

# Enabling Privacy For Distributed Video Coding By Transform Domain Scrambling

Mourad Ouaret, Frederic Dufaux and Touradj Ebrahimi

Institut de Traitement des Signaux

Ecole Polytechnique Federale de Lausanne (EPFL), CH-1015, Lausanne, Switzerland

*{mourad.ouaret,frederic.dufaux,touradj.ebrahimi}@epfl.ch*

## ABSTRACT

In this paper, a novel scheme for video scrambling is introduced for Distributed Video Coding. The goal is to conceal video information in several applications such as video surveillance and anonymous video communications to preserve privacy. This is achieved by performing a transform domain scrambling on both Key and Wyner-Ziv frames. More specifically, the sign of the scrambled transform coefficient is inverted at the encoder side. The scrambling pattern is defined by a secret key and the latter is required at the decoder for descrambling. The scheme is proven to provide a good level of security in addition to a flexible scrambling level (i.e the amount of distortion introduced). Finally, it is shown that the original DVC scheme and the one with scrambling have a similar rate distortion performance. In other words, the DVC compression efficiency is not negatively impacted by the introduction of the scrambling.

**Keywords:** Media Security, Privacy, Scrambling, Transform Domain, Distributed Video Coding.

## 1. INTRODUCTION

Recently, more research is performed in the field of Distributed Video Coding (DVC)<sup>1</sup>, a new paradigm for video compression. In DVC, the source statistics are exploited at the decoder side. In a practical scenario, this implies low power/low complexity encoders. Therefore, DVC is attractive for a wide range of real life applications where the computational power is sparse at the encoder. Hardware surveillance cameras is one important example where DVC can be used to keep the complexity low.

In <sup>2</sup>, the issue of privacy in video surveillance is addressed for JPEG 2000<sup>3</sup> compression. The privacy is ensured using a transform domain scrambling of regions of interest. It is shown that the technique provides a good security level. Furthermore, the scrambling is flexible and allows adjusting the distortion introduced into the image. Nevertheless, there is a small loss in the coding performance and a negligible complexity increase.

A framework for securing JPEG<sup>4</sup> images is introduced in <sup>5</sup>. It allows efficient integration and use of security tools to ensure confidentiality, integrity verification or conditional access. The latter is performed using a scrambling technique on the DCT coefficients.

The problem of scrambling regions of interest for video surveillance to preserve privacy is discussed in <sup>6</sup> and <sup>7</sup>. In addition to JPEG 2000, the case of MPEG-4<sup>8</sup> is also considered. The latter differs from JPEG and JPEG2000 in the different encoding frame modes available in MPEG-4. For JPEG and JPEG2000, only Intra mode is possible. In other words, each frame is encoded on its own without information from its neighboring frames. On the other hand, a frame can be encoded in the predictive mode in MPEG-4. This is why the scrambling has to pay attention not to introduce a drift in the prediction loop. Therefore, the scrambling has to be introduced outside of the motion compensation loop.

In this paper, a scheme ensuring privacy for DVC is introduced. Moreover, secure JPEG is used to encode the Key frames. For the Wyner-Ziv frames, parity bits are generated for the scrambled DCT coefficients by introducing the **DCT coefficient scrambler** prior to the Wyner-Ziv encoder. To recover the descrambled video at the decoder, both side information and reconstructed frame are scrambled as well. Finally, the comparison of the original and modified DVC schemes shows that they have a similar rate distortion performance.

The paper is structured as follows. Initially, the paradigm of DVC and used DVC scheme are introduced in section 2. Then, the scheme is modified by introducing the **DCT coefficient scrambler** at both encoder and decoder side in section 3. In section 4, the security level of the introduced scheme is evaluated. Both schemes

are compared in terms of rate distortion performance in section 5. Finally, some concluding remarks are drawn in section 6.

