International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:10, No:11, 2016

H.264 Video Privacy Protection Method Using Regions of Interest Encryption

Taekyun Doo, Cheongmin Ji, Manpyo Hong

Abstract**—**Like a closed-circuit television (CCTV), video surveillance system is widely placed for gathering video from unspecified people to prevent crime, surveillance, or many other purposes. However, abuse of CCTV brings about concerns of personal privacy invasions. In this paper, we propose an encryption method to protect personal privacy system in H.264 compressed video bitstream with encrypting only regions of interest (ROI). There is no need to change the existing video surveillance system. In addition, encrypting ROI in compressed video bitstream is a challenging work due to spatial and temporal drift errors. For this reason, we propose a novel drift mitigation method when ROI is encrypted. The proposed method was implemented by using JM reference software based on the H.264 compressed videos, and experimental results show the verification of our proposed methods and its effectiveness.

Keywords**—**H.264/AVC, video encryption, privacy protection, post compression, region of interest.

I. Introduction

CCTV surveillance has become one of the indispensable elements in our society, due to the advances in video encoding and processing technologies and computer network. Recently surveillance cameras are widely deployed in public space and even in workplaces such as offices, factories, convenience stores, kindergartens, and so on. Videos collected by surveillance cameras provide real-time monitoring capabilities for various accidents, and have considerable effect of crime prevention by providing evidence of criminals. Though they serve as a useful tool for public safety, however, there are mainly two problems concerning CCTV surveillance. First, collected videos can abuse personal privacy. A monitoring person can access personal information such as people's faces without their agreement or their recognition. Second, the surveillance video can be exposed to unauthorized persons. Surveil-lance cameras that were connected previously to closed network are now mostly using the internet as a communication channel, which can be exploited by anyone who has the internet access and some knowledge of the surveillance system. There is a streaming website that shows video data gathered from arbitrary IP-connected surveillance cameras using default ID/password, and IP/port address scanning actually exists [1]. Thus, there needs an additional measure that can address the problem.

One way to protect privacy in surveillance video is de-identification of ROI which include people's faces and car license plates. Fundamental idea is to obscure only ROI in each frame of the video in a reversible way so that privacy can be guaranteed, and the original objective for monitoring and prevention of accidents can still be achieved simultaneously. Obscuring ROI with reversibility can be achieved by encryption. Only authorized persons such as police officers or public prosecutors with legal rights may restore the original video using the encryption key.

In order to reduce the bandwidth of video transmission, surveillance cameras utilize video encoding algorithms such as H.264/AVC standard. Thus, ROI encryption can be applied before [6], [7], during [4], [8]-[12], and after [2], [13], [14] encoding process. Each of them has its own pros and cons. Most previous works suggest ROI encryption during encoding, but they have a significant disadvantage that the design of internal encoder in surveillance cameras must be modified in order to apply the proposed schemes. Since encoders are typically in a form of built-in hardware, it is substantially expensive to replace all the surveillance infrastructures that are already deployed [2]. Several works suggest ROI encryption before encoding. This approach brings reduction of compression efficiency since scrambled footage disturbs prediction, even though it has advantages that ROIs are easily separable from the footage, and the encryption is independent from compression. On the other hand, ROI encryption after encoding has drift issue that modifying video information in the ROI of encoded bitstream causes errors in the other areas of the inter and intra frames.

In this paper, we present an after-encoding ROI encryption method for H.264/AVC bitstream of surveillance system and a mitigation approach for the drift issue. The present method scrambles ROIs by encrypting only a part of non-zero coefficients that is generated after inverse quantization. The encryption process is independent from compression; therefore, it can be adopted in the existing surveillance systems. In addition, we manage the drift issue by inserting regenerated I-pictures. Details of the approach are mentioned in Section III.

