

# Error Resilience Aware Motion Estimation and Mode Selection for HEVC Video Transmission

Gosala Kulupana, Dumidu S. Talagala, Hemantha Kodikara Arachchi, and Anil Fernando  
 Centre for Vision, Speech and Signal Processing, University of Surrey, United Kingdom  
 Email: {g.kulupana, d.talagala, h.kodikaraarachchi}@surrey.ac.uk, w.a.c.fernando@gmail.com

**Abstract**—Error concealment techniques such as motion copying require significant changes to HEVC (High Efficiency Video Coding) motion estimation process when incorporated in error resilience frameworks. This paper demonstrates a novel motion estimation mechanism incorporating the concealment impact from future coding frames to achieve an average 0.73 dB gain over the state-of-the-art.

## I. INTRODUCTION

Multimedia consumption over mobile terminals has experienced an unprecedented growth during the past decade. However, resource provisioning amongst mobile consumers, in order to cope with the increasing traffic demands, poses many challenges, of which, bandwidth scarcity and network packet errors are the most significant. In the process of overcoming these challenges, video coding is expected to mitigate the bandwidth demand through compression, whereas video error resiliency is expected to combat the packet errors during the transmission. However, the emerging coding standards such as HEVC have mostly disregarded the transmission aspect of the video data, and as a result the consumers' viewing pleasure will be significantly degraded when a coded video is transmitted over a lossy medium such as a wireless channel. This is due to the extensive inter-frame prediction processes that have been introduced to improve its compression efficiency. An effective error resilience framework for HEVC is therefore of utmost importance.

Even though the video error resiliency has been extensively investigated in the literature, few can be immediately adopted for HEVC. Zhihai *et al.* [1] modeled the end-to-end distortion of a H.263 video as a combination of source distortion and channel distortion, which was later extended to incorporate the error propagation impact to motion vector prediction [2] in H.264. With respect to HEVC error resilience, a temporal domain error propagation analysis and mitigation approach was presented in [3]. A similar approach which effectively switches the TMVP (Temporal Motion Vector Prediction) feature in HEVC has also been presented [4]. However this only investigates the effect of error propagation due to TMVP candidates, and the pixel domain temporal correlation and error propagation have not been considered.

Furthermore, both [1], [2] have modeled the total distortion using a slice copy error concealment technique. Thus, they cannot be applied with a more effective concealment technique such as motion copying, without assessing the impact of the estimated motion vectors to the pixels in the future frames.

This work was supported by the EU FP7 project ACTION-TV, funded under the EC's 7<sup>th</sup> Framework Program (Grant Number: 611761).

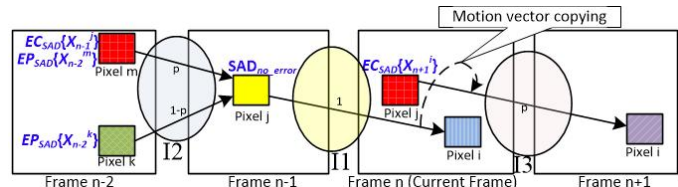


Fig. 1 Proposed motion vector estimation criteria

This paper therefore proposes a novel mechanism to estimate motion vectors that enhance the resiliency of HEVC encoding, considering the error concealment impact from the pixel domain, in both past and future frames.

## II. PROPOSED METHOD

### A. Error Resilient Motion Estimation

The purpose of the standard HEVC encoder is to find the motion vectors which minimize  $J_{motion}$  in (1).

$$J_{motion} = SAD_{no\_error} + \lambda_{motion} \cdot R_{mv} \quad (1)$$

where  $SAD_{no\_error}(n, i) = E\{|X_n^i - \hat{X}_{n-1}^j|\}$ . (The notations that are used during this work are summarized in Table I.)

However, this model cannot be applied in an erroneous channel as this has not evaluated the impact from error propagation. In the proposed motion vector selection criteria, the impact from following three aspects will be evaluated.

1. Quantization distortion in the current block.
2. Error propagation impact from the previous blocks.
3. Error concealment impact for the future coding blocks.