## 2. DISTRIBUTED VIDEO CODING

DVC is the consequence of information-theoretic bounds established by Slepian and Wolf<sup>9</sup> for distributed lossless coding, and by Wyner and Ziv<sup>10</sup> for lossy coding with decoder side information. In a practical scenario, lossy coding is used. In this paper, we consider the DVC architecture from<sup>11</sup> as illustrated in Figure 1. The Wyner-Ziv

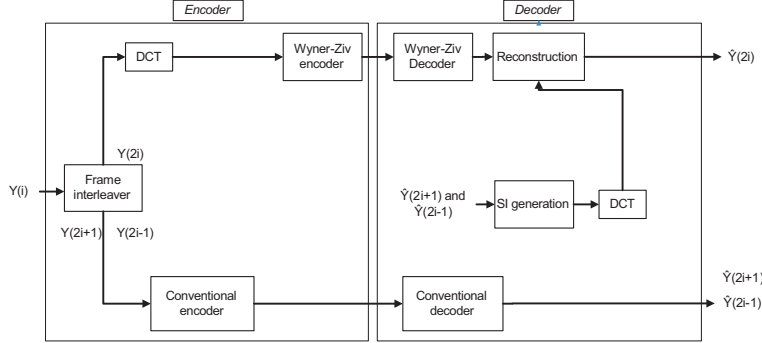


Figure 1. DVC scheme (GOP=2)

encoder operates in the DCT domain. In other words, an interleaved turbo encoder is used to generate parity bits for the quantized DCT coefficients. The case where the Group Of Pictures (GOP) is equal to two is considered and JPEG is used as the Key frame codec.

The conventionally decoded previous and forward Key frames are used to generate side information by motion compensated interpolation. To exploit the side information, the decoder assumes a statistical model, which is a Laplacian distribution of the difference between the individual DCT coefficients of the original Wyner-Ziv frame and the side information. The decoder combines the side information and the received parity bits to recover the original frame. For more details on the used DVC scheme, see <sup>1,10</sup> and <sup>11</sup>.

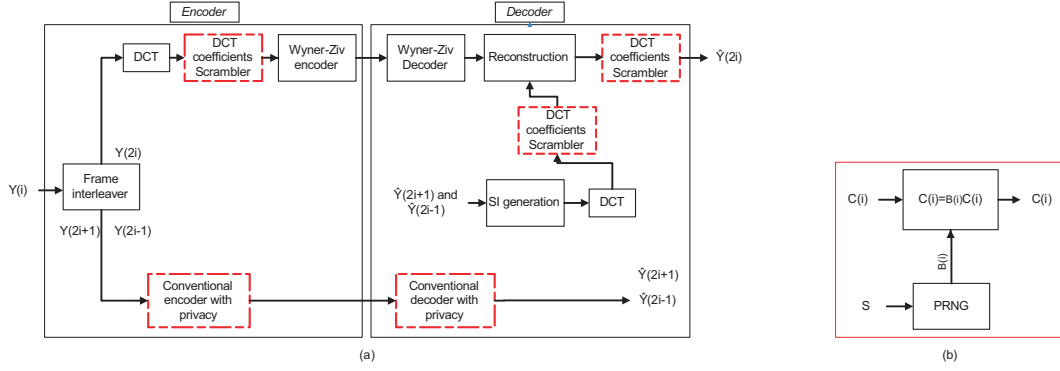
## 3. SCRAMBLING FOR DISTRIBUTED VIDEO CODING

To preserve privacy, the DVC scheme in Figure 1 is modified as shown in Figure 2(a). First, The issue of privacy over the key frames is discussed. For this purpose, secure JPEG<sup>5</sup> is used to encode the Key frames with the scrambling security tool. The latter performs a scrambling in the transform domain on the DCT coefficients. It is driven by a Pseudo-Random Number Generator (PRNG) to pseudo-randomly invert the sign of these coefficients.

