



# HEVC intra-frame drift cancellation matrix

Abhaya Dhathri Arige, Mihai Mitrea, Ismail Boujelbane

## ► To cite this version:

Abhaya Dhathri Arige, Mihai Mitrea, Ismail Boujelbane. HEVC intra-frame drift cancellation matrix. *Journal of Visual Communication and Image Representation*, 2019, 62, pp.56-67. 10.1016/j.jvcir.2019.04.014 . hal-02468475

**HAL Id: hal-02468475**

**<https://hal.science/hal-02468475>**

Submitted on 22 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License



# HEVC intra-frame drift cancellation matrix

A. Arige, M. Mitrea, I. Boujelbane

*IMT – Telecom SudParis, ARTEMIS Department, Evry, France*

---

## Abstract

This paper investigates the possibility of avoiding the HEVC intra-frame drift effect induced by additive modifications of the luma DCT/DST coefficients. The main novelty consists in solving the intra-drift problem in its general form: starting from the equations defining the 35 intra-prediction modes and the DCT/DST computation, a mask multiplicative matrix cancelling the intra-frame drift effect is computed. The relevance of the advanced method with respect to the state of the art studies is illustrated, discussed and quantitatively assessed.

**Keywords:** HEVC; additive modification; intra-drift free; theoretical computation; masking matrix

## 1. Introduction

The High Efficiency Video Coding (HEVC, a.k.a. H.265) standard is a joint project of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) standardization bodies. With respect to its ancestor MPEG-4 AVC (a.k.a. H.264), the main advantage of HEVC is to offer about 50% bit-rate reduction for similar perceptual video quality [1].

As its ancestors, HEVC computes the encoded frames by a sequence of four operations. The initial uncompressed (pixel<sup>\*</sup>) representation is first partitioned into blocks and a prediction operation is performed among and between these blocks. The prediction errors are subsequently transformed (either by DCT – Discrete Cosine Transform or DST – Discrete Sine Transform), quantized and finally entropic (lossless) encoded.

HEVC also inherits the structuring of the video sequences into Groups of Pictures (GOP), *i.e.* a series of successive frames. Any GOP contains one *intra*-coded frame (*I* frame). Additionally, it generally contains *inter*-coded frames that may be of two types: unidirectional predicted frames (*P* frames), and bi-directional predicted frames (*B* frames). While an *I* frame contains the complete information required for its decoding, the *P* and *B* frames require references to other *I/P* or *I/P/B* frames, respectively.

---

<sup>\*</sup> In HEVC specification, *pixels* are referred to as *samples*.

The relations among prediction blocks of a same *I* frame are defined by the so-called *intra*-prediction modes, see Fig. 1 (left) while the relations among prediction blocks in a *P/B* frame and prediction blocks in preceding/succeeding frames are defined by *inter*-prediction modes, see Fig. 1. (right). Consequently, the modification of one element in an encoded block in an *I* frame induces, through decoding, modifications in both subsequent encoded blocks in that *I* frame (the so-called *intra-drift* effect) and in the subsequent *P* and *B* frames (the so-called *inter-drift* effect). Such an intra/inter drift effect can be propagated several times (according to the visual content, encoding configuration, etc.), in a hardly to be predicted way.

The present paper theoretically studies the possibility of cancelling the HEVC intra-frame drift effect generated by additive modifications. It starts by considering the equations describing the intra-prediction modes and the *DCT/DST* computation and it derives the drift cancellation matrix (mask): when multiplying an additive modification of the encoded block by such a matrix, the drift effect is cancelled. The relevance of the advanced method with respect to the state of the art is methodologically discussed and quantitatively assessed in terms of maximal drift generating error inside the altered block as well as in terms of visual quality impact.



Fig. 1 Block prediction principle: intra-prediction (left) and inter-prediction (right)

The paper is organized as follows. Section 2 presents an overview of the state of the art drift avoidance solutions and introduces the HEVC stream syntax elements of interest in drift removal. Section 3 is devoted to method presentation: it first formulates the intra-drift effect as a matrix equation, then solves this equation under the HEVC steam syntax element constraints and finally theoretically brings to light the so-called *residual drift* inner to the HEVC standard. Section 4 illustrates the drift effect and presents the experimental results (both absolute assessment and relative benchmarking against a state of the art solution) while Section 5 concludes the paper and discusses future work directions.

## 2. State of the art

This section is structured at two levels. First (Section 2.1), 9 state of the art studies related to the drift-avoidance in MPEG-4 AVC and HEVC compressed domains are discussed. Secondly (Section 2.2), the main HEVC syntax elements of relevance for the HEVC drift avoidance are presented: coding tree units (CTU), coding tree blocks (CTB), coding unit (CU), prediction modes, transforms, ...

### 2.1. Overview of drift free solutions

The alteration of an encoded block can be induced either by errors occurring during the video transmission, storage and/or decoding or by voluntary operations, such as compressed-stream watermarking applications, where a watermark is imperceptibly and persistently inserted into the encoded blocks for

copyright/integrity verification purposes. Actually, the most recent intra drift-free studies for MPEG-4 AVC [2]-[8] and HEVC [9], [10] standards were conducted in relation to watermarking applications. From the conceptual point of view, intra drift-free studies can be structured into three categories subsequently referred to as *compensation-based* [2]-[5], *selection-based* [6],[7],[9],[10] and *cancellation-based* [8], see Table 1.

The compensation-based methods [2]-[5] consist in estimating the drift induced by the considered modification and in subsequently compensating it by complete decoding/re-encoding operations. This way, the drift can be completely removed at the expense of drastically increasing the computational complexity and of modifying the compressed stream parameters.

The selection-based methods [6],[7],[9],[10] consist in restricting the blocks (and the elements in the blocks) subjected to modifications and/or of imposing constraints on the modification statistics: only the blocks which are not involved in the prediction process and/or only the modifications which are not prone to drift effect are dealt with. This way, no drift effect occurs, the computational complexity is kept constant but the problem is solved only for some particular cases and not in its general statement. Such approaches already proved their efficiency for avoiding the HEVC intra drift. In this respect, the study in [9] considers the 4x4 blocks and individually investigates 5 typologies of intra-prediction modes. By considering the HEVC prediction equations, it is brought to light that 3 or 9 coefficients in each block can be modified by a bipolar additive modification (denoted by  $t$  or  $-t$ , where  $t$  is mentioned to be a small integer). It is stated that a similar approach can be derived for the other block sizes and it is indicated that, in the 8x8 block cases, 2 to 4 coefficients (according to the prediction mode) can be modified without drift effect. The study in [10] reconsiders the results in [9]. It deals only with 4x4 blocks and is based on the idea that preserving the most-right column and the bottom row elements would eliminate the drift. Hence, the 3x3 left-up elements in each block are individually processed in a way similar to the HEVC encoding procedure so as to identify 2 (out of 9) elements that can be modified by a bipolar additive alteration (denoted by  $\delta$  /  $-\delta$ , where  $\delta$  is also mentioned to be a small integer). The improvements with respect to [9] are evaluated by three image quality metrics, namely the PSNR (Peak Signal-to-Noise Ratio), the SSIM (Structural SIMilarity) and the VIFp (Visual Information Fidelity in pixel domain).

Table 1: Synopsis of state of the art approaches

	<b>compensation</b> ♂ <i>general</i> ♂ <i>high complexity</i>	<b>selection</b> ♀ <i>particular</i> ♂ <i>low complexity</i>	<b>cancellation</b> ♂ <i>general</i> ♂ <i>low complexity</i>
<b>AVC</b>	Ma [2], Huo [3], Gong [4], Zhang [5]	Ma [6], Chen [7]	Hasnaoui [8]
<b>HEVC</b>		Chang [9], Gaj [10]	<b>paper challenge</b>

The cancellation-based approach is meant to feature the advantages of both compensation and selection based approaches while getting rid of their disadvantages. The principle consists in pre-processing the additive alteration before applying it to the encoded blocks. In this respect, the study in [8] advances a solution devoted to the MPEG-4 AVC stream, structured on three incremental steps. First, the drift propagation effect induced by the additive modification is expressed as a function of the propagation mode (for each of the 13 MPEG-4 AVC prediction modes). Then, an algebraic (matrix) based equation for the drift effect, generic with respect to the prediction modes peculiarities, is derived. By solving this equation under the MPEG-4 AVC

stream syntax constraints, a cancellation matrix is finally obtained. By left/right multiplying the additive modification by this cancellation matrix, no drift intra drift effect occurs. Yet, the overall energy of the modification may be amplified, thus inducing stronger alterations in the respective block.

