



Secure Separate Bit Plane Image Processing for Distributed Video Surveillance System (DVSS/DVC)

Kien. T.V¹; An. B.L²; Quynh. L.C³

¹Electric Power University, Ha Noi, Vn

²Thieuduong Company, Quangninh, Vn

³Thieuduong Company, Quangninh, Vn

¹ kientv@epu.edu.vn; ³ quynh.lechi@gmail.com

DOI: [10.47760/IJCSMC.2020.v09i09.007](https://doi.org/10.47760/IJCSMC.2020.v09i09.007)

Abstract— For emerging applications such as wireless video cameras, wireless low-power surveillance networks, and disposable video cameras for medical applications. DVC is useful and maybe the best choice. Since the primary objective of DVC is low-complexity video encoding, the bulk of computation is shifted (transmitted) to the decoder, as opposed to a low-complexity decoder in conventional video compression standards such as H.264 and MPEG. Besides the low bit rate, the transmission has to be very secure also. That means the complexity of the used keys should be rather high and the key space has to be large to provide enough individual keys for a large wireless sensor network.

Keywords— DVSS/DVC, compressive sensing, security.

I. INTRODUCTION

First, let have a look at the DVC network [1]

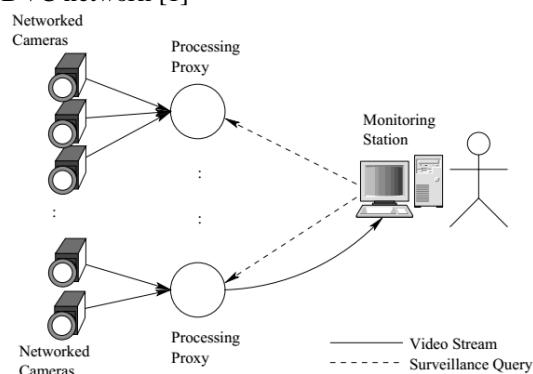


Figure. 1: Architecture of Distributed Video Surveillance System.

Our concern here is distributed video surveillance systems, which consist of a number of video sources, processing proxies, and monitoring stations, connected via a wide area network. Video signals can be taken from either networked cameras or video sensors. They capture, encode and transmit video streams to processing proxies. Processing proxies are computers dedicated to the processing and filtering of incoming video streams, and if needed, relaying them to monitoring stations. The need to relay depends on the queries specified by users.

Generally, the DVC encoder has the following modules [2]

- Adaptive Video Splitter
- Transform
- Quantizer and Bit plane ordering
- LDPCA Encoder and buffer
- H.264 Intra Encoder

In dealing with the DVC network as an important member of the wireless sensor network we should take care of the following issues [2]:

- Taking the measures for a significant reduction in bandwidth and resource usage when a video is transmitted for analysis by vision algorithms and not for human viewers and The resulting image qualities are evaluated via well-known perceptual concept like (PSNR, SSIM) [3,4 ..].
- Finding out the optimization solution, that allows minimization of bit rate under accuracy constraints, or maximization of accuracy under bit rate constraints.

These networks are vulnerable to numerous security threats that can adversely affect their proper functioning due to the distributed nature of these networks and their deployment in remote areas This problem is more critical if the network is deployed for some mission-critical applications such as in a tactical battlefield or medical operations. Recently special attention has been paid to the physical layer security and the PN-masking approach has been proved to be very effective [5-10]. In this regard, we want to point out that choosing the optimal PN-sequences with great length L, good Autocorrelation Function (ACF), high linear complexity (LC), simple hardware implementation, and having a huge keyspace is of great importance!

II. SELECTIVE BIT-PLANE SCRAMBLING ON THE PHYSICAL LAYER

Selective bit-plane scrambling can bring about significant savings in terms of processing time or power. In the literature, bit-plane extraction is sometimes termed as lossless bit plane compression and used for the Mobile environment [7] Selective encryption is starting with the bit plane containing the most significant bit (MSB) of the pixels. Each possible subset of bit planes may be chosen for SE (selective encryption), however, the minimal percentage of data to be encrypted is 12.5% (when encrypting the MSB bit plane only), increasing in steps of 12.5 % for each additional bit plane encrypted.

We assume a 512×512 pixels image to be given in 8bit/pixel (bpp) precision. Each pixel has a gray value between 0 and 255 For example, a dark pixel may have a value of 10 and a bright pixel might have a value of 230. The entire image can be considered as a two-dimensional array (2_D) of pixel values. We consider the 8bpp data in the form of 8-bit planes, each bit plane associated with a position in the binary representation of the pixels. 8-bit data is a set of 8-bit planes. Each bit plane may have a value of 0 or 1 at each pixel, but together all the bit planes make up a byte with a value between 0 to 255 [8].

