Prediction restricted H.264/AVC video scrambling for privacy protection

L. Tong, F. Dai, Y. Zhang and J. Li

Drift error is a critical issue in video scrambling which is widely used to protect privacy in surveillance video. However, how to improve video coding efficiency while preventing drift error in video scrambling has not been covered in most studies. Proposed is a prediction restricted scheme for H.264/AVC privacy protected video scrambling. Experiments show that the proposed scheme dramatically raises coding efficiency.

Introduction: Video scrambling is driven by a key [1–3]. With a legitimate key, people can view the whole scene clearly. Otherwise, the privacy region is invisible with the scene still comprehensible. Fig. 1 (dark part excluded here) illustrates a transform-domain scrambling approach in H.264/AVC. In the encoder, a pseudorandom number generator (PRNG) initialised by a key is used to produce random number sequences. Then the sign of AC coefficients corresponding to privacy regions (assume having been extracted) is pseudorandomly flipped. Note that here unscrambled data is used in the prediction loop (Fig. 1*a*). In the decoder, authorised users descramble the coefficients before the prediction loop, so the same unscrambled data as in the encoder is used for prediction. In this case, the scene is fully reproduced. However, for unauthorised users, scrambled data is used for prediction (Fig. 1*b*) which is different from the one used in the encoder, incurring drift error. There are two cases in H.264/AVC video scrambling: drift error in intra-prediction and in inter-prediction. Drift error destroying comprehension of the scene, if no measures are taken in the encoder, not only the privacy region (solid line box in Fig. 2*a*) but also some part of the non-privacy region was invisible (dashed border in Fig. 2*b*).

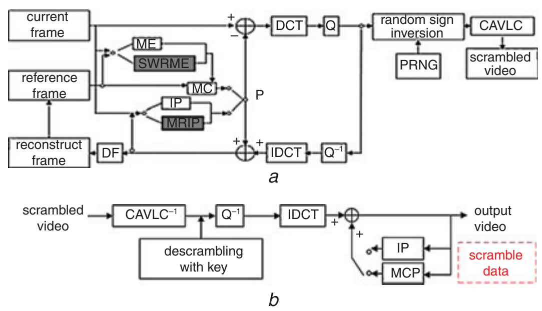

Fig. 1 *Transform-domain H.264/AVC video scrambling*
a Encoder
b Decoder

Fig. 2 *155th frame of Foreman scrambled with drift error*
a Original image
b Drift error in scrambled image

To prevent drift error in intra-prediction, in [1] the flexible macroblock ordering (FMO) mechanism of H.264/AVC was used. However, using FMO, intra-prediction across slice boundaries is not used [4], which would lead to reduction of coding efficiency. For drift error in inter-prediction, the scheme in [1] (abbreviated as FMIC) forced the macroblock (MB), which used scrambled data to predict, to be coded in intra-mode. Without inter-prediction for those forced intra-coding MBs, more data had to be encoded, which also incurred a decrease in coding efficiency.

Proposed prediction restricted video scrambling scheme: To improve coding efficiency while preventing drift error, mode restricted intra-prediction (MRIP) and search window restricted motion estimation (SWRME) are proposed, which modify the original intra-prediction and motion estimation in H.264/AVC (dark part in Fig. 1*a*).

MRIP: H.264/AVC uses intra-prediction from neighbouring pixels in a frame. Three block types (luma 4 × 4, 16 × 16 and chroma 8 × 8) are supported, with each type having different prediction modes. To prevent drift error in intra-prediction, modes using scrambled data to predict are forbidden for unscrambled blocks in MRIP, summarised as follows:

$$Mode_{ip} = \arg\min(COST(Block, Mode))\ Mode \in M \qquad (1)$$

$$M = \begin{cases} M_{scr} & Block \in S_{MRIP} \\ M_{org} & Block \in S_{nMRIP} \end{cases}$$

where Mod_{eip} indicates the intra-prediction mode of the current block, *Block* is the current block and *Mode* is the candidate intra-prediction mode. *COST(Block,Mode)* is the coding cost using *Mode*, the one that has minimal cost is the best mode. M_{org} is the set of intra-prediction modes for the current block type in H.264/AVC *Mscr* is the set of intra-prediction modes for unscramble blocks, which do not refer to scrambled blocks, S_{MRIP} is the set of MRIP control blocks, and S_{nMRIP} is the set of blocks outside the control of MRIP. As a block is predicted from its neighbouring pixels in intra-prediction, S_{MRIP} only involves unscrambled blocks at the boundary of the scrambled region, shown in Fig. 3*a*, where S_p is the privacy region.