Table 1 presents a synoptic view on these three types of drift-avoidance methods for MPEG-4 AVC and HEVC standards. It can be noticed that, to the best of our knowledge, no cancellation-based drift-free HEVC method has been designed. The present paper takes this theoretical and experimental challenge.

## 2.2. HEVC stream syntax elements overview

This section recalls the HEVC syntax elements [1], [11], [12] of relevance for the intra-drift removal.

As earlier mentioned, the first encoding operations are GOP and frame partitioning. An *I* frame in HEVC is partitioned into coding tree units (CTUs), each CTU covering a rectangular area up to 64x64 samples (pixels) depending on the encoder configuration. The CTU regroups information related to the luminance and the chrominance of the samples, presented as the luma and chroma CTB (coding tree blocks), respectively. As the present study focuses on the luma drift effect, only the luma related information is presented in the sequel.

Each CTU is divided into coding units (CUs) that are signaled as intra or inter predicted blocks. A CU is then divided into intra or inter prediction blocks (PB) according to its prediction mode. For residual coding, HEVC supports two modes for partitioning a CU: PART\_2Nx2N and PART\_NxN. The former indicates that the PB size is the same as the CU size, *i.e.* 2Nx2N, while the latter signals the splitting of the CU into four equal-sized PBs.

The HEVC intra-prediction is achieved by 35 prediction modes, numbered by 0 to 34: the planar, the DC and 33 different angular (also referred to as directional) prediction modes. Once the frame is partitioned, the prediction mode is selected so as to minimize the rate-distortion cost. The predicted block is computed from the boundary most-right column and bottom row samples of previously encoded blocks.

The Intra-Planar prediction (the so-called Mode 0) uses the average values of two linear predictions to compute the predicted block samples as follows [11]:

$$\begin{aligned} P_{x,y} &= (PV_{x,y} + PH_{x,y} + N) \gg (\log_2 N + 1) \\ PV_{x,y} &= (N - y) \cdot R_{x,0} + y \cdot R_{0,N+1} \\ PH_{x,y} &= (N - x) \cdot R_{0,y} + y \cdot R_{N+1,0} \end{aligned} \quad (1)$$

where  $x$  and  $y$  are the coordinates of the sample to be predicted,  $N \times N$  is the size of the block while  $R_{x,y}$  and  $P_{x,y}$  stand for the reference and predicted samples, respectively.

The Intra-DC prediction (the so-called Mode 1) assigns the average value of reference samples to the predicted block samples.

The sample prediction equation for horizontal modes, *i.e.*, angular modes 2-17, is defined as follows [11]:

$$P_{x,y} = ((32 - w_x) * R_{i,0} + w_x * R_{i+1,0} + 16) \gg 5 \quad (2)$$

$$\begin{aligned} c_x &= (x * d) \gg 5 \\ x &= (x * d) \& 31 \\ i &= y + c_x \end{aligned}$$

The sample prediction equation for vertical modes *i.e.*, angular modes 18-34, is defined as follows [11]:

$$P_{x,y} = ((32 - w_y) * R_{i,0} + w_y * R_{i+1,0} + 16) \gg 5 \quad (3)$$

$$c_y = (y * d) \gg 5$$

$$w_y = (y * d) \& 31$$

$$i = x + c_y$$

In equations (2) and (3),  $R_{i,0}$  and  $R_{i+1,0}$  represent the reference samples, where  $i$  stands for reference sample index and  $P_{x,y}$  represents the sample to be predicted, where  $x, y$  are the spatial coordinates.  $w_y$  and  $w_x$  are the weighting between the two reference samples corresponding to projected subpixel location in  $R_{i,0}$  and  $R_{i+1,0}$ , which are calculated based on the projection displacement  $d$ .  $c_x$  and  $c_y$  represent pixel parameters corresponding to  $x$  and  $y$  coordinates. The parameters  $c_y$  and  $w_y$  depend only on  $y$  coordinate in vertical mode and  $c_x$  and  $w_x$  depend only on  $x$  coordinate for horizontal modes. ‘ $\gg$ ’ denotes a bitwise shift operation to the right and ‘ $\&$ ’ denotes logical AND operation.

The prediction errors are subsequently transformed by a  $T$  transform<sup>†</sup> calculated as follows [12]:

$$T(Y) = (Cf * Y * Cf^t) * 2^{-(S_{t1}+S_{t2})} \quad (4)$$

where :

- $Y$  is the prediction errors matrix to be transformed; its size can be 4x4, 8x8, 16x16, 32x32.
- $Cf$  is the HEVC forward multiplier matrix. For the transform blocks of size different from 4x4, the elements of the core transform matrices  $Cf$  are derived by approximating scaled  $DCT$  basis functions, under constraints such as limiting the dynamic range for transform computation or maximizing the precision and closeness to orthogonality when the matrix entries are specified as integer values. For the transform blocks of size 4x4,  $Cf$  is derived by approximating scaled  $DST$  basis functions.
- $t$  denotes the matrix transposal operation.
- $S_{t1}$  is calculated as  $(B + M - 9)$ , where  $B$  is the video bit-depth,  $M$  is equal to  $\log_2(N)$  and  $N \times N$  is the transform block size.
- $S_{t2}$  is calculated as  $(M + 6)$ , where  $M$  is equal  $\log_2(N)$  and  $N \times N$  is the block size.

The HEVC core transform matrix is based on the following approximation [12]:

$$Cf^t * Cf * \frac{2^{-(S_{t1}+S_{t2})}}{2^{S_{t3}}} \approx I_{N \times N} \quad (5)$$

where  $I_{N \times N}$  is the  $N \times N$  identity matrix (should not be taken for an  $I$  frame) and  $S_{t3}$  is 5, 4, 3 or 2 according to the transform block sizes: 4x4, 8x8, 16x16, 32x32 respectively. Note that although the  $Cf$  values are different for DCT and DST, equation (5) holds for both types of transforms.

Equations (1) to (5) bring to light that although similar in their principles, MPEG-4 AVC and HEVC are different in their computation, and that conducting a theoretical study on HEVC drift avoidance is not a trivial extension from MPEG-4 AVC case: HEVC has a more elaborated hierarchical structure, allows different

---

<sup>†</sup>In HEVC, both DCT (Discrete Cosine Transform) and DST (Discrete Sine Transform) can be alternatively used; in order to ensure a unitary presentation through the paper, the transform will be denoted by  $T$ .

block sizes to be defined, extends the definition of its prediction modes, changes the definition of the transforms, ... All these structural modifications should be reflected into a new theoretical drift-avoidance study, as presented in Section 3.

### 3. Method presentation

The presentation of the drift-avoidance method is structured in three sub-sections: the intra-drift problem formulation as a matrix equation (Section 3.1), its solving under the HEVC stream syntax element constraints (Section 3.2) and a discussion on the impact of the built-in HEVC approximation (Section 3.3).

#### 3.1. Matrix representation of intra drift effect

Consider a distortion matrix  $\tilde{W}$  to be added to the transformed residual block  $\tilde{R}$ ; the altered block  $\tilde{R}_w$  is obtained as:

$$\tilde{R}_w = \tilde{R} + \tilde{W} \quad (6)$$

At the decoder side, after applying the inverse transform  $IT$  (be it  $IDCT$  or  $IDST$ ), the altered pixel-domain residual block  $R_w$  is computed as following:

$$R_w = IT(\tilde{R} + \tilde{W}) \quad (7)$$

The linearity of the inverse transform function leads to:

$$R_w = R + W \quad (8)$$

where  $R$  is the initial pixel-domain residual block and  $W$  is the pixel domain representation of the distortion matrix.

Depending on the selected intra-prediction mode, the predicted block  $P$  is computed according to one of the equations (1), (2) or (3); then, it is added to the altered residual block to obtain the altered pixel-domain block  $X_w$ :

$$\begin{aligned} X_w &= P + R + W \\ &= X + W \end{aligned} \quad (9)$$

For a block of size  $N \times N$ ,  $X$  and  $W$  can be expressed in matrix form:

$$X = \begin{bmatrix} X_{0,0} & \dots & \dots & \dots & X_{0,N-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ X_{N-1,0} & \dots & \dots & \dots & X_{N-1,N-1} \end{bmatrix} \quad (10)$$

$$W = \begin{bmatrix} W_{0,0} & \dots & \dots & \dots & W_{0,N-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ W_{N-1,0} & \dots & \dots & \dots & W_{N-1,N-1} \end{bmatrix} \quad (11)$$

In order to avoid the drift effect, the additive distortion  $\tilde{W}$  should be preprocessed in such a way that it would

not result in any changing in the most-right column and bottom row of the host block in the pixel domain, as illustrated in Fig. 2. Be  $B_{i,j}$  a current block and  $B_{i,j+1}$ ,  $B_{i+1,j-1}$  and  $B_{i+1,j}$  a set of neighboring blocks that may use the block  $B_{i,j}$  in their prediction. The drift does not take place when the values at the most-right column and the bottom row elements of  $B_{i,j}$  are kept unchanged.