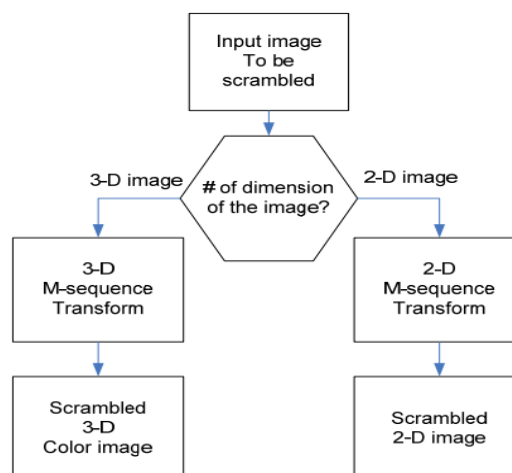


Fig. 2 Multi-dimension image presentation

The steps to construct such a matrix is elaborated as follow: The constant-W two-dimensional matrix is formed as follows [9]

- i. Form a $2^n - 1$ period bipolar m-sequence: $T = 255$ (the transformation from binary to bipolar (ternary) can be done easily and discussed on another occasion).
- ii. Append 0 to the m-sequences to get $T + 1 = 256$ bits sequences (extended m-sequence) since $T = 255$) The extended m-sequences have a good balance property.
- iii. Shape this to form a 256×256 matrix W

To scramble the images (a $16 \times 16, 32 \times 32, 64 \times 64, 128 \times 128$ or 256×256 blocks of the image) Y, perform $Y = Y + W$ where Y is the scrambled image block. This process is repeated until the entire image is scrambled. The logic hardware will perform bit manipulation logic operation on image bit patterns [10].

Let $\{u_i\} = \{u_1, u_2, \dots, u_i\}^L$ be a PN-sequence $\{u_i\}$ (such as m-sequence) The PN-mask vector $\{u_n\}$ is used to scramble the image signal $\{i_n\}$ of interest by component-wise (bit by bit). Physically speaking, $\{u_i\}$ will randomly flips every bit of $\{i_n\}$. For illustration we can use the picture 2,3,4 [11]

