ENSURING PRIVACY IN A DISTRIBUTED VIDEO CODING SURVEILLANCE SCENARIO

J. L. Martínez[1], J.M. Villalón[2], J.H. Parreño[2], P. Cuenca[2] and H. Kalva[3]

[1]Architecture and Technology of Computing Systems Group. Complutense University
Ciudad Universitaria s/n, 28040 Madrid, SPAIN
joseluis.martinez@fdi.ucm.es
[2] Instituto de Investigación en Informática de Albacete (I3A). University of Castilla-La Mancha.
Campus Universitario, 02071 Albacete, SPAIN.
{josemvillalon, pcuenca}@dsi.uclm.es
[3]Department of Computer Science and Engineering. Florida Atlantic University.
777 Glades Rd, Boca Raton, FL, 34341, USA
hari@cse.fau.edu

ABSTRACT

In this paper, a permutation scheme is introduced for Distributed Video Coding. The main goal is to preserve privacy and security in video surveillance video communications and can be adapted for other scenarios. The proposed approach consists to apply permutation based encryption to the DVC scenario. The permutation is defined by a secret key which is required at the decoder for decompression providing security. Simulation results show that the reference architecture and the permutation-based proposed have similar rate distortion performance.

Index Terms— DVC, video surveillance, privacy, identity, selective encryption

1. INTRODUCTION

Some years ago, *Distributed Video Coding* (DVC) appeared in the field of video coding as a solution for low complexity video encoding [1]. In DVC, the source statistics are exploited totally or partially at the decoder side and thus, the complexity is moved from encoders to decoders. DVC paradigm is based on two major Information Theory results: the Slepian–Wolf [2] and Wyner–Ziv theorems [3]. This innovator paradigm offers support for many applications which require low power / low complexity encoders [4]. Video surveillance cameras are one example of such applications where DVC can be used to keep the complexity low at expense of increasing the computation of the monitoring-decoder system.

Video surveillance systems are widely adopted in many strategic places such as banks, public transportation, airports markets or shops. In this framework, privacy rights are beginning to gain importance. Although the perception of security is demanded, the rightful fear of breach of privacy is turning into a significant concern. Selective encryption of faces can be seen as a solution to the privacy invasion. In this case, parts of the video can be encrypted and decrypted only for authorized persons in anytime. This process needs to be reversible and therefore a simpler solution as removing identities by blurring the portions is not acceptable. With selective encryption when an activity is declared as suspicious, the identities can be recovered. In this field, most of the solutions are specific to video compression algorithms used and require modification to the video encoders [5-7]. Furthermore, most of them are based on traditional video codecs which are not the most suitable solutions for this scenario as DVC theory postulates [1]. When encryption or scrambling techniques are introduced the spatial and temporal correlation of video sequence is lower and then, when this scrambled video is encoded with traditional video codecs, it denotes in an increase of bitrate due to poor exploration of the temporal and spatial correlations. On the contrary, in DVC this temporal correlation is exploited at the decoder side and thus, the impact of the bitrate increasing is much lower.

At this point, this paper presents a solution based on DVC that meets the requirements of privacy and security needs. The proposed approach consists to apply permutation based encryption to the DVC scenario. Due to the fact that the temporal correlation in DVC are exploited at the decoder side, the permutation based scheme offers less impact than approaches based on traditional video coding.

Accordingly, this paper is organized as follows: section 2 presents a brief overview of encryption techniques adopted in the literature; section 3 describes the DVC surveillance architecture proposed in this paper; section 4 carries out a performance evaluation for the proposed architecture; and, finally, some final remarks are presented in Section 5.

2. BACKGROUND

Privacy is a big problem in video surveillance due to unsuspecting person´s identity or license tags of vehicles may be recorded and saved on digital media and backups. Law enforcement agencies have legitimate needs to recorder or monitor surveillance video. Therefore the system should achieve a tradeoff between balancing privacy and security. The solutions adopted in this framework are based on reversible encryption with effective key management which presents security, hides identity completely, compression independent, interoperates with communication infrastructure and survives recoding / transcoding as main characteristics.

In 2005, in [5], the authors proposed the use of computer vision techniques in order to select the region of interest such that a face recognition system. With this method, although the privacy is maintained, the security meets are not met due to irreversibility of the encryption process. Then, in [6], *Senior et al*. proposed an invertible cryptographic obscuration in JPEG based on data encryption standard. Finally, in the same year, *Martínez-Ponte et al.* in [7] proposed a medical application for automatic detection, tracking, labeling and obscuring in real time was developed.