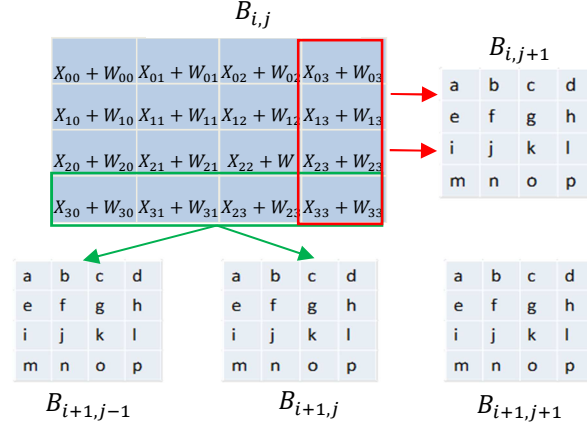


Fig. 2 Intra-frame drift avoidance principle

Consequently, the block distortion system considering the drift free condition in the compressed domain can be formulated according to (12):

$$\begin{cases} \tilde{R}_w = \tilde{R} + \tilde{W} \\ IT(\tilde{W}) = \begin{bmatrix} W_{0,0} & \dots & \dots & W_{0,N-2} & 0 \\ \vdots & \ddots & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & \vdots & 0 \\ W_{N-2,0} & \dots & \dots & W_{N-2,N-2} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{cases} \quad (12)$$

$W$  can be expressed as matrix multiplication:

$$W = E * W_1 * E \quad (13)$$

$$\text{where } W_1 \in M_{N \times N} \text{ and } E = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & 0 \\ 0 & \ddots & \ddots & \vdots & 0 \\ 0 & \dots & \dots & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, the compressed domain solution space considering the drift free distortion (denoted by further  $S_c$ ) can be expressed as:

$$S_c = \{\tilde{W} \mid \tilde{W} = T(E * W_1 * E)\} \quad (14)$$

### 3.2. Intra drift-free matrix

The domain solution space  $S_c$  described by (14) should be further restricted to the HEVC DCT/DST



computation formulas.

Consider  $\widetilde{W}$  belonging to  $S_c$ ; by applying the HEVC transform computation formula (4),  $\widetilde{W}$  can be expressed as:

$$\widetilde{W} = (Cf * (E * W_1 * E) * Cf^T) * 2^{-(S_{t1}+S_{t2})} \quad (15)$$

By considering equations (15) and (5),  $\widetilde{W}$  can be expressed as:

$$\begin{aligned} \widetilde{W} &\approx Cf * E * \left( Cf^T * Cf * \frac{2^{-(S_{t1}+S_{t2})}}{2^{S_{t3}}} \right) * W_1 * \left( Cf^T * Cf * \frac{2^{-(S_{t1}+S_{t2})}}{2^{S_{t3}}} \right) * E * Cf^T \\ &\approx \left( Cf * E * Cf^T * \frac{2^{-(S_{t1}+S_{t2})}}{2^{S_{t3}}} \right) * \left( Cf * W_1 * Cf^T * 2^{-(S_{t1}+S_{t2})} \right) * \left( Cf * E * Cf^T * \frac{2^{-(S_{t1}+S_{t2})}}{2^{S_{t3}}} \right) \end{aligned} \quad (16)$$

From equations (4) and (16),  $\widetilde{W}$  can be expressed as:

$$\begin{aligned} \widetilde{W} &\approx \frac{T(E)}{2^{S_{t3}}} * T(W_1) * \frac{T(E)}{2^{S_{t3}}} \\ &\approx \frac{\tilde{E}}{2^{S_{t3}}} * \widetilde{W}_1 * \frac{\tilde{E}}{2^{S_{t3}}} \end{aligned} \quad (17)$$

Therefore, the HEVC compressed domain solution space  $S_c$  can be defined according to:

$$S_c = \left\{ \begin{array}{l} \widetilde{W} \mid \widetilde{W} = \frac{\tilde{E}}{2^{S_{t3}}} * \widetilde{W}_1 * \frac{\tilde{E}}{2^{S_{t3}}} \\ \text{where } W_1 \in M_{N \times N} \text{ and } E = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \ddots & : & 0 \\ 0 & \ddots & \ddots & : & 0 \\ 0 & \dots & \dots & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \right\} \quad (18)$$

Thus, in order to prevent the intra-drift effect, the distortion induced in the compressed domain should be left and right multiplied by the matrix  $M_d = \tilde{E}/2^{S_{t3}}$ , prior to its addition to the host transformed residual block. Note that  $M_d$  acts as a drift masking matrix: it does not eliminate/correct the additive modification but it ensures that its effect will be restricted to the corrupted block.

### 3.3. Residual intra drift

While equations (16) – (18) are derived solely by matrix computations on HEVC specifications, they also consider equations (5) which is an approximation: in order to obtain beneficial properties (like limiting the dynamic range for transform computation or maximizing the precision and closeness to orthogonality when the matrix entries are specified as integer values), the HEVC transform computation relies on one approximation. Consequently, the intra-drift cancellation (18) we derived may eventually allow modifications of the values at the most-right column and the bottom row elements in the block and, implicitly, a residual (un-cancelled) drift effect. Yet, such modifications are *a priori* expected to be imperceptible. Actually, the HEVC transform computation approximation does not induce *a priori* any visual artifacts (criterion essential to its very acceptance in the standard). Moreover, the prediction operations are linear with respect to the distortions, as demonstrated hereafter by equations (20) and (22): hence, very small modifications are also

expected to have very small visual impact.

The prediction modes 0 and 1 are linear by their very definitions (average values over linear predictors or over samples, respectively).

Considering now a modification  $\varepsilon$  in the reference pixel values used for predicting the pixel sample  $P_{x,y,\varepsilon}$  and apply (2) for sample prediction with horizontal angular modes;  $P_{x,y,\varepsilon}$  can be computed according to:

$$\begin{aligned} P_{x,y,\varepsilon} &= \left( (32 - w_x) * (R_{i,0} + \varepsilon) + w_x * (R_{i+1,0} + \varepsilon) + 16 \right) \gg 5 \\ &= \left( (32 - w_x) * R_{i,0} + w_x * R_{i+1,0} + 16 \right) \gg 5 + 32 * \varepsilon \gg 5 \end{aligned} \quad (19)$$

Therefore,  $P_{x,y,\varepsilon}$  can be expressed as:

$$P_{x,y,\varepsilon} = P_{x,y} + \varepsilon \quad (20)$$

where  $P_{x,y}$  is the predicted pixel value if no modification is induced on the reference pixel values.

A similar behavior can be identified for the vertical prediction modes, defined by equation (3):

$$\begin{aligned} P_{x,y,\varepsilon} &= \left( (32 - w_y) * (R_{i,0} + \varepsilon) + w_y * (R_{i+1,0} + \varepsilon) + 16 \right) \gg 5 \\ &= \left( (32 - w_y) * R_{i,0} + w_y * R_{i+1,0} + 16 \right) \gg 5 + 32 * \varepsilon \gg 5 \end{aligned} \quad (21)$$

yielding the same relation as in (20):

$$P_{x,y,\varepsilon} = P_{x,y} + \varepsilon \quad (22)$$

#### 4. Illustrations and experimental results

While the main contribution of the paper consist in theoretically computing the drift cancellation (masking) matrix expressed by (18), the present section has as objectives to illustrate the drift effect, to objectively evaluate the effectiveness of the advance method and to benchmark it against the state of the art solutions.

The experimental results reported in this Section 4 are carried out on corpus of 8 video sequences, at 3840x2160 resolution, encoded by the HEVC main profile. These 8 video sequences processed in the experiments sum-up to 375 *I* frames: this size ensures the 95% statistical relevance of the results (as detailed for each experiment here-after). They are encoded with the following configuration: Rate Control / qCompress: CRF-28.0 / 0.60, Keyframe min / max / scenecut / bias: 5 / 5 / 40 / 5.00. The rest of the encoding parameters are left to the encoder choice, thus ensuring a variety of the block sizes, prediction modes and quantizing steps. The processed video content covers heterogeneous indoor and outdoor scenes, with a large variety of motion content as well as with arbitrarily lighting conditions, as illustrated in Fig. 3. In order to cross-check the generality of our results with respect to the type of content, a second (cross-checking) corpus of the same size, resolution and encoded principles but representing a different content is considered in Section 5. Concerning the content choice, the same heterogeneity principle is kept, as illustrated in Fig. 4.