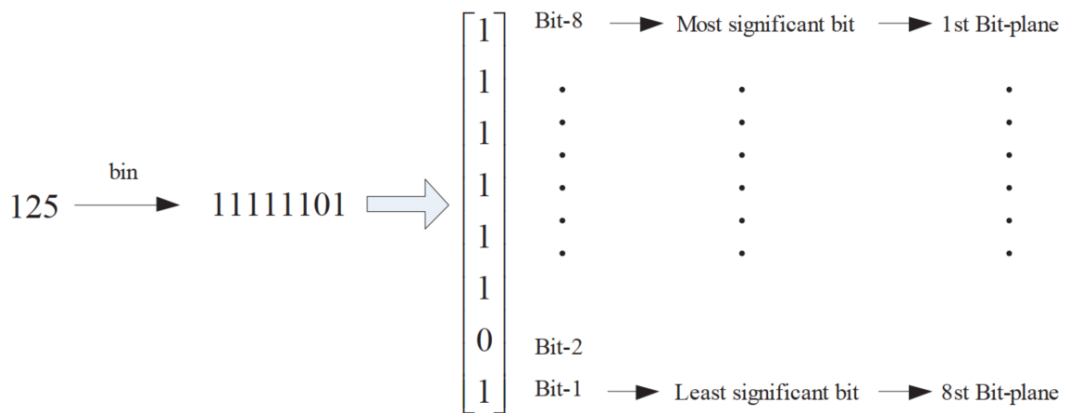


Fig. 3 Bit-plane decomposition

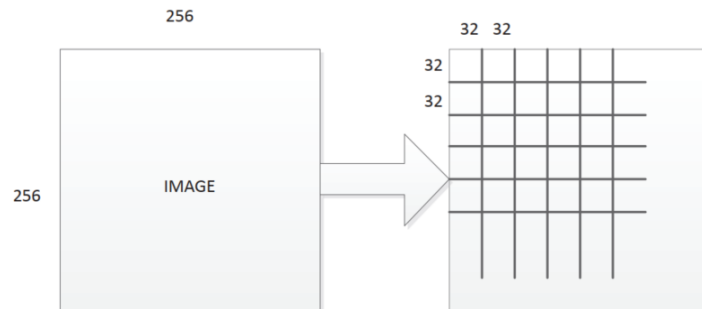


Fig. 4 Image size division

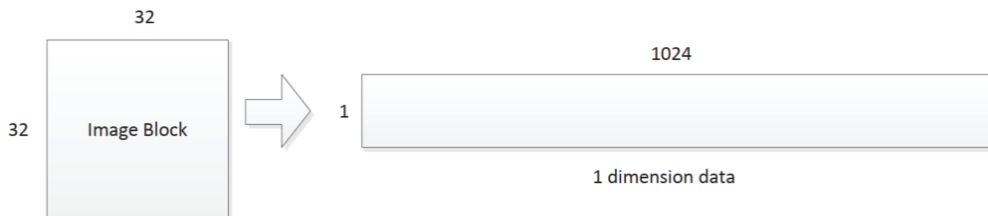


Fig. 5 Two dimension image data transfer to one dimension data

The result of two dimensions of image data transformation into one-dimension data gives us the data bitstream.

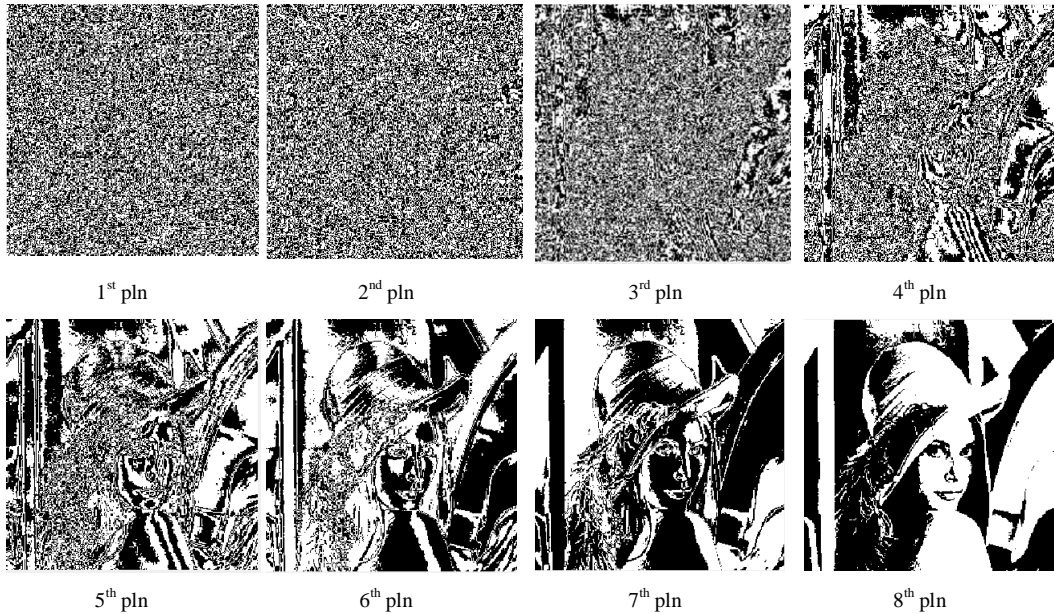


Fig. 6 Bit plane of Lena

A. Bit Stream Representation By D-Transform

Time multiplexing (Hardware oriented method: D-transform) representation and analysis of the interleaved structure

In fact, The delay operation in D-transform is nothing but time multiplexing and can be visualized as shown in Fig. 7.

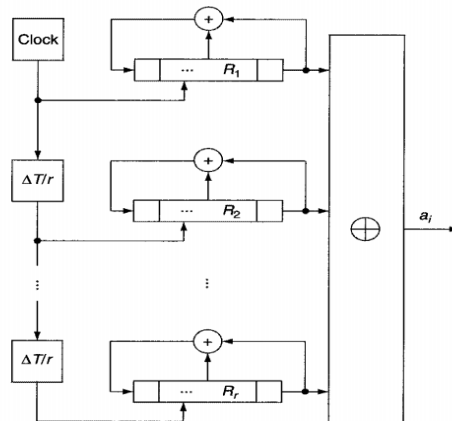


Fig. 7 Time multiplexing presentation of interleave

Definition (D-transform): The D-transform of a sequence $\{b_n\}$ over $GF(p)$ is denoted by $D[b_n]$ or F and designed by:

$$D[b_n] = F = \sum_{i=1}^n b_i D^i \tag{1}$$

Example 3: let $\{b_n\} = 010111$, D-transform of b_n is $D(b_n) = D + D^3 + D^4 + D^5$.

