in watermarking is the host signal, with the insert data providing copyright protection. In steganography the object to be transmitted is the inserted message, and the cover signal serves as a harmless mimic chosen fairly arbitrarily by the user based on its technical suitability.

In addition, the existence of the watermark is frequently declared openly, and any attempt to remove or invalidate the inserted content contributes the host useless. The critical requirement for steganography is undying and algorithmic un detectability. Robustness against bitter attack and signal processing is not the primary cover, as it is for watermarking. The difference between Steganography and Watermarking with respect the three parameters of payload, un detectability and robustness can be understood from Figure 1.2

As initiate steganography assigns with hiding of information in some cover source. On the other hand, Steganalysis is the art and science of detecting messages secret using steganography; this is similar to cryptanalysis applied to cryptography. The goal of steganalysis is to identify conjecture packages, determine whether or not they have a payload encoded into them, and, if possible, recuperate that payload. Hence, the major challenges of effectual steganography are: -

1. **Security of Hidden Communication:** In order to circumvent raising the instinct of eavesdroppers, while evading the demanding screening of algorithmic detection, the secret contents must be invisible both perceptually and statistically.

2. **Size of Payload**: Unlike watermarking, which of necessity to inserted only a small amount of copyright information, steganography focus at secret communication and therefore usually requires enough embedding capacity? Requirements for higher payload and secure communication are frequently inconsistent. Depending on the specific application scenarios, a tradeoff has to be seek.

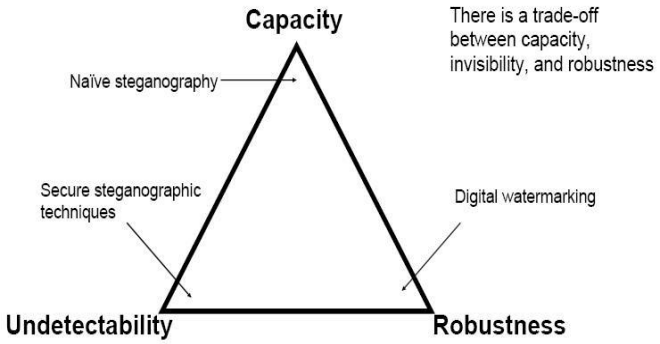


Figure-1.1 Tradeoff between embedding capacity, undetectability and robustness in data hiding.

One of the possible ways of categorizing the present steganalytic attacks is on the following two categories:

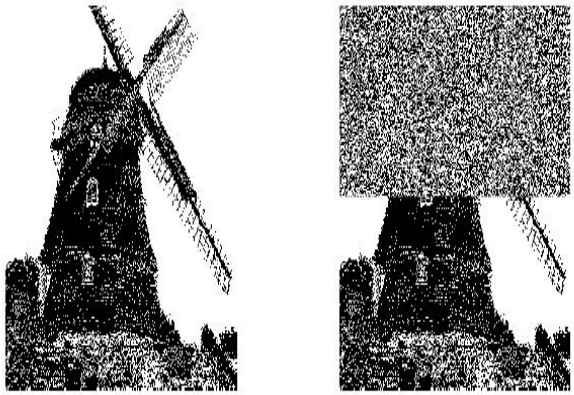


Figure 1.2 Visual attacks for detecting hidden messages in an image layer

1. Visual Attacks: These methods try to recognize the presence of information by optical observation either by the uncovered eye or by a computer. The attack is based on estimate the inserting coated of an image (say a bit plane) and then optical inspecting that layer to look for any uncommon modifications in that layer as shown in Figure 1.2.

2. Statistical Attacks: These methods use first or higher arrangement statistics of the image to reveal small scale adaptation in the statistical conduct caused by steganographic inserting and hence can successfully detect even small amounts of inserting with very high perfection.
These class of steganalytic attacks are further classified as 'Targeted Attacks' or 'Blind Attacks' as explained in specific in the next few section

1.2 A Steganographic Framework

Any steganographic system can be studied as shown in Figure 1.4. For a steganographic algorithm having a stego-key, given any cover image the inserting process generates a stego image. The removal process takes the stego image and using the shared key applies the inverse algorithm to extract the secret message.