The block modification under HEVC stream syntax constraint is achieved by software tools we developed in C/C++ so as to complement the HEVC reference software available at [13].



Fig. 3 Illustration of the content included in the corpus processed in Section 4, for assessing the residual drift effect



Fig. 4 Illustration of the content included in the corpus processed in Section 5, for cross-checking the results presented in Section 4

#### 4.1. Illustrations

The drift cancellation mechanism is illustrated in Fig. 5 by representing the absolute difference between the luminance components of the original  $I$  frame and of the decoded modified frame; these differences are illustrated as dark-blue to light-yellow pseudo-color images (the so-called *parula* colormap, available in both matlab and python). Hence, a dark blue pixel in Fig. 5 indicates no difference at that location between the original HEVC and the drift affected HEVC sequences. Conversely, a light yellow pixel in Fig. 5 indicates a large local difference (up to 255) between the two investigated HEVC sequences.

Consider an  $I$  frame and a transformed block of fixed size (the cases  $4 \times 4$ ,  $8 \times 8$ ,  $16 \times 16$  and  $32 \times 32$  are successively investigated) that is randomly chosen from the processed video corpus. Be  $M$  the maximum value of this block. Consider a distortion matrix of the same size as the considered block, sampled from a Gaussian generator, whose mean value and standard deviation are set to 0 and  $\alpha M/3$ , respectively. In other words, with a 99.7% probability, the modifications will be smaller than  $\alpha M$ . In illustrations in Fig. 5,  $\alpha = 0.1$ . Assume such a distortion is directly added to the considered transformed block, see the left column in Fig. 5: even a small alteration (smaller than a tenth of the maximal value in that block) of a unique transformed block results in an avalanche of modifications in the decoded frame. As expected, the larger the block size, the larger the drift effect. Assume now the case in which the alteration is left and right multiplied by the drift cancellation matrix (18) prior to its addition to the considered block. The results are presented in the right column of Fig. 5: this time, the modifications are kept inside the considered block and the drift effect can no longer be perceived. The only exception occurs in the case of  $32 \times 32$  blocks, where a small residual drift appears.

Note that the usefulness of the drift-avoidance method advanced with this paper is even higher when multiple blocks (of possible different sizes) are affected inside a same  $I$  frame, as illustrated in Fig. 6: by its very principle, the drift-masking matrix (18) ensures the reparability of the drift effects in case of multiple block errors. While Fig. 6 considers the same type of representation (the *parula* colormap) and the same type of random modification (sampled from a Gaussian generator, whose mean value and standard deviation are set to

0 and  $M/30$ ) as in Fig. 5, it illustrates the case in which a 4x4 and a 32x32 blocks are modified inside a same  $I$  frame. In the case in which the drift is not avoided, the effects of the two individual effects can be even amplified (see Fig. 6 left). However, when applying the drift avoidance method, the modification can be localized and no mutual drift amplification is encountered (see Fig. 6 right).

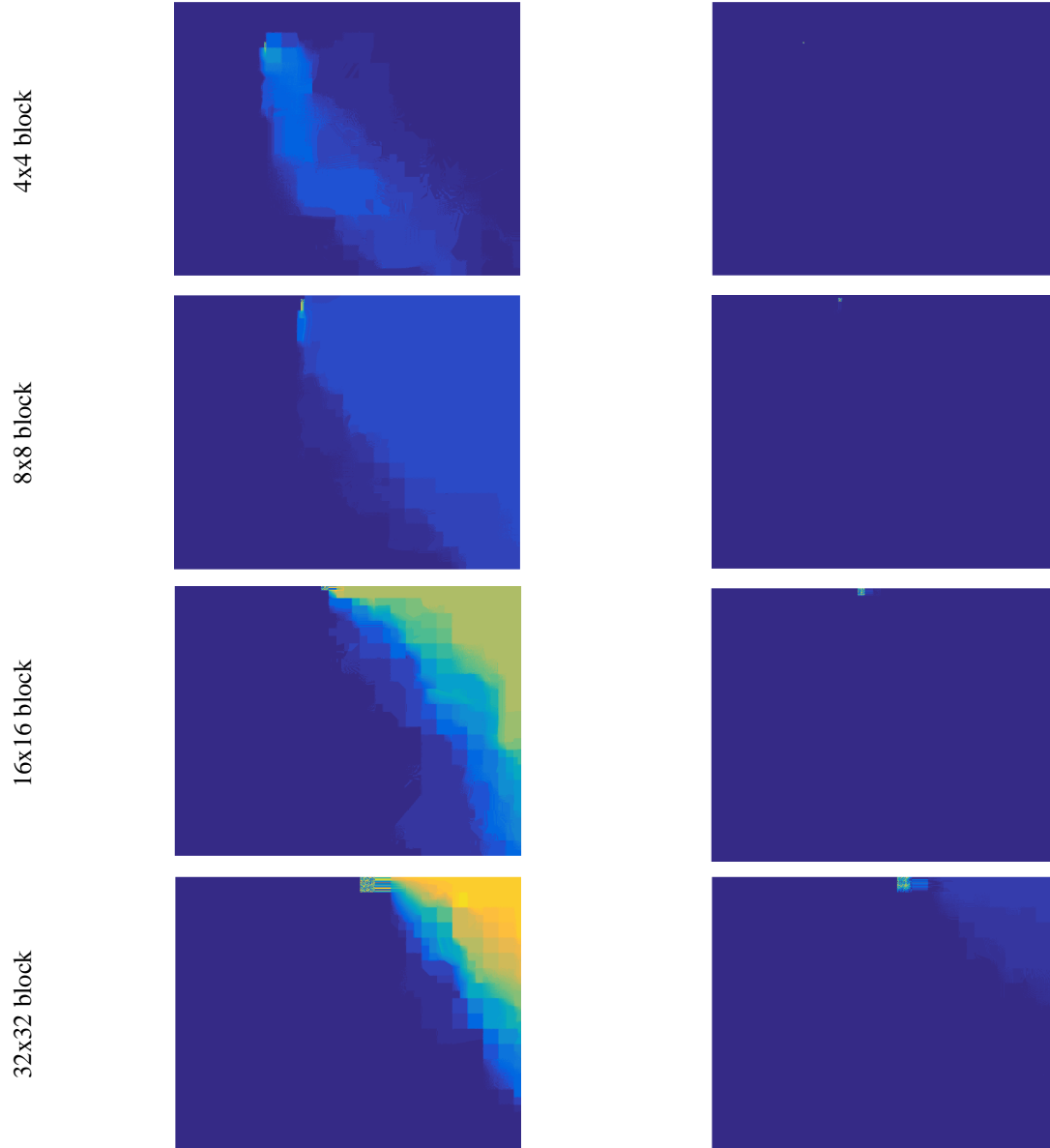


Fig. 5 Illustration of drift (left column) and drift-free (right column) modification of one transformed block in an  $I$  frame; the *parula* colormap is considered, with dark blue the lowest and light yellow the largest value



Fig. 6 Illustration of drift (left column) and drift-free (right column) modification of two transformed block in an  $I$  frame; the *parula* colormap is considered, with dark blue the lowest and light yellow the largest value

#### 4.2. Method performances

The evaluation targets two different criteria. First, it focuses on the residual drift effect and assesses the maximal value in the most-right column and the bottom row elements in the modified block (see the discussion in Section 3.C and the illustrations of the 32x32 blocks in Fig. 5). Secondly, it evaluates the overall visual quality effect of the advanced drift-free method; in this respect, the three quality metrics considered in [10] are considered, namely the PSNR, the SSIM and the VIFp.

##### 4.2.1 Assessing the maximal value in the most-right column and the bottom row of the modified block

As discussed in Section 3.3, the only approximation when deriving the noise cancellation matrix (18) comes from the very HEVC transform computation, see equation (5). Such an approximation can result (in some cases) in a residual drift effect, as illustrated in Fig. 5 for the 32x32 block. Hence, the first experiment is devoted to assessing the maximal value present in the most-right column and the bottom row in a processed block: in the absence of the residual drift, such a value is 0.