The inverse transform of D is $D^{-1} = \{b_n\}$

The D-transform of the generator sequence $\{b_n\}$ of a linear feedback shift register (LFSR) is then given by:

$$b(D) = \frac{S(D)}{G(D)} \tag{2}$$

Where $G(D)$ of degree n is the generating polynomial of a LFSR and $S(D)$ of degree $\leq n-1$ specifies the initial condition corresponding to a particular shifted version of $\{b_n\}$. When $G(D)$ is primitive, the LFSR

sequence is an m-sequence and there are $2^n - 1$ polynomials $S(D)$ corresponding to $2^n - 1$ values of the initial states of that LFSR.

B. Scrambling (masking) effect of bit manipulation.

Pseudo-noise (PN) masking technique has been proven to be an effective technique against unauthorized data collection (eavesdropping). With random-like PN-masked carriers, the SS signal of interest behaves like white noise. Therefore, PN masked SS can efficiently prevent illegitimate data extraction by unauthorized users without knowledge of PN masks. Mathematically, the whitening effect can be explained like this.

1) $\{1\}$ and $\{0\}$ distribution:

Let $\{O\}$ be the output signal of the scrambler, $\{I\}$ be the image sequences and $\{b\}$ represents LFSR sequences. The probabilities of '0' and '1' bits in $\{O\}$ are $P_0(0)$ and $P_0(1)$ respectively. The probabilities of '0' and '1' bits in $\{I\}$ are $P_I(0)$ and $P_I(1)$ respectively. Similarly, the probabilities of '0' and '1' bits in $\{U\}$ are $P_U(0)$ and $P_U(1)$ respectively. Since scamler is a linear system in $GF(2^n)$, we can apply the super position rule and have:

$$\{O\} = \{I\} + \{b\} \text{ mod } 2 \tag{3}$$

So, the probability of '1' in $\{O\}$ equals the probability of '1' in XOR –operation:

$$P_0(1) = P_I(1).P_u(0) + P_I(0).P_u(1) \tag{4}$$

Since $\{u\}$ satisfies the balance condition, we get:

$$P_u(1) \approx P_u(0) \approx \frac{1}{2} \tag{5}$$

Under the assumption that $\{I\}$ and $\{U\}$ are statistically independent, we have

$$P_0(1) = P_I(1) \cdot \frac{1}{2} + P_I(0) \cdot \frac{1}{2} = \frac{1}{2} [P_I(1) + P_I(0)] \approx \frac{1}{2} \text{ Or } P_0(1) \approx P_0(0) \approx \frac{1}{2} \tag{6}$$

So, any input the sequence with unbalanced distribution will become balanced.

2) *Autocorrelation function (ACF):*

ACF of $\{O\}$ in binary form is calculated as: $R(k) = \frac{A-D}{A+D}$

where: A and D are the agreement and disagreement positions between $\{O\}$ and its shift version respectively. So, ACF is:

$$R(k) = \frac{A-D}{A+D} = \frac{A+D-2D}{A+D} = 1 - 2 \frac{D}{A+D} = 1 - 2P_0(1) \approx 0 \tag{7}$$

In other word $\{O\}$ looks like noise with:

$$P_0(1) \approx P_0(0) \approx \frac{1}{2} \tag{8}$$

$$R(k) = 0$$

This noise looking effect ensures the security of the transmission. The scrambling scheme has a low-security level if the intended receiver and the unauthorized receiver have similar BER. If the intended receiver has low BER and the unauthorized receiver has much higher BER, then the embedding scheme has a high-security level. If the unauthorized receiver has BER as high as 0.5 which means that what he received is garbage, then the scrambling scheme has perfect security [10].

C. Complexity of PN-sequences:

The LFSR is very well studied and can be effectively described in $GF(2^n)$ and easily hardware implemented (FPGA) [12,13]. However, the sequences generated by LFSR normally have low linear complexity (LC). This shortcoming can be overcome by choosing a suitable interleaved structure, which can maintain the best ACF but can increase the LC by hundred times! [11,12]. Furthermore, this structure can provide a huge number of keys needed for Wireless sensor networks.