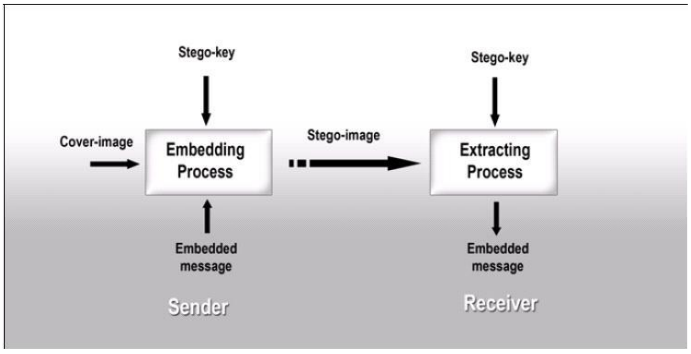
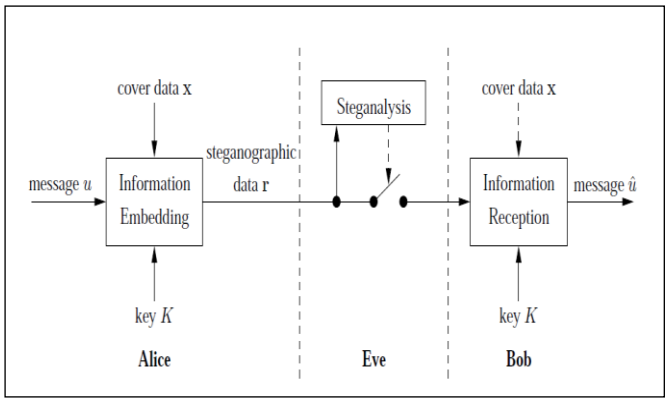


Figure 1.3 A generalized steganographic framework

This system can be explained using the 'prisoner's problem' (Figure 1.4) where Alice and Bob are two patients who wish to communicate in order to hatch an abscond plan. However, communication between them is examined by the supervisor, Wendy. To send the secret message to Bob, Alice inserts the hidden message 'm' into the cover object 'c', to obtain the stego object 's'. The stego object is then sent via the public channel. In a solid steganographic framework, the technique for inserting the message is unknown to Wendy and shared as a secret between Alice and Bob. In private key steganography Alice and Bob share a hidden key which is used to inserted the message. The hidden key, for example, can be a password used to seed a pseudo-random number generator to select pixel locations in an image cover-object for inserting the hidden message. Wendy has no knowledge about the hidden key that Alice and Bob share, even though she is aware of the algorithm that they could be engage for inserting messages. In public key steganography, Alice and Bob have private-public

key set and know each other's public key. In this thesis we enclose ourselves to private key steganography only.



II. RELATED WORK

In 2006, Chang and Wu proposed an adaptive data-hiding method based on VQ [1]. The Chang–Wu method consists of three procedures: the codeword grouping procedure, the data embedding procedure, and the data extracting procedure. In the grouping procedure, the method iteratively selects an initial codeword with the smallest index as the seeding codeword. Then, it finds out all the closest codeword to the seeding codeword with a given threshold (TH), and merges the code words into the same group. The loop continues until all the code words are put into some groups. In the data-embedding procedure, the secret data is embedded into the index table. The capacity for embedding secret data depends on the size of the group to which the current encoding codeword belongs.