The rest of this paper is composed as follows. Related works concerning ROI encryption for privacy protection of surveillance video are introduced in Section II. In Section III, details of our present approach for ROI encryption and drift mitigation are described. Section IV shows our experiment results through actual implementation based on a reference software. Analysis of the result is also given. Finally, we conclude the paper in Section V.

Taekyun Doo and Cheongmin Ji are graduate students at Ajou University, Suwon-si 16499 S. Korea (e-mail: tglft4@ajou.ac.kr, zardmin@ajou.ac.kr).

Manpyo Hong is a professor of Ajou University Suwon-si 16499 S. Korea (phone: +82-31-219-2438, fax: +82-31-219-1614, e-mail: mphong@ ajou.ac.kr).

II. Backgrounds

A. H.264 Encoder and Bitstream Structure

H.264/AVC is a widely adopted video compression standard in many applications for recording, compressing, and deploying the video contents which show a better performance compared to the other previous video compression standards such as MPEG-2 or H.263. H.264 is divided into parts of spatial and temporal domains. There is a high dependency on many parts of coded blocks of footage and video sequences, H.264 works removing its duplications through spatial and temporal predictions. Fig. 1 shows a diagram of common H.264 encoder. In spatial domain, blocks are compressed using already encoded blocks of the current picture through the intra prediction. In the temporal domain, blocks are compressed using previous or future encoded picture through the inter prediction. Finally, raw video is compressed into H.264 after discrete cosine transform (DCT), quantization, and entropy coding are performed [3].

Since H.264 standard has been developed for network transmission, compressed bitstream constitutes a NAL (Network Abstraction Layer) units to transmit video contents effectively. H.264 bitstream is divided into two parts. First one is the parameter information NALs which contain information about video and parameters for decoding process. The other one is video coding layer (VCL) NALs which contain compressed video data. Fig. 3 shows overall structure of H.264 bitstream.

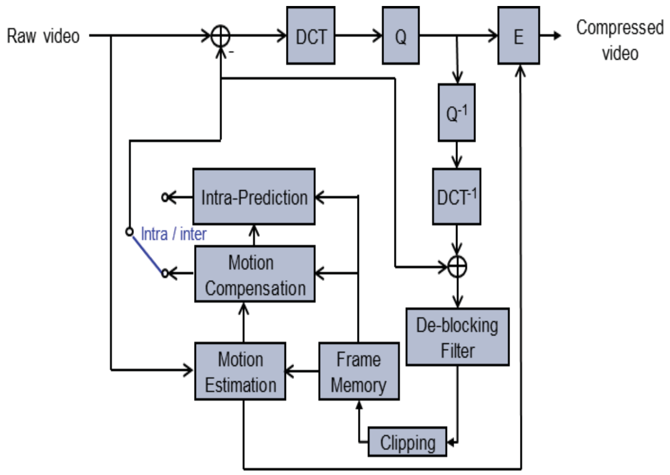


Fig. 1 H.264 video encoder structure. DCT is discrete cosine transform, Q is quantization, E is entropy coding and Q, DCT to the power of –1 is inverse function

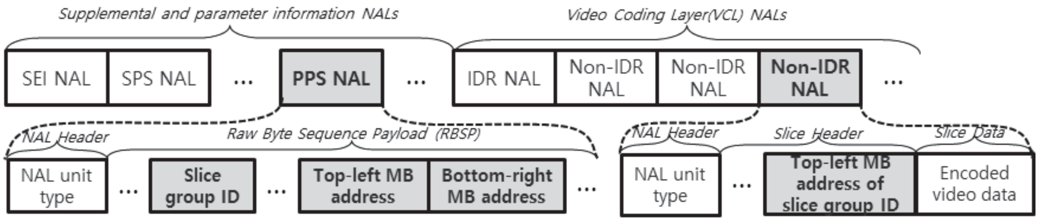


Fig. 2 H.264 bitstream structure defined as ROI. SEI: supplemental enhancement information, SPS: sequence parameter set, PPS: picture parameter set, IDR: instantaneous decoding refresh, Non-IDR: non-instantaneous decoding refresh [4]