TABLE I
LC OF OPTIMAL PN-SEQUENCE

Order	GF(2 ¹⁶)	GF(2 ⁸)					
		101110001	100011101	110101001	100101011	101101001	100101101
1	1011010000000001	1024	16	256	64	64	256
2	1000000000101101	16	1024	64	256	256	64
3	1001110000000001	64	256	1024	16	256	64
4	1000000000111001	256	64	16	1024	64	256
5	1111110000000001	64	256	128	128	1024	16
6	1000000000111111	256	64	128	128	16	1024
7	1100101000000001	1024	16	256	64	64	256

Note that LC of the nonlinear interleaved sequences increase significantly while ACF remains ideal. For example:

- With n=12, LC jumps from 12 to 192 due to the effect of nonlinear operation (when the subsequence 1100001 is replaced by its reciprocal 1000011)
- With n=14, LC jumps from 14 to 488 (when the subsequence 11000001 is replaced by its reciprocal 10000011)
- With n=16, LC jumps from 16 to 1024 (when the subsequence 100011101 is replaced by its reciprocal 101110001)
- With n=18, LC jumps from 18 to 2304 (when the subsequence 1000011011 is replaced by its reciprocal 1101100001) For details see [11,12]

Remark 1: as we already know, the number of primitive polynomials of degree n is: $\frac{\Phi(p^n - 1)}{n}$.

The Hughes XPD/KPD stream cipher algorithm uses a 61-bit LFSR. There are about 1024 different primitive polynomials (approved by NSA) stored in a table that are key selected for the XPD/KPD. But, there are in fact 37, 800, 705, 069, 076, 950 degree-61 primitive polynomials, not just 1024. A greater choice in the selection of primitive polynomials is clearly a requirement to ensure the strength of stream cipher implementations [13]. It is

not at all easy to carry out the exhaustive search the $\frac{\Phi(p^n - 1)}{n}$ primitive polynomials.

Taking into account four criteria: hardware simplicity, the large value of LC, best possible ACF, huge key space dimension, the nonlinear interleaved structure may be the best choice for this purpose! To the best of our knowledge, there has not been any such publication for DVC!

III. SIMULATION

Randomization effect of scramblers: in this section, some sensitive information (on telecom network performances and human picture) are taken for testing and the result are evaluated by perceptual concept such as Structural Similarity (SSIM) index, PSRN, [93,4,14-19]