In this respect, for each  $I$  frame, the cases of 4x4, 8x8, 16x16 and 32x32 blocks are successively investigated. The distortion matrix has the same size as the processed block and is sampled from a Gaussian generator, whose mean value and standard deviation are set to 0 and  $\alpha M/3$ , respectively. In the experiments, 10 different  $\alpha$  values, evenly distributed between 0.1 and 1, are successively considered.

The maximal values in the most-right column and the bottom row of the modified block are plotted in Fig. 7, as a function of  $\alpha$ . In Fig. 7, the values corresponding to the un-masked (drift) case are labeled by “U” and plotted in circles • while the values corresponding to the advanced method (drift-free) are labeled by “M” and plotted in up-triangles ▲. The color of the plots differs according to the size of the block: blue for 4x4, orange for 8x8, green for 16x16 and purple for 32x32.

Fig. 7 shows that the residual drift results in modifications that depend, as expected, on both the  $\alpha$  value and, for a fixed  $\alpha$  value, on the size of the block.

In order to assess the effectiveness of the advanced method in the reduction of the modifications, the ratio of the initial (unmasked) values to the corresponding drift-masked values is computed and represented in Fig. 8 as a function of  $\alpha$  (the block-size is considered as parameter). It is thus brought to light that for 4x4 blocks, the dynamics is reduced 20 to 80 times (depending on the  $\alpha$  value). For the other sizes of the blocks, the reduction ratio ranges between 5 and 12. It should also be noticed that the reduction is more effective in the most disturbing cases, namely for large  $\alpha$  values.

Note that the values in Fig. 7 are obtained by averaging the values obtained for each  $I$  frame. The experiments show that all the 95% confidence intervals (regardless the  $\alpha$  value or the size of the block) are smaller than 1 (i.e., the 95% absolute error is always smaller than 0.5), thus ensuring the statistical relevance of the values reported in Fig. 7.

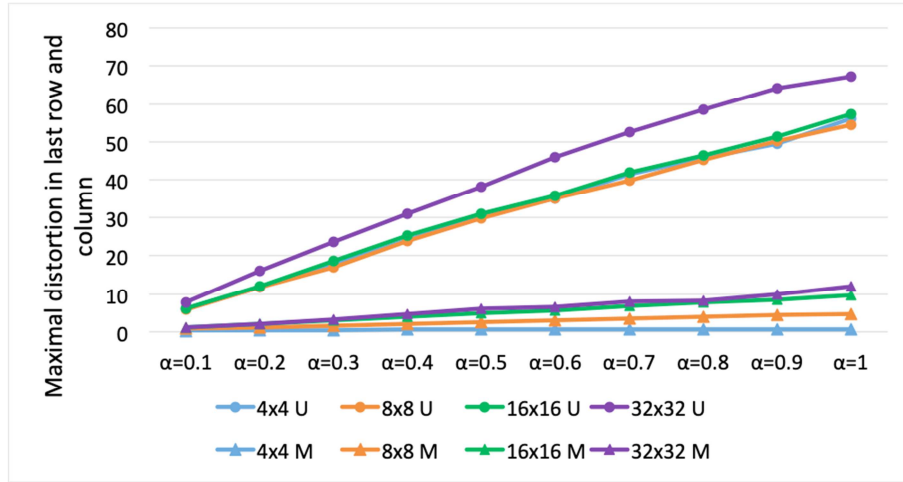


Fig. 7 Maximal distortion values in the most-right column and the bottom row of the modified block; the plot color differs with the size of the block: blue for 4x4, orange for 8x8, green for 16x16 and purple for 32x32

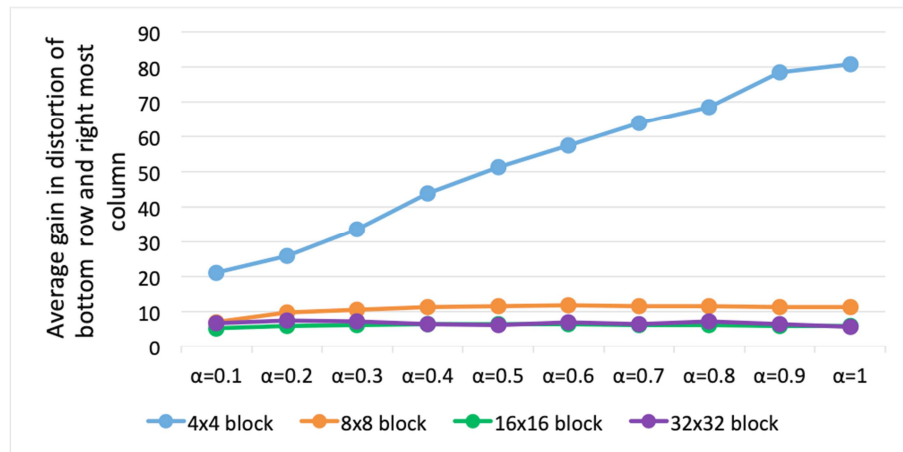


Fig. 8 The reduction ratio of the maximal values in the most-right column and the bottom row of the modified block; the plot color differs with the size of the block: blue for 4x4, orange for 8x8, green for 16x16 and purple for 32x32

#### 4.2.2 Assessing the visual impact of the advanced method

In addition to the residual drift, the advanced method also modifies the power of the additive modification (by left/right multiplication). In the sequel, the combined effect of these two impairments in

visual quality is evaluated.

In order to assess the visual impact of the advanced method, the three quality metrics considered in the state of the art study [10], namely the PSNR, the SSIM and the VIFp are computed. In this respect, for any of the three metrics, the unmasked and masked drift affected HEVC sequences are compared against the reference HEVC encoded sequence. The results are presented in Figs. 9, 10 and 11.

The experimental conditions are kept the same as in Fig. 7:

- the same 8 video sequences are processed;
- the cases of 4x4, 8x8, 16x16 and 32x32 blocks are successively investigated;
- the distortions are sampled from a Gaussian generator of 0 mean and  $\alpha M/3$  standard deviation;
- 10 different  $\alpha$  values, evenly distributed between 0.1 and 1, are considered.

The plots in Figs 9, 10 and 11 are organized the same way as Fig. 7:

- the values plotted in  $\bullet$  and labeled by “U” represent the drift, un-masked case while the values in up-triangles  $\blacktriangle$  and labeled by “M” represent to the advanced method;
- the same color mapping is considered blue for 4x4, orange for 8x8, green for 16x16 and purple for 32x32.

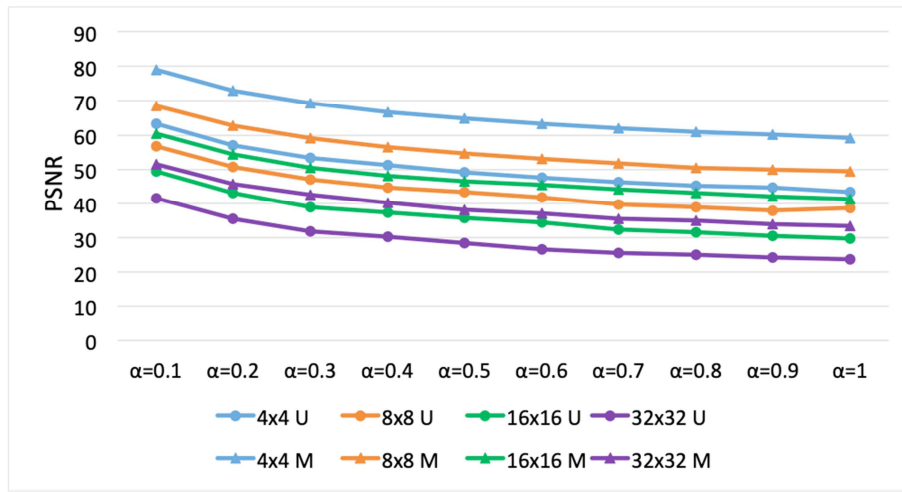


Fig. 9 Assessing the visual impact of the advanced method: PSNR values; the plot color differs with the size of the block: blue for 4x4, orange for 8x8, green for 16x16 and purple for 32x32