B. ROI Encryption Approaches

When it comes to ROI encryption, video encryption approaches can be categorized into three types based on encryption timing; pre-compression, in-compression (or joint-compression), and post-compression [5]. Since pre-compression encryption [6], [7] is performed independently from compression, it has advantages that encryption process of ROI can simply be done. However, it removes duplicated attributes in footage due to the encrypted ROI area, therefore compression efficiency will be decreased [5]. In-compression encryption methods [4], [8]-[12] propose ROI encryption process in conjunction with compression process. They encrypt elements that are produced in the encoding process such as motion compensation and estimation, DCT, quantization, or entropy coding. These approaches have advantages of relatively easy implementation because encryption process is controlled by encoder and it is easy to manage the error propagation that starts from the encrypted areas into the non-encrypted ones through spatial and temporal prediction [2]. In order to apply them, however, all the encoders have to be modified, which is impractical. Post-compression encryption methods [2], [13], [14] perform ROI encryption on the compressed bitstream. They have an advantage that encryption can be applied without modifying the encoders. However, when ROIs are encrypted after compression, a standard decoder cannot obtain proper video that only its ROI areas are obscured because of inter/intra predictions. In order to apply the post-compression ROI encryption method, it needs to address this issue.

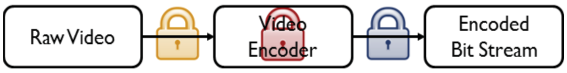


Fig. 3 Three approaches of video encryption

C. Drift Error

Post-compression ROI encryption methods inevitably encounter the drift issue since compressed bitstreams have spatial and temporal dependency. Given compressed bitstream with encrypted ROIs, non-ROI areas can be decoded by referencing encrypted ROI areas during decoding process. Due to this problem, error is propagated out of ROI area, and we cannot obtain only ROI-encrypted video. This issue is called drift in the other words. There are two kinds of drifts, the first one is spatial drift caused by intra prediction, the other one is temporal drift caused by inter prediction.

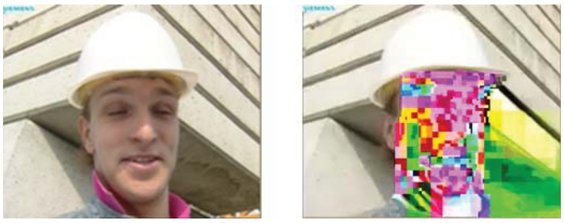

(a) (b)

Fig. 4 Example for spatial drift due to change macroblock (a) original frame and (b) modified frame

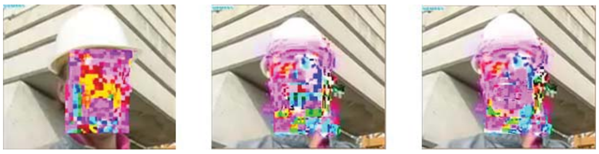

(a) (b) (c)

Fig. 5 Example for temporal drift due to encryption (a) first of ROI encrypted frame, (b) fifteenth frame, (c) thirtieth frame

Fig. 4 shows an example of spatial drift. Macroblocks that are outside of the ROI but adjacent to it are affected by the macroblocks within the encrypted ROI. During decoding, therefore, drift is propagated out of the ROI area due to intra prediction. Fig. 5 shows example of temporal drift. In group of picture (GOP) in H.264 standard, I-picture is generated first, and P and B-pictures are generated through inter prediction for compressing video bitstream. If an encrypted ROI in a picture moves, then the next picture's non-ROI area that previously belonged to the ROI is affected by the encrypted ROI area in the previous picture. Eventually, temporal drift occurs along with ROI traces.

One of the method to prevent spatial drift is flexible macroblock ordering (FMO) [15] that error resilient feature of H.264 standard. FMO allows dividing a picture into regions called slice group that consists of macroblocks. In FMO, ROIs can be defined as slices, and encoder is able to decode slice groups independently. The properties of FMO enables spatial drift to be contained within the slice group.