Some year later, in 2008, in [8], the authors presented a bitstream domain (i.e. mean after motion- compensation and DCT-quantization processes and before the entropy coding) encryption technique using secure shape and texture set partitioning in hierarchical trees. This technique fails because any transcoding technique requires decoding the entire frame first. Later, *Dufaux et al.* in [9] presented a scrambling technique which consists of detecting the region of interest and then the signs of selected transform coefficients are scrambled. Therefore, the decoded video will have blocky region unless a proper key is used for de-scrambling. Basically, the key is to know the pseudo randomly inverting pattern. This approach was applied to MPEG-4 in [9], H.264/AVC in [10] and DVC in [11]. As far as the authors of this paper know, this is the only work related to privacy on DVC. All these scrambled based [9-11] approaches fail into they are compression algorithm specific and cannot survive transcoding.

All the existing approaches are video compression specific, need modification to the encoders and decoders, need customized video codecs, forces a video compression solution on a system and cannot survive any modification to the encoded video. Therefore, in order to solve that problem, the present approach is applied before the compression algorithm and then survives transcoding and recording other formats (such as DVC to H.264 [12]). The present approach uses provably secure permutation based encryption and this enables flexible surveillance system that can select compression algorithms that suite their needs.

3. PROPOSED ARCHITECTURE

3.1. Overall Architecture

This section presents a DVC scenario designed to meet the requirements presents in a video surveillance system. Figure 1 shows the system. The Region of Interest (ROI) module performs the object detection. The procedure of determining object is out of the scope of this work. Based on this selection, the ROI permutation module, this part of the frame is encrypted using a key. The video with the encrypted parts can be encoded and decoded with DVC video codec and played on any standard display but these parts will remain encrypted. The procedure of how these parts are encrypted which is based on permutations will be explained later.

As part of the DVC encoder, the frame split module (1) separates the input frames into Key frames (K) (those that will be encoded using traditional video codecs) and Wyner-Ziv (WZ) frames (those that will be encoded following the Wyner-Ziv paradigm). The K frames are encoded using H.264/AVC Intra mode (2) [13]. On the other hand, WZ frames are sent to the WZ encoding procedure, where the information is firstly quantized (3a), bitplanes are extracted (3b); in (3c) each bitplane is independently channel encoded and several parity bits are calculated, which are stores in a buffer (3d).

On the decoder side, initially K frames are decoded by a H.264 decoder [13] (4). From these frames, it is calculated a Side Information (SI) [1] (5), which represents an estimation for each non-present original WZ frame. For this estimation, the Correlation Noise Model (CNM) module (7a) generates a Laplacian distribution, which models the residual between SI and original frame. Afterwards, SI and CNM is sent to the turbo decoder, which correct differences of SI and original frame by means of iterative decoding, which requests several parity bits to the encoder through the feedback channel. Finally, decoding bitplanes are reconstructed in module (7c).

Once the video is decoded, the video remains encrypted and a key is necessary for decrypting the video. Therefore the video can be displayed with encrypted region (those individuals who has not the key) and also perfectly by who knows the key. Since encryption rearranges the pixel in a block, the correlation of these blocks is decreased. By using this technique over traditional video codecs leads to increase in bitrate and the amount of increase depends on the content and the number of block encrypted because the correlation is exploited at the encoder side. On contrary, in DVC, due mainly to the fact that the correlation is exploited at the decoder the impact of bitrate increases is lower as it will be shown in next section.