Before discussing the privacy over Wyner-Ziv frames, the utility of the **DCT coefficients scrambler** box is explained. It takes as input DCT coefficients organized in 4x4 blocks. Thus, each block contains 16 coefficients, where the top-left one is the DC coefficient. The rest are AC coefficients. If a block is scrambled, some of its AC coefficients signs are inverted. For this purpose, a PRNG is used to decide which coefficients signs are inverted. For this purpose, a random sample  $B(i) \in \{-1, +1\}$  is multiplied by each AC coefficient within the scrambled block. This is illustrated in Figure 2(b) where  $s$  is the seed initializing the PRNG and  $C(i)$  is the input transform coefficient.

At the encoder, a segmentation mask depending on the scene scenario is required in order to decide which block should be scrambled. Otherwise, the scrambling can be just applied to the whole image (i.e. All DCT blocks). It is obvious that the seed parameter  $s$  constitutes the secret key that should be transmitted safely to the decoder. At the decoder, the seed  $s$  is used as input for the PRNG to generate the same sequence of numbers generated at the encoder. Thus, the same coefficients are inverted again to recover the original ones. In other words, descrambling comes down to applying the same **DCT coefficients scrambler** with the same key  $s$ . This should be applied to the reconstructed DCT coefficients at final stage. In addition, since the parity bits are generated

for the scrambled coefficients at the encoder, the side information DCT coefficients should be scrambled as well. This is because the DCT of the side information is an estimate of the DCT frame for which parity bits are generated. This would prevent spending more rate in the Wyner-Ziv decoding process.



**Figure 2.** a) DVC scheme with privacy (GOP=2). b) The **DCT coefficients scrambler** multiplies the DCT coefficient  $C(i)$  by the random number  $B(i)$ .

#### 4. SECURITY ISSUES

JPEG uses a  $8 \times 8$  block DCT while the Wyner-Ziv frames are transformed by a  $4 \times 4$  DCT-like transform. So in case the scheme is attacked by brute force, the latter will require a maximum of  $2^{63}$  and  $2^{60}$  operations per  $16 \times 16$  block for the Key and Wyner-ziv frames respectively. This makes the scrambling slightly less secure for the Wyner-Ziv frames. On the other hand, it has a better precision with respect to the region to be scrambled due to the smaller transform block size.

The reason why the scrambling is applied only to AC coefficients is to have the visual perception of the scrambling as shown in Figure 3, where the privacy is ensured and at the same time the scene is understood. The amount of distortion introduced is controlled by the total number of AC coefficients to which the scrambling is applied to. The smaller this number the less the introduced distortion is.

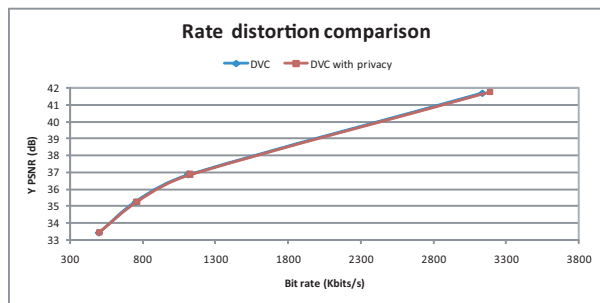


**Figure 3.** a) Secure JPEG. b) Wyner-Ziv transform domain scrambling.

#### 5. RATE DISTORTION PERFORMANCE COMPARISON

In this section, the effect of the scrambling on the compression efficiency is studied. For this purpose, the sequence *hall monitor* is used to evaluate the rate distortion performance of both DVC schemes. It simulates a video surveillance scenario case. The spatio-temporal resolution used is CIF@30 fps.