There have been several studies on post-compression ROI encryption [13], [14], [16], [17]. However, they have common problem: decrease of compression efficiency that they introduce is not suitable for real-time video encryption and they still cannot prevent temporal drift.

III. Proposed Approach

As mentioned in Section I, we propose ROI encryption method that does not need to modify encoders. It performs the encryption and compression separately. Fig. 7 represents a scenario of privacy protection using ROI encryption in a surveillance system. We propose a method of ROI encryption in Section IIIA and a drift mitigation method in Section IIIB.

A. ROI Encryption

In H.264, each slice constitutes NALs. In order to encrypt ROI, a NAL unit which contains ROI needs to be encrypted. In our encryption method, we adopt selective encryption of ROI texture because the visual concealment is maximized [4] when the texture of video is encrypted. Fig. 6 shows block diagram of ROI encryption system that we propose. Since the NAL is generated after entropy coding in compression, header information will be scrambled, and decoding will fail if encryption is performed over the entire NALs. To solve this problem and preserve a format compliant way, the proposed ROI encryption method selectively encrypts the ROI macroblocks which are obtained through inverse-entropy coding from the compressed bitstream. In addition, the energy is converged on DC coefficient, whereas AC coefficients are lowered in the matrix after DCT. In the end, almost all AC coefficients are changed to be zero since quantization. Using this feature, encoder performs the run-length coding. For this reason, the coding efficiency of entropy coding is decreased if the entire matrix is encrypted. Thus, the proposed method selectively encrypts only nonzero coefficients to prevent decrease of coding efficiency [18].

For the visual hiding of ROI, any cryptographic method such as Data Encryption Standard (DES), Advanced Encryption Standard (AES), Rivest, Shamir and Adleman (RSA) can be used. However, we apply Rivest's Cipher (RC4) for our experiments. Although DES, AES, and RSA algorithms ensure the high security level, they are not suitable for real-time video encryption because they have high computational cost. In addition, block cipher algorithms have padding problem which incurs slight increase of the size of cipher-text. On the other hand, RC4 has advantages that are small size of code, remarkable simplicity, and fast encryption speed rather than the other algorithms. Furthermore, since stream cipher does not need padding on the contrary to the block ciphers, RC4 is suitable to be used in real-time video encryption.

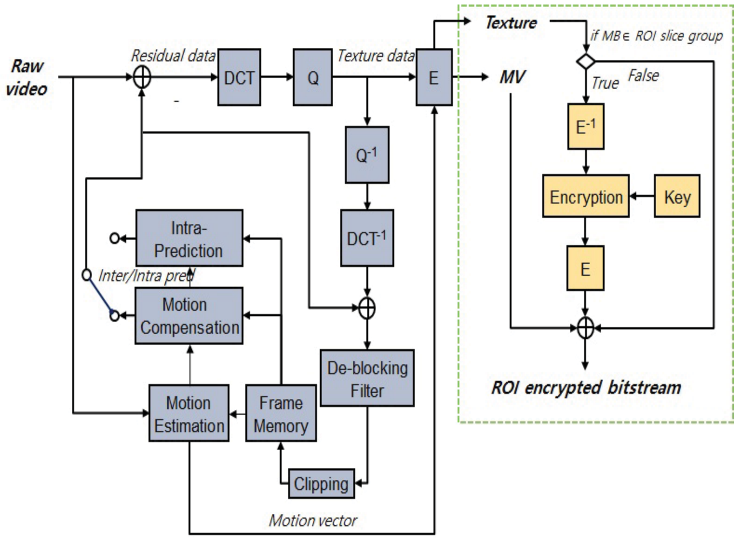


Fig. 6 Proposed ROI encryption method for H.264 bitstream compressed in common encoder