In 2009, Shie and Lin also proposed two similar data hiding methods using SOC [2]. Themethods embed either an unfixed or fixed number of secret bits in each block that is encoded in SOC. If the number of embedding secret bits is not fixed, two parameters, a seed key *k*, and an integer *I* are used to generate a random number, ranging from 1 to *i*. On the other hand, if the embedding secret bits is fixed, the parameter *p* is the secret bits in each SOC code. The BR of [2] is shown as follows:

In 2009, Wang and Lu proposed a paper entitled "A path optional lossless data-hiding scheme based on VQ joint neighboring coding" [3]. Each adjacent block of the target blockis given a number, called the position flag bits The method chooses the adjacent blocks of the target block on the index table to perform joint neighboring coding according to their corresponding position flag bits. Then the secret data is embedded into the cover image using a given initial key and the secret data content. Two paths are used to increase the capacity: when Path 1 is selected, 2 bits of the secret data are hidden in each block on average, and when Path 2 is selected, 3 bits of the secret data are embedded in each block. In 2010, Chen and Chang proposed a high-capacity, image hiding method based on SMVQ using adaptive indices [5]. In the Chen–Chang embedding procedure, as shown in Fig. 12, two thresholds THSC and THSMVQ are proposed to determine the hiding strategy, where THSC determines how the state codebook is constructed and THSMVQ determines which encoding type, VQ or SMVQ, is used. If VQ is chosen, the method partitions the codebook into three groups: G0, G1, and G−1. Among the code words of group G0 and G1, the code words in the same index are set to be very similar. The residual code words that are not similar to others are classified into group G−1. While embedding the hidden data, if the to be- embedded secret bit is 0 and the closest code word belongs to G0, then the closest code word is outputted. However, if the closest code word belongs to G1, then it outputs the corresponding code word in G0

In 2010, Wang *et al.* [4] proposed a reversible, VQ-based data-hiding method using index difference [4]. Their method can be divided into three parts: preprocessing, embedding, and extracting-recovering.

In2012 Patrick Bas and Teddy Furon proposed methodology to assess the security levels of watermarking schemes, proposed in [11], [12], [13], [14], [15], poorly captures T. Kalker's definition. In a nutshell, the methodology proposed in these papers is based on C. E. Shannon's definition of security for symmetric crypto-systems [16]. The security level is defined as the amount of uncertainty the attacker has about the secret key. This is measured by the equivocation which is the entropy of the key knowing some observations such as contents watermarked with the same technique and the same secret key.

In 2013 **T.Lakshmi1,P.Ganga Bhavani2** proposed digital image processing, while processing images here we need to follow certain criteria. Combining data into a carrier for conveying secret messages that should be confidentially is the technique of data hiding [6], [7]. After embedding, pixels of cover images will be modified and deformation occurs. The distortion caused by data embedding is called the embedding distortion [8]. A good data-hiding method must be capable of evading visual and statistical detection [9] while providing an adjustable payload [10].

In 2014, Kumar have proposed a new steganography method called JMQT based on modified quantization table. This steganography method is differentiating with steganography method JPEG-Jsteg. Two parameters namely volume and Stego-size has been differentiated. More data can be inserting using this method as compared to JPEG-JSteg. So JMQT provides better volume and JPEG-Jsteg furnish better stego-size.

A. Algorithm Used

4.3.2 Spatially Desynchronized Steganographic Algorithm (SDSA)

The algorithm is summarized below.
Algorithm Spatially Desynchronized Steganographic Algorithm (SDSA)
Input: Cover Image
Input Parameters: Rows and Columns to be cropped (u, v), Block size (m × n), Quantization Matrix (Q)

Output: Stego Image Is
Begin

1. Partition the cover image I into $\hat{I}(U,V)$ and $I^{\delta}(U,V)$ by cropping u topmost rows and v leftmost columns.
2. Perform m×n non-overlapping block $\bar{I}_{u,v}$. Let us denote this set of blocks $P^{(m\times n)}_{\bar{I}_{u,v}}$.
3. Choose a set of blocks from $P^{(m\times n)}_{\bar{I}_{u,v}}$ (using a key shared by both ends) and perform the embedding in each of the selected blocks using any standard DCT based steganographic scheme. The quantization matrix Q which is a shared secret is used for obtaining the quantized coefficients.
4. Apply DE quantization and Inverse Discrete Cosine Transform (IDCT) to the set of blocks used for embedding in Step 3.
5. Join $I^{\delta}_{u,v}$ with the resulting image obtained at Step 4. This combined image is the output stego image Is which is compressed using JPEG compression and communicated as the stego image.