A. On Network management pictures

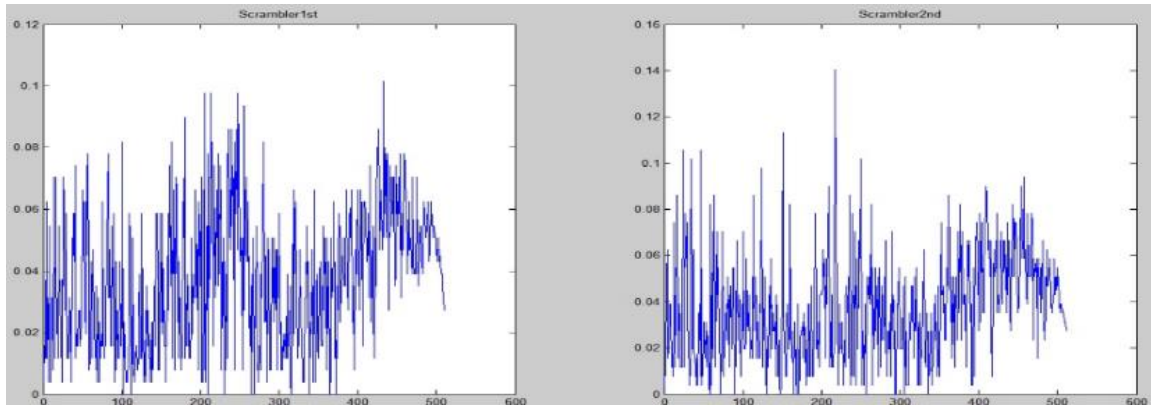


Fig. 8 Power spectrum of the scrambler's input and output

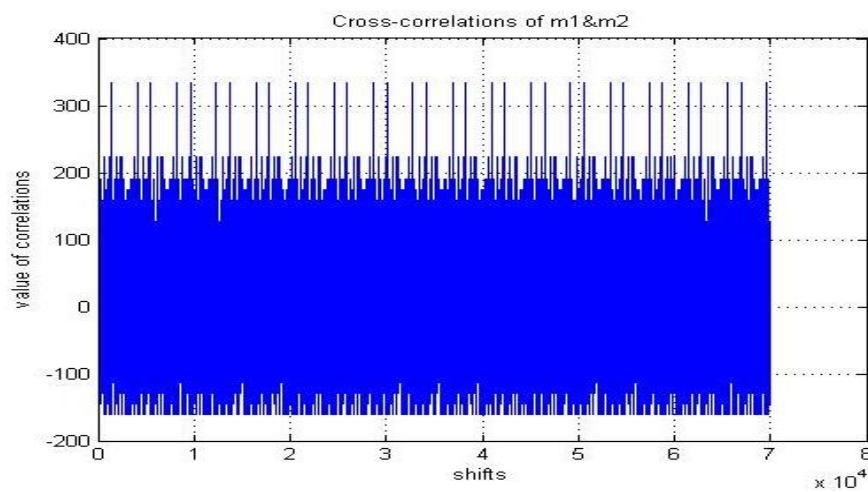
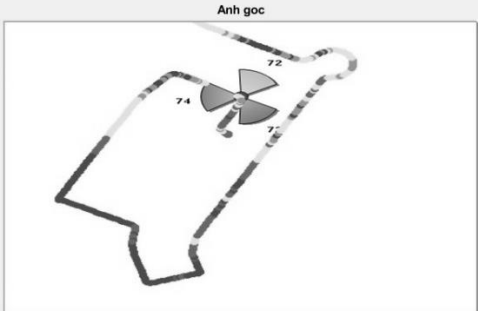
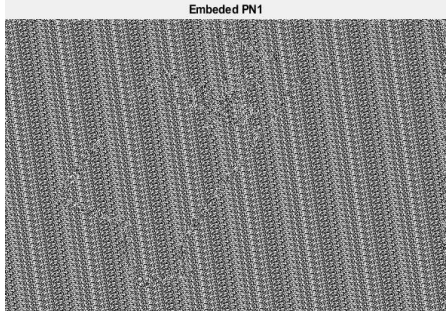


Fig. 9 CCF between two sequences



Anh goc

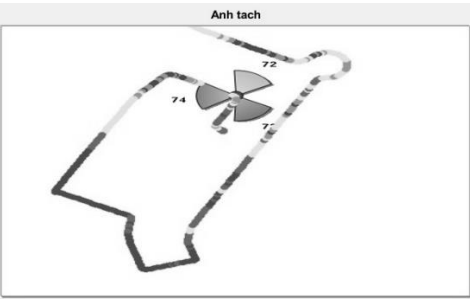
Original picture



Embeded PN1

PN- masked picture

MSE for embedding	231.9703
PSNR for embedding	56.3592
MSE for extracted	254.9648
PSNR for extracted	55.414
The SSIM value PN image	0.0077911
The SSIM value PN image	1