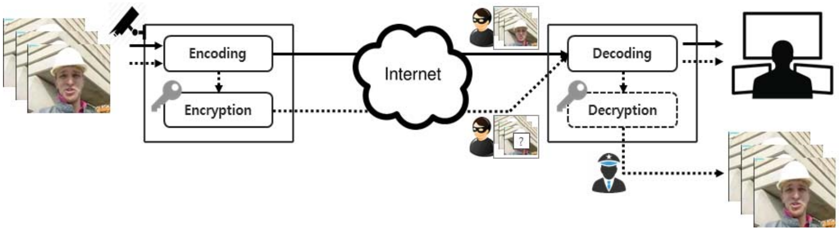


Fig. 7 Scenario of privacy protected video transmission. Unencrypted cctv video bitstream is parsed to encryption black box. Surveillance staffs and attackers can see privacy protected video, only officer who having key can decrypt original videos

After encrypting ROI and performing re-entropy coding, video bitstream with encrypted ROI is generated with a format compliant way. And visually the ROIs are protected when it is decoded.

B. Drift Mitigation

In intra prediction, each macroblock can be encoded using blocks that are already encoded within the current picture. However, this feature brings the spatial drift problem when ROI is encrypted. To prevent spatial drift, we use FMO's error resilient feature that is mentioned in Section IIC. Using FMO, ROI and non-ROI area are decoded separately since each area constitutes an independent slice group. It means intra prediction is not allowed between ROI slices and non-ROI slices. Shortly, it can prevent spatial drift. In this paper, the ROI is defined as the FMO map type 2 known as 'foreground and leftover' [15]. Detailed results of this are referred to in Section IV.

Although spatial drift is eliminated, temporal drift problem is remained. Temporal drift occurs due to inter prediction, therefore slice grouping method cannot eliminate temporal drift. Fig. 8 shows an illustration of temporal drift. Since ROI moves, non-ROI area in the current picture references encrypted ROI area in the previous picture, and drift error propagation occurs. To mitigate temporal drift, we propose periodic insertion of regenerated I-pictures when one or more ROIs appear in video.

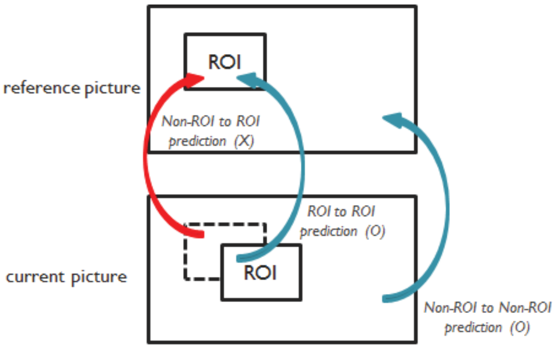

reference picture
ROI
Non-ROI to ROI prediction (X)
ROI to ROI prediction (O)
current picture
ROI
Non-ROI to Non-ROI prediction (O)

Fig. 8 Moving ROI affects temporal drift error propagation

Temporal drift problem remains until appearance of a new I-picture (IDR picture) in a video sequence. Temporal drift cannot be removed completely because pictures in compressed video bitstream have high dependency between each other. However, if GOP length becomes shorter, dependency between pictures are refreshed more frequently. Therefore, the temporal drift can be mitigated by inserting I-pictures periodically when ROI is detected. Fig. 9 shows insertion of I-picture that regenerated using previous I and P-pictures. After insertion, texture data of ROIs are encrypted as the method that is mentioned in Section IIIA. Then, it can obtain a drift-mitigated video bitstream with ROI encryption.