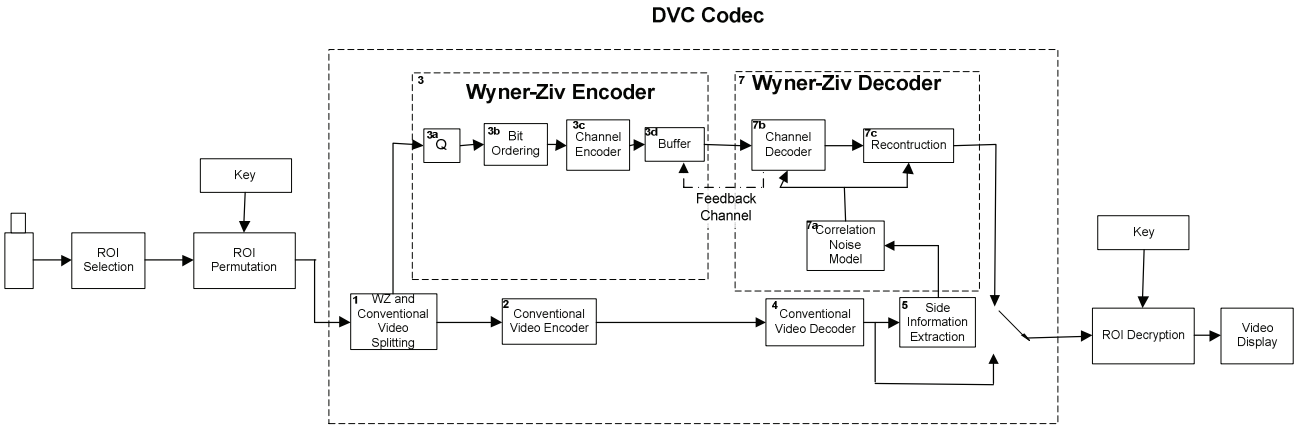


Figure 1. Proposed Architecture

3.2. Encryption

In our system, the encryption is performed before encoding and it makes this system compression independent. The encryption used in this DVC-based scenario is based on permuting pixel values using pseudo-random permutations. The generating process of these permutations is based on logarithmic signatures [14], and uses a secret pass as a key. The regions to be encrypted in a video are encrypted on a block basis. In this work, the block size was fixed to 16x16 in a trade-off between encryption strength and loss in correlation. A small block size is easy to attack because of small set of permutation possible. A larger block size significantly increases the encryption strength. In traditional video codecs, using larger block size result in a higher bitrate because of loss in correlation; on contrary, in DVC this has not a great impact due to the fact that the correlation are exploited at the decoder side. Moreover, keeping the encryption block size fixed is also necessary for avoid to transmit additional information to decoders. As the regions could not fit multiple of 16x16; the ROI is expanded to an area with width and height that are integral multiples of 16. The 16x16 blocks that cover the selected regions are determined and then a block based encryption is applied. Figure 2 shows an example of an original frame and its encrypted-faces version.

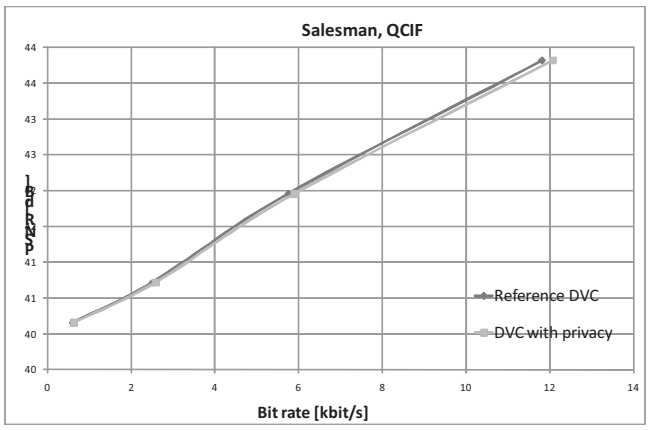
(a) (b)

Figure 2. a) original frame b) frame with encrypted faces

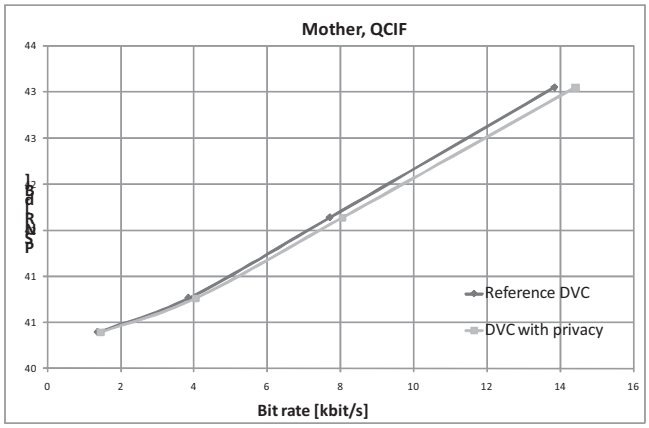
For each of the 16x16 block to be encrypted in a frame, a sequence of pseudorandom permutations (α_t) is applied to the cleartext sequence of blocks to yield the encrypted block sequence. Each key choice yields a sequence of random permutations (α_t) of periodicity $(16^2!) \approx 8.6 \times 10^{506}$. The reader could get more details in [14]. The size of the theoretical key-space (the number of logarithmic signatures generator) by far exceeds (2! x 3! x ··· x 256!) which make a brute force attack impossible. The encryption key can be generated dynamically based on the frame number and block number. The encryption key can be varied on a per-block or per-frame basis if desired. In our scheme, different blocks within a frame are encrypted with different and distinct elements of the random sequence of permutations. The frame with the encrypted blocks is then encoded using the DVC encoding paradigm.