Anh tach

Recovered picture

Perceptual indices

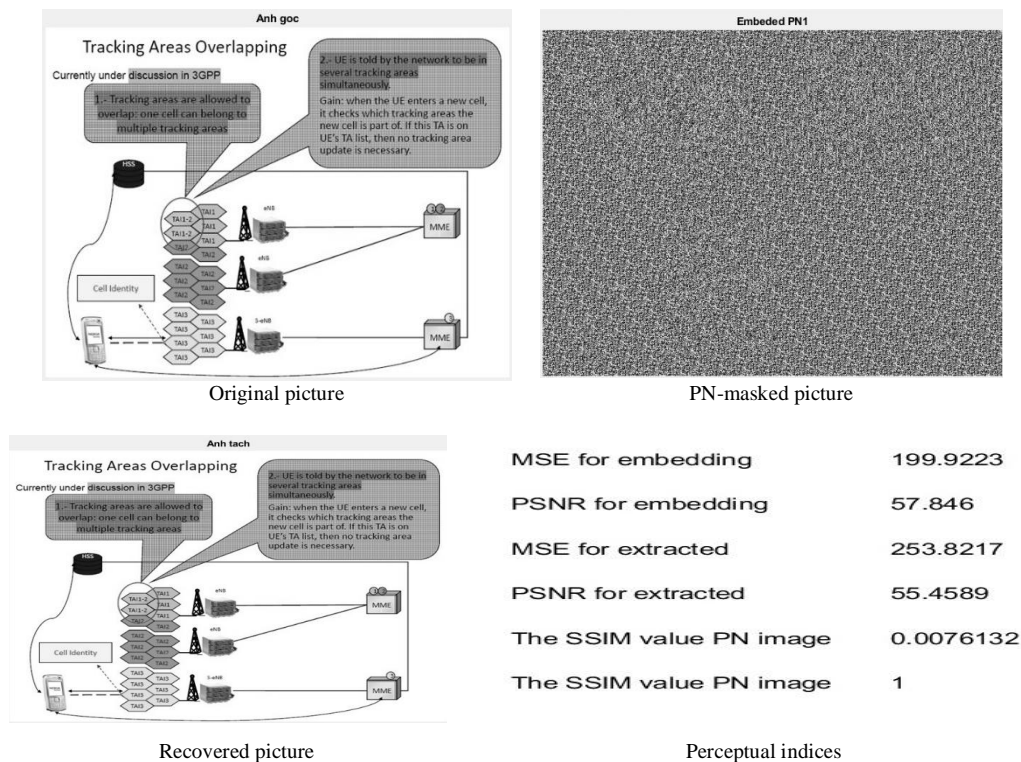


Fig. 10 Original masked and recovered pictures and perceptual indices

B. On human picture

[22-25] new image/video coding approach combining the CS theory into the traditional discrete cosine transform (DCT) based coding method to achieve better compression efficiency for a spatially sparse signal. The sparsity of the signal can be proved by the calculated gradient and showing that significant pixel value variations only occur at a few pixels (little state change of the binary stream). According to the groundbreaking work by Candes et al. [22] and Donoho [23], who showed that under certain conditions (sparsity), a signal can be precisely reconstructed from only a small set of measurements. The CS principle provides the potential of the dramatic reduction of sampling rates, power consumption and computation complexity in digital data acquisitions.

For image and video coding, the signal is 2-D form and few blocks can be defined sparse by the above criterion. However, blocks containing a sharp edge or several edges are more common. For these blocks, the ℓ_1 -norm criterion cannot be directly applied since a large percentage of pixels are not zero-valued. Instead, these blocks can be defined as gradient sparse, i.e. significant pixel value variations only occur at a few pixels. For a $n \times n$ block, the gradient can be defined as [24][25]:

$$D_{ij}s = \begin{pmatrix} D_{h,ij}s \\ D_{v,ij}s \end{pmatrix}. \quad (9)$$

Its horizontal components is

$$D_{h,ij}s = \begin{cases} s_{i+1,j} - s_{ij} & i < n \\ 0 & i = n \end{cases} \quad (10)$$

and its vertical component is

$$D_{v,ij}s = \begin{cases} s_{i,j+1} - s_{ij} & i < n \\ 0 & i = n \end{cases} \quad (11)$$

Thus, the total variations of s is simply the sum of the magnitudes of this discrete gradient at every point:

$$TV(s) = \sum_{i,j} \sqrt{(D_{h,ij}s)^2 + (D_{v,ij}s)^2} = \sum_{i,j} \|D_{ij}s\|_2 \quad (11)$$

In word, the above forms show the vertical and horizontal state change of the image signal. The value are significantly smaller than the binary stream bit rate. That means combing CS with separate bit plane techniques can dramatically reduce the sampling rates, power consumption and computation complexity in digital data acquisitions. For example [11] the 32×32 block (rows and columns) the 8th bit plane of Lena picture can be recovered with $M = 4.K = 4.81$ elements where K is the sparsity determined by D-matrix. In [22,25] the sampling rate can be reduced to 1/3 without affecting the quality of the picture!. In our experiments, we take K as the maximal K of the blocks and the perceptual indices (PSNR, MSSSI) will be acceptable!



Fig. 11 Image matrix (1 pixel = 8bit)

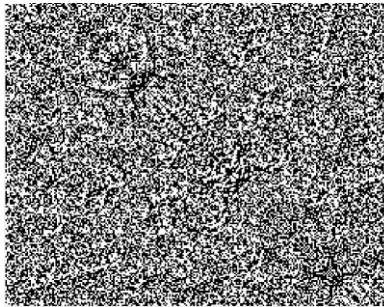


Fig. 12 Measurement Matrix (200×256) constructed from PN-sequences



Fig. 13 8th bit plane of lena picture

-sparse presentation K= 1024 [11]



Fig. 14 Masked 8th bit plane of lena picture (matrix 200×256)

-sparse transformation: $K= 109$



Fig. 15 8th bit plane reconstruction: and PSNR : 54.1423 and SSIM : 0.990529 MSE 0.158187

To assess the perceptual quality of the data-masked images the measures like MSE, PSNR, SSIM [3,4,14-19] are used



Fig. 16 7th bit plane of lena picture



Fig. 17 7th bit plane of transform differential



Fig. 18 Masked . 7th bit plane of lena picture (matrix 200×256)



Fig. 19 Reconstructed 7th bit plane of lena picture and PSNR : 81.0344 and SSIM : 0.999538 MSE 0.0101782

IV. CONCLUSIONS

In this contribution, we propose and analyse the SECURE separate bit plane image processing for DSVC. To reduce the bit rate we employed the separate bit-plane extraction and compressive sensing technique. It can be seen clearly how much the bit rate is reduced. On the transmission media, the nonlinear interleaved sequences are used for masking the signal. Without the scrambling key (nonlinear interleaved sequences) the attacker will definitely get garbage. Regarding the Best ACF, high complexity, and large key space (enough for millions of sensors) these scrambling sequences may be the best choice.