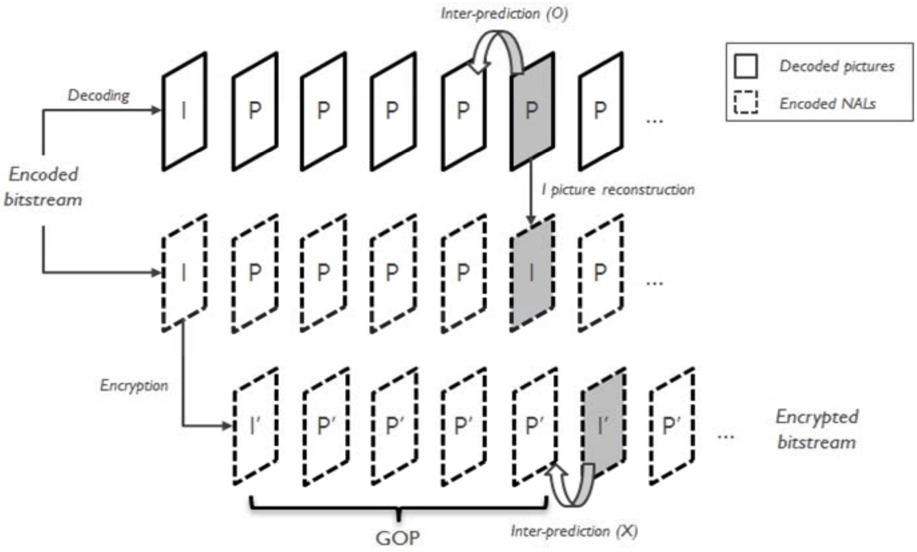
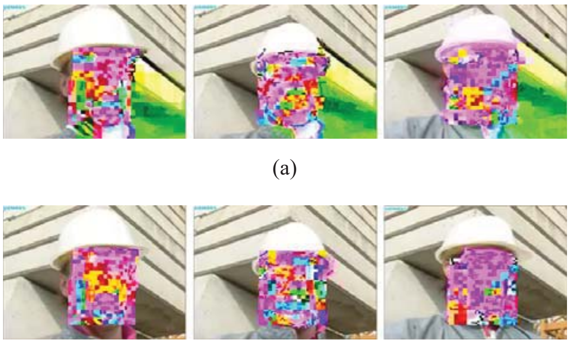


Fig. 9 Insertion of regenerated I-picture (I coded slice) between P-pictures

IV. Experimental Results

We implemented the proposed methods in JM v10.2 reference software [19] for verification of ROI encryption and drift mitigation. The test video "Foreman", which has QCIF (176*144) resolution and 30 fps, is used for the experiment, and one ROI area is defined as 64*80 size. Video sequences are encoded in Baseline Profile with an IPPP. GOP structure. All I-pictures are IDR and the number of reference frames is 1. In these experiments, we assume that surveillance cameras have ROI detecting and tracking functions.



(a)

(b)

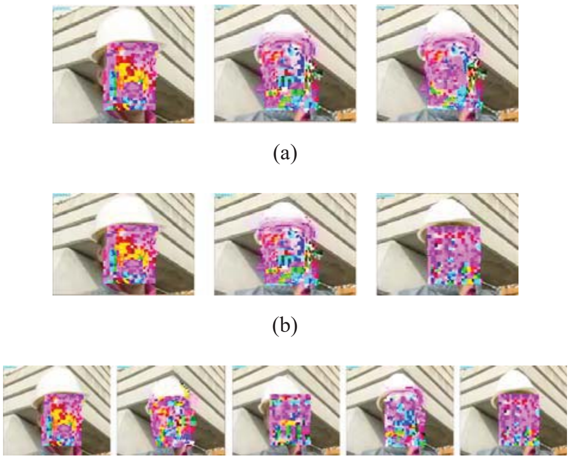
Fig. 10 ROI encrypted video (a) no using FMO, (b) using FMO

Fig. 10 shows two different results of decoded test videos according to the applying FMO. In Fig. 10 (a), spatial drift is occurred since ROI area is referenced by non-ROI area. Applying FMO however, Fig. 10 (b) shows that spatial drift is eliminated due to ROI, and non-ROI macroblocks are decoded independently.