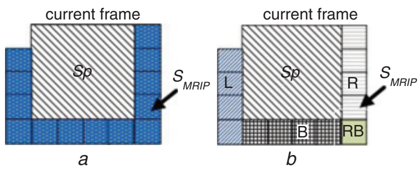

Fig. 3 *MRIP control blocks and their partition*
a MRIP control blocks
b Partition of MRIP control blocks

Here we consider the intra-prediction for a 4 × 4 block which has nine intra-prediction modes (Fig. 4). In this case, M_{org} comprises nine modes. According to MLIP, for blocks in S_{MRIP}, Mode0 is used in block set L, R and RB, as shown in Fig. 3*b*. L, B, RB and R are subsets of S_{MRIP} which is composed of 4 × 4 blocks. Mode1 is used in L, B and RB. Mode3 is used in R and RB. From mode4 to mode8, the same method is used to determine the block set they are used in. Mode2(DC) is a special case, which is not used for blocks in B and R. Regarding DC mode for blocks in B, A-D is set unavailable and I-L is used to predict in the encoder. In the decoder which does not know A-D is unavailable, prediction is formed by A-D and I-L, leading to error. If Mode2 is forbidden in B, this case will not occur, the same is true for R.

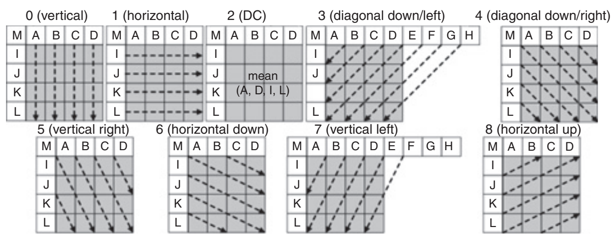

Fig. 4 *Nine intra_4 × 4 prediction modes*

As analysed above, according to M_{scr} of different blocks, S_{MRIP} is divided into L, B, R and RB (Fig. 3*b*). For variant block sizes, modes used for block set L, B, R and RB are illustrated in Table 1. As 16 × 16 and 8 × 8 blocks are in the same situation, we use one column to illustrate this case. Compared with FMO, only some intra-prediction modes for blocks in S_{MRIP} are forbidden, which has less influence on coding efficiency.

SWRME: In H.264/AVC, for the current block, the motion estimation algorithm searches an adjacent area in reference frames for a matching one, as follows:

$$b_r = \text{argmin}(SAD(B, B(x, y))) \quad B(x, y) \in S \qquad (2)$$

where b_r is the reference block, SAD is the sum of absolute differences between the current and the matching block, B is the current block, $B(x,y)$ is the search block, and S indicates the set of blocks in the search window.

Table 1: Intra-prediction modes used for MRIP control blocks

Block set	M_{scr} for 4 × 4 block	M_{scr} for 16 × 16 block and 8 × 8 block
L	Vertical, horizontal, DC, vertical right, diagonal down/right, horizontal down, horizontal up	Vertical, horizontal, DC, plane
B	Horizontal, horizontal up	Horizontal
R	Vertical, diagonal down/left, vertical left	Vertical
RB	Vertical, horizontal, DC, diagonal down/left, vertical left, horizontal up	Vertical, horizontal, DC, plane

To prevent drift error in inter-prediction with lower coding cost, SWRME modifies the set of blocks S in (2), which is only restricted to non-scrambling blocks, expressed as follows:

$$S = \begin{cases} S_w \cap S_{np\_r} & B \in S_{np} \\ S_w & B \in S_p \end{cases} \tag{3}$$

where S_w is the block set of the search window in H.264/AVC, S_p, S_{np} are the privacy region and the non-privacy region in the current frame, respectively. S_{np_r} is the block set of the non-privacy region in the reference frame (Fig. 5). For MBs in the non-privacy region, a suitable block in $S_w \cap S_{np\_r}$ which has minimal SAD will be found after motion estimation. For MBs in the privacy regions, the same motion estimation as H.264/AVC is carried out. As no MB is forced to code in intra-mode, coding efficiency of SWRME obviously exceeded that of FMIC.