The following two scenarios are compared. The first one is where no scrambling is applied, which corresponds to the original scheme scenario. The second one is where scrambling and descrambling are applied respectively at the encoder and decoder. It's obvious from the plot in Figure 4 that both schemes have a similar rate distortion



DVC					
Intra		Wyner-Ziv		Total	
Bit rate (Kbits/s)	Y PSNR (dB)	Bit rate (Kbits/s)	Y PSNR (dB)	Bit rate (Kbits/s)	Y PSNR (dB)
401.2712	33.446277	95.46	33.42	496.7312	33.4331385
538.2828	35.150265	211.4	35.35	749.6828	35.2501325
732.4192	36.845116	380.02	36.87	1112.4392	36.857558
2303.0208	42.679045	834.38	40.7	3137.4008	41.6895225

DVC with privacy					
Intra		Wyner-Ziv		Total	
Bit rate (Kbits/s)	Y PSNR (dB)	Bit rate (Kbits/s)	Y PSNR (dB)	Bit rate (Kbits/s)	Y PSNR (dB)
401.2716	33.446277	99.64	33.41	500.9116	33.4281385
538.2836	35.150265	220.22	35.35	758.5036	35.2501325
732.418	36.845116	392.12	36.87	1124.538	36.857558
2303.0344	42.679045	885.29	40.77	3188.3244	41.7245225

Figure 4. The rate distortion performance of DVC with and without privacy.

performance. Thus, introducing the scrambling does not have a negative impact on the compression efficiency of the DVC scheme.

The tables in Figure 4 show the rate distortion points computed to compare both schemes. It can be seen that the scrambling disturbs the compression efficiency of the Wyner-Ziv codec more than the JPEG codec. The Wyner-Ziv decoder requires the laplace model relating the side information and the original frame DCT. When the side information DCT coefficients are scrambled, their statistics are disturbed which explains the slight increase in Wyner-Ziv rate for similar quality.

## 6. CONCLUSION

In this paper, an efficient scrambling scheme for DVC is introduced. It is based on transform domain scrambling. For the Key frames, secure JPEG is used. Further, a way to perform scrambling for the Wyner-Ziv codec is introduced. The scrambling preserves privacy with a sufficient level of security and a flexible scrambling level. It is shown that the scrambling is not worsening the DVC compression efficiency.

## REFERENCES

1. B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," in *Pro-ceedings of the IEEE*, **93**, pp. 71–83, January 2005.
2. F. Dufaux, M. Ouaret, Y. Abdeljaoued, A. Navarro, F. Vergnenegre, and T. Ebrahimi, "Privacy enabling technology for video surveillance," in *Proceedings of SPIE, Mobile Multimedia/Image Processing for Military and Security Applications*, S. S. Agaian and S. A. Jassim, eds., **6250**, May 2006.
3. A. Skodras, C. Christopoulos, and T. Ebrahimi, "The jpeg 2000 still image compression standard," *IEEE Signal Processing Magazine* **18**, pp. 36–58, September 2001.
4. G. Wallace, "The jpeg still picture compression standard," *Communications of the DCM* **34**(4), pp. 31–44, 1991.
5. F. Dufaux and T. Ebrahimi, "Toward a secure jpeg," in *Proceedings of the SPIE, Applications of Digital Image Processing XXIX*, A. G. Tescher, ed., **6312**, September 2006.
6. F. Dufaux and T. Ebrahimi, "Scrambling for video surveillance with privacy," in *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop*, IEEE Computer Society, 2006.
7. F. Dufaux and T. Ebrahimi, "Region-based transform-domain video scrambling," in *Proceedings of SPIE, Visual Communications and Image Processing 2006*, J. G. Apostolopoulos and A. Said, eds., **6077**, January 2006.
8. T. Ebrahimi and F. Pereira, *The MPEG-4 Book*, Prentice Hall, 2002.
9. J. Slepian and J. Wolf, "Noiseless coding of correlated in-formation sources," *IEEE Trans. on Information Theory* **19**, July 1973.
10. A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Information Theory* **22**, January 1976.
11. C. Brites, J. Ascenso, and F. Pereira, "Improving transform domain wyner-ziv video coding performance," in *International Conference on Acoustics, Speech and Signal Processing*, (Toulouse, France), May 2006.