23 DCA crops 4 rows and 4 columns from the top and the left of the image and uses the resting portion of the image for supposing the cover image statistics. To disturb this calibration step, at the time of embedding u rows and v columns (where u; v 6= 4 or any multiple of 8) should be cropped from the left and the top of the image. Thus the cover image is spatially desynchronized before actual embedding is done. During steganalysis, the steganalyst uses the stego portion of the image itself as a reference for estimating the cover image statistics and hence is not able to distinguish the cover image statistics from the stego image statistics. It should be noted that if u consecutive rows and v consecutive columns are cropped from an image, then u; v 6= 4 or any multiple of 8 because such kind of cropping will realign the blocks of the partitioned image I2 with the original cover image and hence in effect there won't be any desynchronization during embedding.

Also since the embedded image undergoes JPEG compression before being communicated to the decoding end, some of the embedded data bits might get lost in the process because of the quantization step during JPEG compression. This quantization loss occurs for almost all the DCT domain embedding schemes. We try to circumvent this problem by 'embedding data mainly in the low-frequency DCT coefficients. Also embedded data can be made secure by adding some redundant bits in the data stream and using error-control coding techniques. This problem of using error-control coding for securing the data bits has been addressed in albeit at the cost of low embedding rate. We would like to mention here that in our implementations of QIM, YASS and SDSA we have not included any error-control technique. Hence we shall be comparing the raw versions of the three schemes. In the next section we verify our claim using statistical hypothesis testing.

IV. Simulation Results

5.1 Result & Discussion:

To demonstrate the performance of the new mechanism, different types of images used in the simulations are shown in Figure 2. The image quality is evaluated by both the human eye and the peak signal–to-noise rate (*PSNR*).

1. The *PSNR* formula is described as follows

$$\mathrm{PSNR} = 10 \times \log_{10}(m \times n/\mathrm{MSE})$$

Where

$$\mathrm{MSE} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left[I(i,j) - S(i,j) \right]^2$$

and *m* and *n* are the width and length of image

Fig. 5.1 Test Image

TABLE 5.1 PSNR values with various VQ block size at VQ compression

Images	VQ block size4	VQ block size8	VQ block size16
Peppers	34.3941	27.6867	21.3304
Leena	33.5457	27.2975	22.4519
Barbara	31.965	26.4989	21.126
Airplane	35.0654	28.5519	24.2228
Baboon	33.987	28.1785	24.3468
Boat	34.5713	28.0243	23.66

TABLE 5.2 MSE values with various VQ block at VQ compression

TABLE 5.3 PSNR values with various VQ block size at embedding

Images	VQ block size4	VQ block size8	VQ block size16
Peppers	64.1897	68.4774	71.5149
Leena	70.461	70.5798	71.967
Barbara	69.9663	71.1327	67.8491
Airplane	69.2877	69.2155	71.979
Baboon	62.2877	71.7152	7.1727
Boat	64.421	62.5932	68.7431

TABLE 5.4MSE values with various VQ block size at embedding

Images	VQ block size4	VQ block size8	VQ block size16
Peppers	23.6413	98.1185	369.3651
Leena	26.1004	102.0181	330.2489
Barbara	37.5585	126.8415	380.2856
Airplane	16.0513	71.2903	196.6021
Baboon	21.6777	85.4366	206.4531
Boat	21.1223	87.7801	245.9213