Table I represents the bitrate overhead according to applying FMO. Average overhead is increased 3.49% because encoder does not perform intra prediction between slice groups when applying FMO, so compression efficiency is decreased.

TABLE I
Average Bit Rate and Bit Rate Overhead by Apply FMO

Without FMO	Applying FMO	Bit rate overhead
141.98 kbit/s	146.94 kbit/s	3.49%



(a)

(b)

(c)

Fig. 11 ROI encrypted video with mitigated temporal drift (a) only one I picture in the first frame (b) I picture period 30 (c) I picture period 15

Fig. 11 shows results of applying temporal drift mitigation method proposed in Section III *B*. Fig. 11 (a) video contains only one I-picture at first of video sequence and the others are P picture. As shown in the result, temporal drift is occurred as the ROI is moved to the left. Figs. 11 (b) and (c) show the results of applying temporal drift mitigation method; In Fig. 11 (b), I picture is inserted once in 30 frames. In Fig. 11 (c), I-picture is inserted once in 15 frames. Though there still remains temporal drift, error propagation is mitigated due to the inserted I pictures that refresh the dependency between pictures.

Table II represents the bitrate overhead according to applying proposed method. Assuming that Fig. 11 (a)'s bitrate overhead is 0, the overhead is increased 5.42% and 11.52%, respectively when the I picture is inserted once in 30 frames and 15 frames. Additionally, encoding time increases 1.56% and 5.37% respectively to obtain ROI encrypted bitstream. In this experiment, the shorter period of the I-picture insertion, the more often temporal drift is eliminated. Even though drift error is less propagated by inserting more I-pictures, there is side effect that bitrate overhead increases because I-picture is coded without prediction, so it needs relatively more bits than P-picture.

TABLE II
AVERAGE BIT RATE, OVERHEAD AND ENCODING TIME INCREASE ACCORDING TO APPLY PROPOSAL METHOD

I picture period	Bit rate	Overhead	Encoding time
No applied	146.94 Kbit/s	0 %	0 %
Once per 30 picture	161.87 kbit/s	5.42%	+1.56%
Once per 15 picture	180.51 kbit/s	11.56%	+5.37%

Table III contains PSNRs of decoded pictures shown in Fig. 11 for representing the visual hiding effect by applying ROI encryption. According to the results, the average PSNR difference between original and encrypted video is approximately 26. This result represents that the tested ROI encrypted video has sufficient visual hiding effect.

TABLE III
VISUAL HIDING EFFECT BY APPLY ENCRYPTION (PSNR)

	PSNR Y	PSNR U	PSNR V
Original video	37.39 db	40.61 db	41.78 db
Encrypted video	13.68 db	14.50 db	14.15 db
Difference	23.71 db	26.11 db	27.63 db

V. CONCLUSION

In this paper, we propose a H.264 video privacy protection method using ROI encryption in real time. In this method, the encryption process is separated from compression. In addition, we propose the drift mitigation method for the compressed bitstream encryption. The experimental results show verification of proposed methods and their tradeoff. In the tested video, the ROI occupies approximately 20% of footage and the ROI appears in the entire video sequence. In real world, however, the size of ROIs would be smaller than the one in our experiment and they would appear less frequently. Therefore, an overhead lower than the one measured in the experiments is expected, especially when the methods are adopted in the existing surveillance system. For future work, we consider control of the I picture insertion ratio according to the speed of ROIs since it is expected that it would be able to improve the coding efficiency.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (2015R1D1A1A01060793).