4. PERFORMANCE EVALUATION

In order to evaluate the performance of the proposed architecture, two QCIF sequences were taken into account. These sequences are composed by 300 frames and they have with different movement and complexity features. The frame rate employed was 30 fps. The codec used is based on pixel domain Stanford architecture [1] and GOP length of 2. Number of BPs work as quantification, and values from 1 to 4 were considered to evaluate different rate-distortion points. Face detection for the experiments was done manually and face regions are input to the system. The proposed approach increases a little bit the bitrate of the encrypted and encoded video compared to the video encoded without encryption. Figure 3 shows the proposed encryption DVC architecture has a performance very close to the reference DVC architecture in all tested QCIF sequences. This is mainly because exploiting the redundancy at the decoder side does not affect too much in terms of bitrate increasing. On contrary, in traditional video codecs such as H.264 the due to the loss of correlation in the encrypted blocks leads to larger non-zero coefficients and quantization of these large coefficients increases the distortion in these blocks [15].



a)



b)

Figure 3. Rate Distortion coding efficiency comparison.
a) News, b) Mother

5. CONCLUSIONS

In this paper, an efficient privacy scheme based on permutation for DVC is introduced. The proposed architecture is proven to provide a good level of security in addition to an acceptable RD penalty loss. The compression efficiency of the architecture depicted in this paper is not negatively impacted by the introduction of the permutations.

6. ACKNOWLEDGES

This work has been jointly supported by the Spanish MEC and European Commission FEDER funds under grants "Consolider Ingenio-2010 CSD2006-00046" and "TIN2009-14475-C04-02/04", and by JCCM funds under grant "PEII09-0037-2328" and "PII2I09-0045-9916.

7. REFERENCES

[1] B. Girod, A. Aaron, S. Rane and D. Rebollo-Monedero, "Distributed video coding", Proceedings of the IEEE, Volume 93, Issue 1, January 2005

[2] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder", IEEE Transactions on Information Theory, Vol. 22, pp. 1-10, Jan. 1976

[3] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources", IEEE Transactions on Information Theory, pp. 471-480, Vol. 19, July 1973

[4] F. Pereira, L. Torres, C. Guillemot, T. Ebrahimi, R. Leonardi, S. Klomp, "Distributed Video Coding: Selecting the most promising application scenarios", Image Communication, v.23 n.5, p.339-352, June, 2008

[5] T. E. Boult, "PICO: privacy through invertible cryptographic obscuration", in Proceedings of the Computer Vision for Interactive and Intelligent Environment, pp. 27–38, November, 2005

[6] A. Senior, S. Pankanti, A. Hampapur, et al., "Enabling video privacy through computer vision," IEEE Security and Privacy, vol. 3, no. 3, pp. 50–57, May, 2005

[7] I. Martínez-Ponte, X. Desurmont, J. Meessen, and J. Delaigle, "Robust human face hiding ensuring privacy," in Proceedings of the Workshop on the Integration of Knowledge, Semantics and Digital Media Technology, Montreux, Switzerland, April 2005.

[8] K. Martin and K. N. Plataniotis, "Privacy protected surveillance using secure visual object coding", IEEE Transactions of Circuits and Systems for Video Technology, vol. 18, no. 8, pp. 1152–1162, August, 2008

[9] F. Dufaux and T. Ebrahimi, "Scrambling for privacy protection in video surveillance systems", IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 8, pp. 1168–1174, August, 2008

[10] F. Dufaux and T. Ebrahimi, "H.264/AVC Video Scrambling for Privacy Protection", in Proceedings of IEEE International Conference on Image Processing, San Diego, October, 2008.

[11] M. Ouaret, F. Dufaux and T. Ebrahimi, "Enabling Privacy For Distributed Video Coding by Transform Domain Scrambling", in Proceedings of SPIE Visual Communications and Image Processing, San Diego, USA, 2008.

[12] J.L. Martinez, H. Kalva, G. Fernandez-Escribano, W.A.C. Fernando and P. Cuenca, "Wyner-Ziv to H.264 video transcoder", in Proceedings of IEEE International Conference on Image Processing, pp. 2941 – 2944, Cairo, November 2009

[13] ISO/IEC International Standard 14496-10:2003, "Information Technology – Coding of Audio – Visual Objects – Part 10: Advanced Video Coding"

[14] S. S. Magliveras and N. D. Memon, "Algebraic properties of cryptosystem PGM," Journal of Cryptology, vol. 5, no. 3, pp. 167–183, 1992

[15] P. Carrillo, H. Kalva and S. Magliveras, "Compression independent object encryption for ensuring privacy in video surveillance", IEEE International Conference on Multimedia and Expo, pp. 273 – 276, Hannover, Germany, June, 2008.