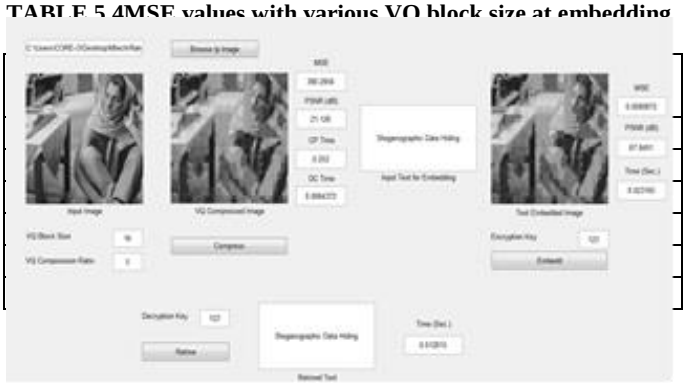


Figure 5.2 after VQ compression when block size is 16

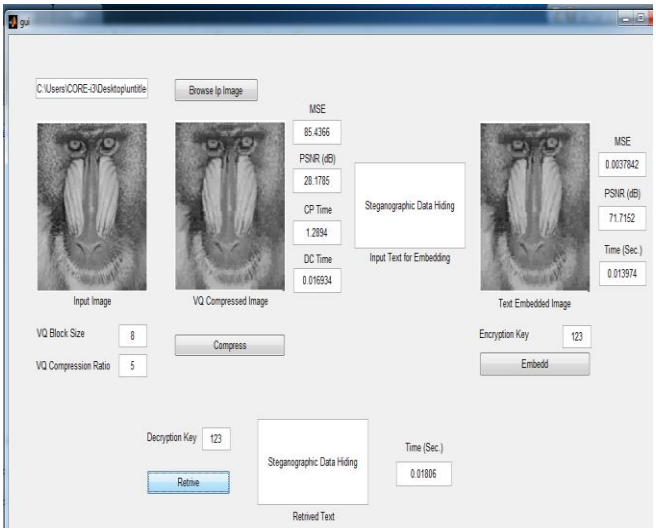


Figure 5.3 after VQ compression when block size is 8

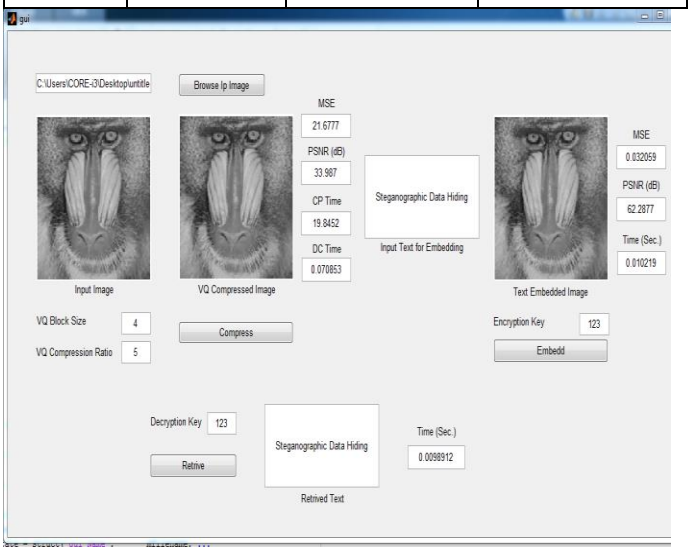


Figure 5.4 after VQ compression when block size is 4

V. CONCLUSION

During the course of this work, some inherent directions of future research were identified. Firstly, it was noticed that most of the steganographic research work till date has been towards designing algorithm which generate stego images which are as close to the cover as possible. All the algorithms study the nature of the cover image while ignoring the message bit stream. It may be feasible to design some encoding functions, which given a cover image and an embedding algorithm can adjust the message stream such that it becomes more useful for embedding than the original bit stream. This kind of steganography can be effective even in the "Active Warden Framework" of steganography because firstly the modified message stream will be recommending less artifacts in the cover. Secondly, even if the embedding algorithm is known to the attacker, the exact message sequel cannot be reconstructed unless the attacker has the knowledge of the encoding function.