In the future, we need to introduce an authentication measure to enhance the security and some algorithms to detect the movements for applications in DVSC.

ACKNOWLEDGEMENT

The authors express their gratitude to “THIEUDUONG” company for their financial support to carry out this project!

REFERENCES

- [1] Xiaojing .Y et al, *A Distributed Visual Surveillance System*, University of Nevada 2014 pp 1-6.
- [2] Pavel .K et al, *Critical Video Quality for Distributed Automated Video Surveillance*, National University of Singapore 2005 pp 1-10.
- [3] Pascal .L, *QoS and QoE in the Next Generation Networks and Wireless Networks*, lorenz@ieee.org 2015
- [4] Ning M. et al, *Distributed video coding scheme of multimedia data compression algorithm for wireless sensor networks*, Ma EURASIP Journal on Wireless Communications and Networking (2019) pp 1-9.
- [5] Jaydip Sen, *A Survey on Wireless Sensor Network Security International Journal of Communication Networks and Information Security*, (IJCNIS) Vol. 1, No. 2, August 2009 pp55-78.
- [6] Amitav. M et al, *Principles of Physical Layer Security in Multiuser Wireless Networks*, A Survey 20 Jan 2014 pp 1-22.
- [7] Nazneen M. G et al, *Selective Bit plane Encryption For Secure Transmission Of Image Data In Mobile Environment*, International Journal Of Scientific & Technology Research Volume 2, Issue 6, June 2013 Pp 92-96.
- [8] Yicong Z Et Al, *An Image Scrambling Algorithm Using Parameter Based M-Sequences*, Tufts University, Pp 1-8.
- [9] Raymond.B.W Et Al, *A Watermarking Technique For Digital Imagery*, Further Studies Cerias Tech Report 2007-44 Purdue University Pp1-8.
- [10] Ming L, *Secure spread-spectrum data embedding with PN-sequence masking Signal Processing*, Image Communication 39 (2015) 17–25
- [11] *Bit-plane Image Coding Scheme Based On Compressed Sensing*, Applied Mathematics & Information Sciences vol. 6, No. 3, 721-727 (2012)
- [12] Quynh. L.c, *On the Comparative study of Some Mathematical Tools for Specific Sequences Design*, Journal of Information Engineering and Applications Vol.5, No.12, 2015 pp1-12
- [13] Kien t.v et al, *FPGA Implementation of optimal PN_sequences by time_multiplexing technique Technique*, Springer Nature Switzerland AG 2020 K.-U. Sattler et al. (Eds.): ICERA 2019, LNNS 104, pp. 373–380, 2020.
- [14] Nirmal R. Saxena and Edward J. McCluskey, *Primitive Polynomial Generation Algorithms Implementation and Performance Analysis*, 2004 Stanford University.
- [15] Arezou S. P, *Digital Watermarking of Non-media data stream (applications)*, PhD thesis RMIT University Melbourne, Australia June, 2017.
- [16] Vranješ M. et al, *Objective Video Quality Metrics*, University of Osijek Croatia 2006.
- [17] Santi P. M. et al, *Spread Spectrum Watermarking: Principles and Applications in Fading Channel Watermarking*, Volume 1 2012 pp-85-105.
- [18] Yusra A. Y et al, *Comparison of Image Quality Assessment:PSNR, HVS, SSIM, UIQI*, International Journal of Scientific & Engineering Research, Volume 3, Issue 8, August-2012 pp 1-5.
- [19] Zhou W. et al, *Image Quality Assessment: From Error Visibility to Structural Similarity*, IEEE Transactions on Image Processing · VOL. 13, NO. 4, APRIL 2004 pp 1-15
- [20] Candes E et al, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, vol. 52, pp.489–509, Feb. 2006.
- [21] Donoho, D.Let al, *Compressed sensing*, IEEE Trans. Inform.Theory, vol. 52, pp. 1289–1306, July 2006.
- [22] Candes E et al, *L1-MAGIC: Recovery of sparse signals via convex programming*, 2005.

- [23] Yifu .Z, Zhibo Ch, *A Novel Image/Video Coding Method Based On Compressed Sensing Theory*, Acoustics, Speech, And Signal Processing July 2017. Pp 1361-1364.
- [24] Vishal M. P. et al, *Sparse Representations and Compressive Sensing for Imaging and Vision*, Springer 2013.
- [25] Hui W, *Measurement Matrix Construction for Large-area Single Photon Compressive Imaging*, Sensors 2019, 19, 474 pp1-12.