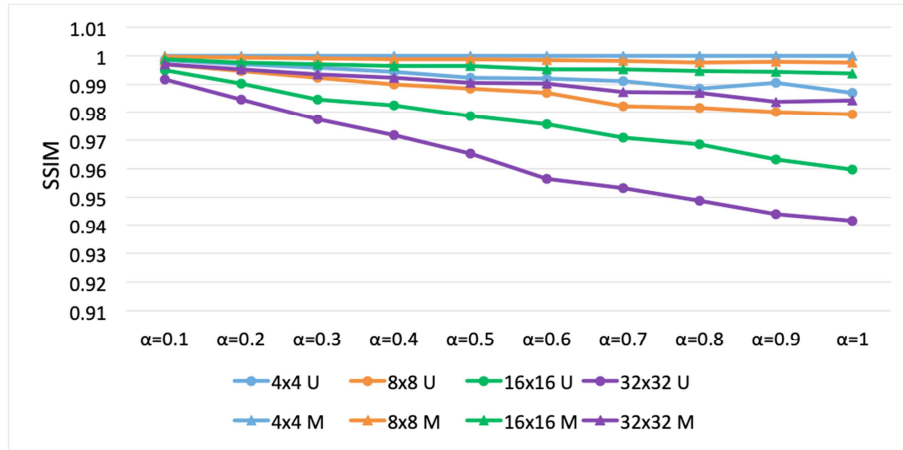


Fig. 10 Assessing the visual impact of the advanced method: SSIM values; the plot color differs with the size of the block: blue for 4x4, orange for 8x8, green for 16x16 and purple for 32x32

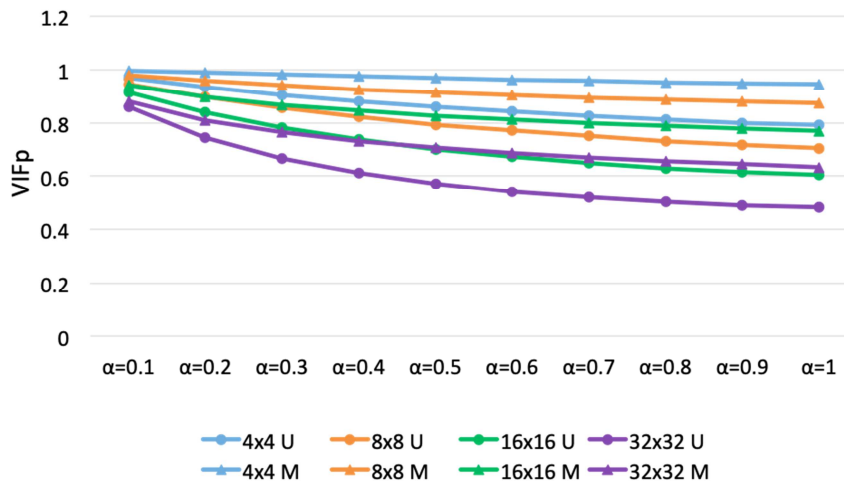


Fig. 11 Assessing the visual impact of the advanced method: VIFp values; the plot color differs with the size of the block: blue for 4x4, orange for 8x8, green for 16x16 and purple for 32x32

By investigating the Figs. 9, 10 and 11 it can be noticed that, as a general trend, the image quality is a decreasing function of  $\alpha$  and, for fixed values of  $\alpha$ , of the block size. Yet, the advanced method ensures lower limits (corresponding to  $\alpha = 1$  and 32x32 blocks) for PSNR, SSIM and VIFp of 33.5 dB, 0.985 and 0.985, respectively.

When compared to the initial (unmasked drift) case, Figs. 9, 10 and 11 bring to light that the advanced method ensures:

- gains in PSNR of 16 dB, 12 dB, 11 dB and 10 dB for the 4x4, 8x8, 16x16 and 32x32 blocks, respectively; note that these gains are quite constant with  $\alpha$  (variations lower than 10% being encountered);



- gains in the SSIM (expressed by the closeness to the ideal 1 value) that vary with both  $\alpha$  and the size of the block, from 0.005 (corresponding to 4x4 blocks and  $\alpha = 0.1$ ) to 0.05 (corresponding to 32x32 blocks and  $\alpha = 1$ );
- gains in VIFp (also expressed by the closeness to the ideal 1 value) having the same trend as the SSIM and varying in the same limits.

All the values reported in Figs. 9, 10 and 11 are obtained as average over the 8 video sequences. For each of the 10 studied  $\alpha$  values and for each of the 4 block sizes defined by HEVC, the length of the 95% confidence interval for PSNR is always lower than 1 while the lengths of the 95% confidence intervals corresponding to SSIM and VIFp are always lower than 0.01. Hence, as the 95% absolute errors are lower than 0.5 and 0.005, it can be stated that the values reported in Figs. 9, 10 and 11 are statistical relevant.

We emphasize that the image quality evaluation presented in this Section 4.2.2 considers the differences between unmasked/masked drift affected sequences and the initial, HEVC encoded sequence. The quite unusual high PSNR values can be explained by the fact that the experimental conditions consider quite sparse modifications (a single block per  $I$  frame).

#### 4.2.4 Benchmarking the advanced method against the state of the art method in [10]

As mentioned in the state of the art section, to the best of our knowledge, the most effective HEVC drift avoidance state of the art method is advanced by S. Gaj *et. al.* in [10]: therefore, the drift-free method advanced in the present paper is compared to it.

The main difference appears at the methodological level. The Gaj method belongs to the selection type: it considers only the 4x4 blocks and only the case in which 2 out of the 9 top-most coefficients are affected by a bipolar  $\delta$  and  $-\delta$  alteration, where  $\delta$  is a small integer. Although not addressing the general drift problem, such an approach can meet its objective, namely solving the drift issue for watermarking. Note that under the watermarking framework, the modification dynamic is expected to be small (so as to ensure transparency) and the modified positions in a block to be sparse (the compressed domain watermarking generally feature low data payload [10], [14]).

The method advance in the present paper belongs to the cancellation category and solves the problem in its generality: any block size, any number of modified coefficients, any type of modifications (not only bipolar) and any dynamic (no restriction to “small integers”) can be dealt with.

For the experimental comparison, the Gaj method was also implemented in C/C++ (in addition to the HEVC reference software) and the following set-up is considered for benchmarking:

- as restricted by the Gaj method, only 4x4 blocks are processed;
- the Gaj method is applied for a bipolar, 2 values matrix,  $\delta / -\delta$ ; in the experiments, 10 values of  $\delta$ , ranging from 1 to 10, are successively considered;
- the drift-free method is applied to an arbitrarily structure matrix, with the same average power as the Gaj matrix; actually, in order to grant more generality to the results, for each  $\delta$  value, 50 random distortion matrices are generated and the values reported in this section for the drift-free method are obtained by averaging the values obtained on that 50 matrices.

The experiments follow the same criteria as in Section 4.2, namely assessing the maximal value in the most-right column and the bottom row of the modified block (see Fig. 12) and assessing the visual impact of the residual intra-frame drift (see Fig. 13, 14 and 15). Note that this time the abscissa correspond to the  $\delta$  value while the ordinates are the same as in Figs. 7, 9, 10 and 11, respectively.

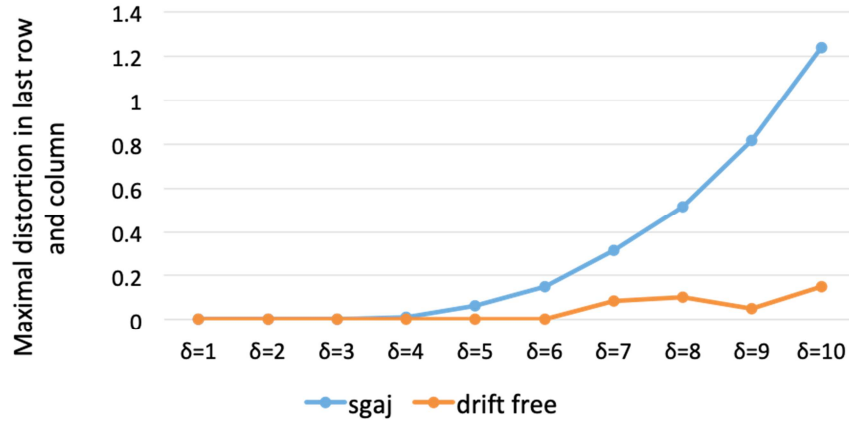


Fig. 12 Benchmarking the drift-free method against the S. Gaj method [10]: maximal values in the most-right column and the bottom row of the modified block; the plot in blue corresponds to the state of the art solution while the plot in orange to the method advanced in the paper