A possible direction of research can be codifying as a problem of "Image Retrieval". It is established on searching for a suitable cover image given a message sequel and the embedding algorithm. This may be possible through manage a

huge image database and given any message sequel and a pseudo-random key for generating embedding locations, we hunt for a cover image from the database that will generate a stego image with minimal amount of changes. The change benchmark considered for searching can be dependent on the

features used by the corresponding steganalytic attacks. Some other possible ideas can be imitated from the field of "Visual Cryptography" which encrypts a message by circulate the decoding key into different images such that the message can be crumbled only by a proper combination of these images.

REFERENCES

[1] C. C. Chang and W. C. Wu, "Hiding secret data adaptively in vector quantisation indextables," IEE Proc. Vision, Image Signal Process., vol. 153, no. 5, pp. 589–597, 2006.

[2] S. C. Shie and S. D. Lin, "Data hiding based on compressed VQ indices of images," Comput. Standards Interfaces, vol. 31, no. 6, pp. 1143–1149, 2009.

[3] J. X. Wang and Z. M. Lu, "A path optional lossless data hiding scheme based on VQ joint neighboring coding," Inform. Sci., vol. 179, no. 19, pp. 3332–3348, 2009.

[4] Z. H. Wang, C. C. Chang, K. N. Chen, and M. C. Li, "An encoding method for both image compression and data lossless information hiding," J. Syst. Softw., vol. 83, no. 11, pp. 2073–2082, 2010.

[5] C. C. Chen and C. C. Chang, "High capacity SMVQ-based hiding scheme using adaptive index," Signal Process., vol. 90, no. 7, pp. 2141– 2149, 2010.

[6] J. Fridrich, Steganography in Digital Media: Principles, Algorithms,and Applications. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[7] A. Cheddad, J. Condell, K. Curran, and P. McKevitt, "Digital image steganography: Survey and analysis of current methods," Signal Process.,
vol. 90, pp. 727–752, 2010.

[8] P. Comesaña, L. Pérez-Freire, and F. Pérez-González, "Fundamentals of data hiding security and their application to spread-spectrum analysis," in 7th Information Hiding Workshop, IH05, Barcelona, Spain, June 2005, Lecture Notes in Computer Science, Springer Verlag.

[9] F. Cayre and P. Bas, "Kerckhoffs-based embedding security classes for WOA data-hiding," IEEE Transactions on Information Forensics and Security, vol. 3, no. 1, March 2008.

[10] L. Pérez-Freire and F. Pérez-González, "Spread spectrum watermarking security," IEEE Transactions on Information Forensics and Security, vol. 4, no. 1, pp. 2–24, Marsh 2009.

[11] C. E. Shannon, "Communication theory of secrecy systems," Bell System Technical Journal, vol. 28, pp. 656–715, 1949.

[12] "Data Matrix barcode FAQ & tutorial," accessed Jan 2008. [Online]. Available:

[13] R. Vill´an, S. Voloshynovskiy, O. Koval, and T. Pun, "Multilevel 2D bar codes: Towards high capacity storage modules for multimedia security and management," IEEE Transactions on Information Forensics and Security, vol. 1, no. 4, pp. 405–420,

[14] D. Parikh and G. Jancke, "Localization and Segmentation of A 2D High Capacity Color Barcode," Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on, pp. 1–6, 2008.

[15] O. Bulan, V. Monga, and G. Sharma, "High capacity color barcodes using dot orientation and color separability," in Proc. SPIE: Media Forensics and Security XI, E. J. D. III, J. Dittmann, N. D. Memon, and P. W. Wong, Eds., vol. 7254,

International Journal of Engineering Technology and Applied Science

(ISSN: 2395 3853), Vol. 2 Issue 9 September 2016

Paper ID: IJETAS/SEPTEMBER/2016/380