REFERENCES

[1] Yun Seong Ko, Kwang Hyuk Park and Chang Soo Kim. “Problem Analysis and Countermeasures Research through Security Threat Cases of Physical Security Control Systems.” Journal of Korea Multimedia Society Vol. 19, No. 1 (2016): 51-59

[2] Unterweger, Andreas, and Andreas Uhl. "Slice groups for post-compression region of interest encryption in H. 264/AVC and its scalable extension." Signal processing: image communication 29.10 (2014): 1158-1170

[3] T. Wiegand, G.J. Sullivan, G. Bjontegaard and A. Luthra. "Overview of the H. 264/AVC video coding standard." IEEE Transactions on circuits and systems for video technology 13.7 (2003): 560-576.

[4] Yeongyun Kim, Sung Ho Jin, Tae Meon Bae and Yong Man Ro. "A selective video encryption for the region of interest in scalable video coding." TENCON 2007-2007 IEEE Region 10 Conference. IEEE, 2007.

[5] Liu, Fuwen, and Hartmut Koenig. “A survey of video encryption algorithms.” Computers & Security 29.1, 2010

[6] Carrillo, Paula, Hari Kalva, and Spyros Magliveras. "Compression independent object encryption for ensuring privacy in video surveillance." 2008 IEEE International Conference on Multimedia and Expo. IEEE, 2008.

[7] Sk. Md. Mizanur Rahman, M. Anwar Hossain, Hussein Mouftah, A. El Saddik and Eiji Okamoto. "A real-time privacy-sensitive data hiding approach based on chaos cryptography." Multimedia and Expo (ICME), 2010 IEEE International Conference on. IEEE, 2010.

[8] Lingling Tong, Feng Dai, Yongdong Zhang and Jintao Li. "Restricted H. 264/AVC video coding for privacy region scrambling." 2010 IEEE International Conference on Image Processing. IEEE, 2010.

[9] Dufaux, Frédéric, and Touradj Ebrahimi. "A framework for the validation of privacy protection solutions in video surveillance." Multimedia and Expo (ICME), 2010 IEEE International Conference on. IEEE, 2010.

[10] Hosik Sohn, Esla T. AnzaKu, Wesley De Neve, Yong Man Ro and Konstantinos N. Plataniotis. "Privacy protection in video surveillance systems using scalable video coding." Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on. IEEE, 2009.

[11] Zeng, Wenjun, and Shawmin Lei. "Efficient frequency domain selective scrambling of digital video." IEEE Transactions on Multimedia 5.1 (2003): 118-129.

[12] Dubois, Loic, William Puech, and Jacques Blanc-Talon. "Smart selective encryption of cavlc for h. 264/avc video." 2011 IEEE International Workshop on Information Forensics and Security. IEEE, 2011.

[13] Andreas Unterweger, Kevin Van Ryckegem, Dominik Engel and Andreas Uhl. "Building a post-compression region-of-interest encryption framework for existing video surveillance systems." Multimedia Systems (2015): 1-23.

[14] Iqbal, Razib, Sharmeen Shahabuddin, and Shervin Shirmohammadi. "Compressed-domain spatial adaptation resilient perceptual encryption of live H. 264 video." Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on. IEEE, 2010.

[15] Datong Chen, Yi Chang, Rong Yan and Jie Yang. "Tools for protecting the privacy of specific individuals in video." EURASIP Journal on Applied Signal Processing 2007.1 (2007): 107-107.

[16] Dufaux, Frederic, and Touradj Ebrahimi. "Scrambling for privacy protection in video surveillance systems." IEEE Transactions on Circuits and Systems for Video Technology 18.8 (2008): 1168-1174.

[17] Unterweger, Andreas, Jan De Cock, and Andreas Uhl. "Bit Stream Based Encryption for Regions of Interest in H. 264/AVC Videos with Drift Minimization. "

[18] Peng, Fei, Xiao-wen Zhu, and Min Long. "An ROI privacy protection scheme for H. 264 video based on FMO and Chaos." IEEE transactions on information forensics and security 8.10 (2013): 1688-1699.

[19] Tourapis, Alexis Michael, Karsten Sühring, and Gary Sullivan. "H. 264/MPEG-4 AVC reference software manual." Geneva, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q 6 (2007).