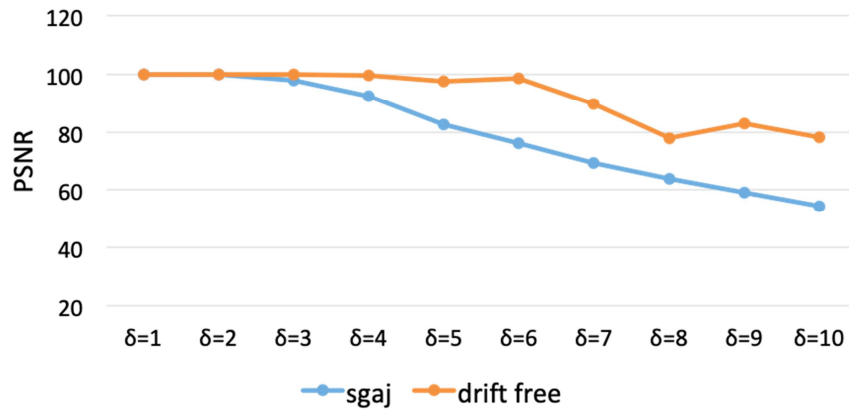


Fig. 13 Benchmarking the drift-free method against the S. Gaj method [10]: the visual impact expressed by PSNR values; the plot in blue corresponds to the state of the art solution while the plot in orange to the method advanced in the paper

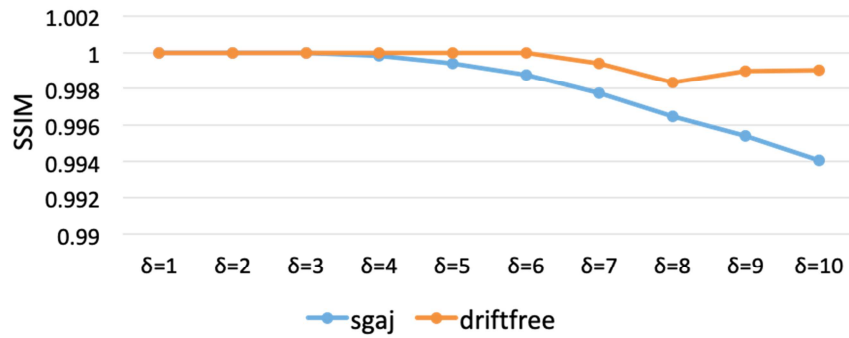


Fig. 14 Benchmarking the drift-free method against the S. Gaj method [10]: the visual impact expressed by SSIM values; the plot in blue corresponds to the state of the art solution while the plot in orange to the method advanced in the paper

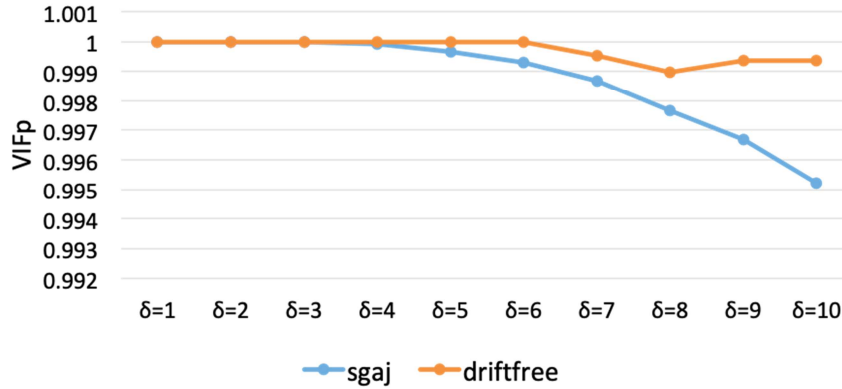


Fig. 15 Benchmarking the drift-free method against the S. Gaj method [10]: the visual impact expressed by VIFp values; the plot in blue corresponds to the state of the art solution while the plot in orange to the method advanced in the paper

Fig. 12 considers the maximal modification in the most-right column and the bottom row. It should be first noticed that the Gaj constraint of choosing “small  $\delta$  values” is very restrictive in practice and considers only very small modifications (very small power for the modifications generating the drift): the ordinates varies from 1 to 70 in Fig. 7 while it is restricted to 0 to 1.23 in Fig. 12. For  $\delta$  values of 1, 2 3 and 4, the Gaj method and the advanced drift-free method provide equally good performances. Yet, for  $\delta$  values larger than 5, the drift-free method outperforms the state of the art method, by factors ranging between 10 and 30 (these factors are computed as the ratio of the maximal values to the two methods). The 95% error corresponding to the results reported in Fig. 12 is lower than 0.05.

Fig. 13, 14 and 15 bring to light a quite similar behavior as in Fig. 12: while the two methods have similar performances for the smallest  $\delta$  values, the advanced drift free method outperforms the state of the art method after a certain  $\delta$  limit. For instance, when considering the PSNR as criterion in image quality evaluation, Fig. 13 shows that the two methods provide similar performances for  $\delta = 1$  and  $\delta = 2$ . While the PSNR value difference is only of 2dB for  $\delta = 3$ , it increases from 5 dB to 24dB when  $\delta$  varies from 4 to 10. Note that the 95% error corresponding to the results is lower than 0.5 in Fig. 13 and lower than 0.001 in Fig. 14 and 15.

As a final quantitative result, the variation in the bitstream size is also investigated. In this respect, the relative variation with respect to the original (unmodified) bitstream is computed for both the advanced drift free method and for the state of the art method in [10], as illustrated in Fig. 16. Fig. 16 shows that the advanced method does not practically impact the stream size: the relative differences between the original and the drift compensate HEVC streams are lower than 0.4%. For comparison, the same relative differences reach 0.18% for the investigated the state of the art method [10].

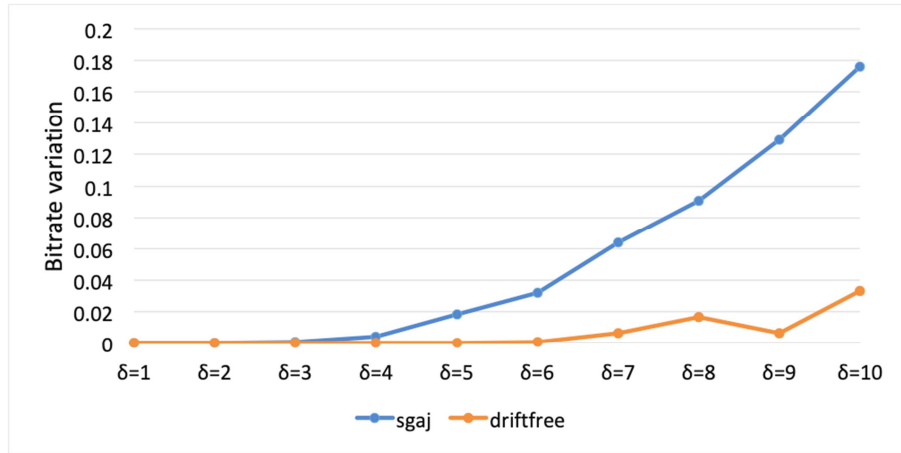


Fig. 16 Benchmarking the drift-free method against the S. Gaj method [10]: the relative bitstream size variation; the plot in blue corresponds to the state of the art solution [10] while the plot in orange to the method advanced in the paper

## 5. Conclusion and perspectives

The present study considers the HEVC compression standard and investigates the possibility of cancelling the intra drift effect induced in the  $I$  frames by additive alterations. Starting from the HEVC formulas describing the intra-prediction modes and the DCT computation, a drift cancellation (masking) matrix is derived: by left-right multiplying the additive alteration prior to its addition, the drift effect is theoretically cancelled. This theoretical contribution is complemented with experimental illustrations and results related to the gain obtained with respect to the basic, no-drift elimination case.

With respect to MPEG-4 AVC drift cancellation methods [8], the main novelty consists in solving a related yet different theoretical problem: as HEVC has a more elaborated hierarchical structure and a completely different definition for its syntax elements (different block sizes, extended prediction modes, changes the definition of the transforms, ...), a new demonstration is required so as to derive the mask multiplication matrix (18).

With respect to the state of the art HEVC drift avoidance methods [9], [10], the main contribution consists in solving the intra-drift problem in its general form: the advanced drift elimination matrix holds for any block-size, any prediction mode and any structure / dynamic of the modification. It is also experimentally shown that when restricting the advanced method to the applicative perimeter of the state of the art method in [10], equally good or better results are obtained.