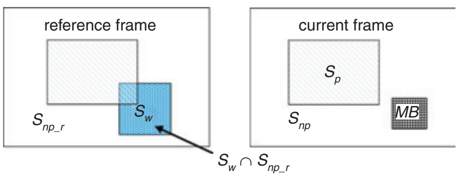

Fig. 5 *Search window restricted motion estimation*

Results: Experiments were carried out on JM15.1 under the baseline profile. Standard video test sequences Hall, Foreman and Carphone of QCIF format were used. The sign of AC coefficients corresponding to privacy regions is pseudorandomly flipped to protect privacy. Average bit rate and rate distortion performance of the proposed scheme are compared to the H.264/AVC without scrambling and the FMO and FMIC schemes [1]. Results are shown in Table 2 and Fig. 6. Table 2 shows the proposed scheme leading to bit rate overhead from 1.63 to 7.33%, with FMO and FMIC from 51.52 to 141.58% at worst. It can be observed from Fig. 6 that the proposed scheme can have minimal effect on coding efficiency, which obviously outperforms FMO and FMIC.

Table 2: Average bit rate (kbit/s) comparison with QP 28

Sequence	Original	Proposed (overhead)	FMO and FMIC (overhead)
Hall	71.66	72.83 (1.63%)	173.12 (141.58%)
Foreman	132.09	141.77 (7.33%)	200.14 (51.52%)
Carphone	77.71	82.71 (6.43%)	172.23 (121.63%)

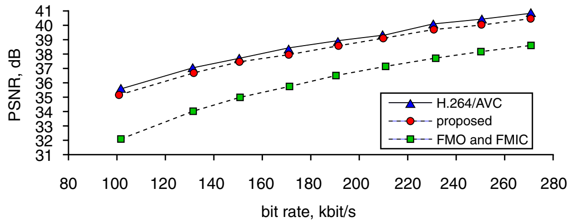

Fig. 6 *Rate distortion performance comparison for Foreman sequence*

Conclusion: MRIP and SWRME are proposed to improve coding efficiency while preventing drift error in H.264/AVC video scrambling. As some intra-prediction modes are available across privacy region boundaries and no MB is forced to code in intra-mode, the proposed scheme has higher coding efficiency than others. The performance of the proposed scheme is confirmed by experiments. In addition, SWRME is applicable to other motion estimation based coding standards.

Acknowledgments: This work was supported by the National Basic Research Program of China (973 Program, 2007CB311100) and the National Nature Science Foundation of China (60802028).

18 July 2009
doi: 10.1049/el.2010.2068
One or more of the Figures in this Letter are available in colour online.

L. Tong, F. Dai, Y. Zhang and J. Li (*Multimedia Computing Group, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, People's Republic of China*)

E-mail: tonglingling@ict.ac.cn

L. Tong: Also at the Graduate University of Chinese Academy of Sciences, Beijing, People's Republic of China

References

1 Dufaux, F., and Ebrahimi, T.: 'H.264/AVC video scrambling for privacy protection'. IEEE Int. Conf. on Image Processing, San Diego, CA, USA, October 2008, pp. 1688–1691
2 Martin, K., and Plataniotis, K.N.: 'Privacy protected surveillance using secure visual object coding', *IEEE Trans. Circuits Syst. Video Technol*, 2008, **18**, (8), pp. 1152–1162
3 Dufaux, F., and Ebrahimi, T.: 'Scrambling for privacy protection in video surveillance systems', *IEEE Trans. Circuits Syst. Video Technol*, 2008, **18**, (8), pp. 1168–1174
4 Wiegand, T., Sullivan, G.J., Bjntegaard, G., and Luthra, A.: 'Overview of the H.264/AVC video coding standard', *IEEE Trans. Circuits Syst. Video Technol*, 2003, **13**, (1), pp. 560–576