The experimental results reported in Section 4 are obtained by processing a video corpus whose size is set so as to ensure their statistical relevance (in the sense of their 95% errors). Yet, in order to also check the stability of the results with respect to the video content itself, a second corpus (of the same size and with the same encoding principles) is composed by sequences sampled from [15]. The comparison between the drift-free cancelation method and the S. Gaj [10] state of the art method is resumed on the new corpus. No statistical relevant differences are encountered between the results reported in Section 4 and the new obtained results when considering (1) maximal values in the most-right column and the bottom row of the modified block – see Fig. 12, (2) the SSIM metric – see Fig. 14 and (3) the VIFp metric – see Fig. 15. Although the PSNR results obtained on the new corpus are statistically different with respect to the ones reported in Fig. 13,

the same trend is kept: the state of the art method [10] and the new advanced method features equally good results for  $\delta = 1$  and  $\delta = 2$ , while the new advanced method outperforms the state of the art method by gains between 3 and 20 dB for larger  $\delta$  values.

As a final quantitative result, it can be stated that the advance drift cancelation method does not impact the stream size, Fig. 16: the relative differences between the original and the drift compensate HEVC streams are lower than 0.4% while the same relative difference can reach 0.18% for the investigated the state of the art method [10].

Note that the study we carried out takes place in the compressed domain and is independent with respect to any pre- and post-processing operations carried out before, during or after HEVC encoding/decoding. Consequently, our results are independent and can be used in conjunction with respect to the recent studies [16], [17] carried out in pixel (uncompressed domain) or at an intermediate encoding step and demonstrating the possibility of using neural networks for increasing the overall HEVC compression efficiency. Specifically, while our study allows intra-drift elimination, the studies in [16], [17] rather consider generic visual artifacts the lossy encoding standards are likely to induce (mainly at low rates) like blocking, blurring, ringing, ...

Future work will be done in order to integrate the drift cancelation matrix (18) under the watermarking applicative framework and to complete this study with an investigation of the inter-drift effect.

Watermarking [18] aims at persistently and perceptually associating some extra information (a watermark) with some original content. By subsequently recovering the watermark from modified (attacked) replicas of the watermark content, several applications can be ensured, like propriety right identification, piracy tracking down or video authentication, to mention but a few. While the insertion procedure itself can be of various types (spread spectrum, side information, *etc.*) it was shown that it can be explicitly or implicitly modeled by the addition of a white, Gaussian noise [18]. Inserting the mark directly into a compressed content (*e.g.* MPEG-4 Part 2, MPEG-4 AVC, HEVC) ensures faster and more flexible solutions. Yet, as a compressed stream is a priori expected to feature no visual redundancy, and as its content observes to strict syntax constraints, any modification of its elements is expected to result in visual artifacts, thus turning the watermarking imperceptibility into a hot research topic. Under this framework, the present study provides a solution for ensuring drift free watermarking: the drift simulations presented in this paper correspond to the additive, white Gaussian noise addition modeling the mark insertion while the 10 values of  $\alpha$  cover the targeted watermark dynamics. Moreover, the drift masking matrix (18) is known both at the mark insertion and detection, and the study in [14] demonstrated that inserting a watermark following a *m*QIM (multiple Quantization Index Modulation) method can ensure robustness against matrix multiplication operations; hence, although the inserted value is modified by the drift avoidance mask, it can still be detected. Consequently, we target an end-to-end HEVC watermarking study, in which the drift avoidance mask matrix (18) will be considered as a transparency optimization tool.

A second direction of our future work will be to study the possibility of avoiding the inter drift effect. In this respect, the joint intra- and inter-drift effect should be modeled as a multidimensional matrix problem to be solved under 0-borders constraints. Of course, the complexity of the problem is much higher: the intra-prediction mode number is to be multiplied by the inter-prediction mode numbers, the various possible sizes for the blocks in *I*, *P* and *B* frames should also be considered, *etc.*

Beyond these two future work directions, our results can also be considered as a ground for studying the intra-drift problem for compression standards related to HEVC, as the multiview and 3D HEVC extensions [19], [20]. At a glance, such an encoder considers information describing the multi-views, the associated depth maps and the camera parameters. The encoding structure is composed of a base (independent) view, which is

coded independently of other views using a conventional HEVC encoder and of several dependent views, which can be coded with additional coding tools [20]. Hence, while our study covers the intra-drift effect related to the base view, extensions are required so as to deal with its effects on the depend views. Note that the difficulty of the extension is twofolded. First, at the theoretical level, the study should start from equation (14) and reconsider its instantiations (15)–(18). On the other hand, in order to ensure coding efficiency, the statistical dependencies between video texture and depth are exploited in the multiview and 3D HEVC [19]: hence, an in-depth quality evaluation should be conducted for the reconstructed multi-view content.

## References

- [1] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, “Overview of the high efficiency video coding (HEVC) standard,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, 2012.
- [2] X. Ma, Z. Li, J. Lv, and W. Wang, “Data hiding in H.264/AVC streams with limited intra-frame distortion drift,” *Proc. - 1st Int. Symp. Comput. Netw. Multimed. Technol. CNMT 2009*, pp. 0–4, 2009.
- [3] W. Huo, Y. Zhu, and H. Chen, “A controllable error-drift elimination scheme for watermarking algorithm in H.264/AVC stream,” *IEEE Signal Process. Lett.*, vol. 18, no. 9, pp. 535–538, 2011.
- [4] X. Gong and H. M. Lu, “Towards fast and robust watermarking scheme for H. 264 video,” in *10th IEEE International Symposium on Multimedia*, 2008, pp. 649–653.
- [5] L. Zhang, Y. Zhu, and L. M. Po, “A novel watermarking scheme with compensation in bit-stream domain for H.264/AVC,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2010, pp. 1758–1761.
- [6] X. Ma, Z. Li, H. Tu, and B. Zhang, “A data hiding algorithm for h.264/AVC video streams without intra-frame distortion drift,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 10, pp. 1320–1330, 2010.
- [7] W. Chen, Z. Shahid, T. Stütz, F. Autrusseau, and P. Le Callet, “Robust drift-free bit-rate preserving H.264 watermarking,” *Multimed. Syst.*, vol. 20, no. 2, pp. 179–193, 2014.
- [8] M. Hasnaoui and M. Mitrea, “Drift-free MPEG-4 AVC semi-fragile watermarking,” in *Proc. SPIE*, vol. 9028, p. 90280F.
- [9] P.-C. Chang, K.-L. Chung, J.-J. Chen, C.-H. Lin, T.-J. Lin, “A DCT/DST-based error propagation-free data hiding algorithm for HEVC intra-coded frames,” *J. Vis. Commun. Image R.*, vol. 25 (2014) 239–253.
- [10] S. Gaj, A. Kanetkar, A. Sur, P.K. Bora, “Drift-compensated robust watermarking algorithm for H.265/HEVC video stream,” *ACM Trans. Multimedia Comput. Comm Appl.*, vol. 13 (2017), no. 1, Article 11.
- [11] J. Lainema, F. Bossen, W. J. Han, J. Min, and K. Ugur, “Intra coding of the HEVC standard,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1792–1801, 2012.
- [12] M. Budagavi, A. Fuldseth, G. Bjontegaard, V. Sze, and M. Sadafale, “Core transform design in the high efficiency video coding (HEVC) Standard,” *IEEE J. Sel. Top. Signal Process.*, vol. 7, no. 6, pp. 1029–1041, 2013.
- [13] <https://hevc.hhi.fraunhofer.de/trac/hevc/browser>; last time visited on December 21, 2018.

- [14] M. Hasnaoui, M. Mitrea, "Multi-symbol QIM video watermarking," *Signal Processing: Image Communication*, vol. 29, Issue 1, pp. 107-127, 2014.
- [15] [http://medialab.sjtu.edu.cn/resources/resources\\_subdataset.html](http://medialab.sjtu.edu.cn/resources/resources_subdataset.html); last visited on December 21, 2018.
- [16] X. He, Q. Hu, X. Han, X. Zhang, W. Lin, "Enhancing HEVC compressed videos with a partition-masked convolutional neural network," *Intl. Conf. Image Processing (ICIP)*, 2018.
- [17] Y. Dai, D. Liu, F. Wu: A convolutional neural network approach for post-processing in HEVC intra coding. *MMM 2017*: 28-39
- [18] I. Cox, M. Miller, J. Bloom, *Digital Watermarking*, Morgan Kaufmann Publishers, 2002.
- [19] Tech G., Chen Y., Müller K., Ohm J.-R., Vetro A., Wang Y.-K., "Overview of the multiview and 3d extensions of high efficiency video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 35-49, 2016.
- [20] Shen L., Li K., Feng G., An P., Liu Z., "Efficient intra mode selection for depth-map coding utilizing spatiotemporal, intercomponent and inter-view correlations in 3d-hevc," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4195-4206, 2018.