Fig.1 illustrates the motion vector selection criteria over an error prone channel assuming a motion copy based error concealment. Here,  $EC_{SAD}\{X_{n-1}^j\} = |X_{n-1}^j - \hat{X}_{n-2}^m|$  corresponds to the absolute difference of the concealment error, whereas  $EP_{SAD}\{X_{n-1}^j\} = p \cdot EC_{SAD}\{X_{n-1}^j\} + p \cdot EP_{SAD}\{X_{n-2}^m\} + (1-p) \cdot EP_{SAD}\{X_{n-2}^k\}$  is the expected value of the propagated error of the  $j^{\text{th}}$  pixel in the  $(n-1)^{\text{th}}$  frame.  $p$  is the packet error rate (PER). The arrows in the Fig.1 signify the probabilities of the associated error terms (denoted in blue colour) and the three circles correspond to the three impacts described earlier in this section.

TABLE I  
TERMS AND NOTATIONS

$X_n^i$ -original pixel $i$ of frame $n$ .	$j$ - location of the reference pixel of the pixel $i$
$\hat{X}_n^i$ -reconstructed pixel $i$ of frame $n$ .	$k$ - location of the reference pixel of the pixel $j$
$\tilde{X}_n^i$ -decoded pixel $i$ of frame $n$ .	$m$ - location of the concealment pixel of the pixel $j$
$n$ - current frame	$h$ - location of the concealment pixel of the pixel $i$
$i$ - location of the currently considered pixel	

The derivation of motion vectors over an erroneous channel (obtained from Fig. 1) leads to the minimization of (2).

$$J'_{motion} = SAD_{error} + \lambda_{motion} R_{mv}, \quad (2)$$

where  $SAD_{error}(n,i) = E\{|X_n^i - \tilde{X}_{n-1}^j|\} + p.E\{|X_{n+1}^i - \tilde{X}_n^j|\}$ . Mathematically (2) can be expressed as,

$$SAD_{error}(n,i) = SAD_{no\_error} + EP_{SAD}\{X_{n-1}^j\} + p[EC_{SAD}\{X_{n+1}^i\} + p.EC_{SAD}\{X_n^j\} + p.EP_{SAD}\{X_{n-1}^k\}] \quad (3)$$

After eliminating the higher order terms of  $p$ , (3) can be approximated as (4), to be used in the remainder of this work.

$$SAD_{error}(n,i) = SAD_{no\_error}(n,i) + EP_{SAD}\{X_{n-1}^j\} + p.EC_{SAD}\{X_{n+1}^i\} \quad (4)$$

### B. End-to-end Distortion Modeling

The standard HEVC encoder performs Rate Distortion Optimization (RDO) in order to derive the best coding modes. RDO entails deriving the coding modes which minimize

$$J_{mode} = SSE_{no\_error} + \lambda_{mode}(R_{mv} + R_{residual}) \quad (5)$$

Similar to (1), (5) also does not consider the impact from the error propagation during coding mode selection. Therefore the RDO function in erroneous situations can be expressed as,

$$J'_{mode} = SSE_{error} + \lambda_{mode}(R_{mv} + R_{residual}), \quad (6)$$

where  $SSE_{error}(n,i) = E\{(X_n^i - \tilde{X}_n^i)^2\} = (X_n^i - \hat{X}_n^i)^2 + p.E\{(\hat{X}_n^i - \tilde{X}_{n-1}^h)^2\} + (1-p).E\{(\hat{X}_{n-1}^j - \tilde{X}_{n-1}^j)^2\} = SSE_{no\_error} + EP_{SSE}\{X_n^i\}$

$EP_{SSE}\{X_n^i\} = p.EC_{SSE}\{X_n^i\} + p.EP_{SSE}\{X_{n-1}^h\} + (1-p).EP_{SSE}\{X_{n-1}^j\}$  is the error propagation distortion and  $EC_{SSE}\{X_n^i\} = (X_n^i - \hat{X}_n^i)^2$  is the concealment distortion. Furthermore, after considering the impact of the HEVC CU (coding unit) size, the above error model can be refined as,

$$EP_{SSE}\{X_n^i\} = k.(EP_{SSE}\{X_n^i\})^t, \quad (7)$$

where  $k = \log_2(\text{CU size})$  and  $t$  is an empirically evaluated 0.8. This can be intuitively explained; for example, if large CUs are encoded in the inter mode, the error propagation increases. Hence larger CUs are less likely to be coded in the inter mode.

## III. RESULTS AND DISCUSSION

The proposed algorithm is implemented in the HM 16.2 reference software. Each video sequence is encoded with slices having 6 CTUs (Coding Tree Units). All the simulations are performed for CIF videos at 25fps. Fig. 2 illustrates the PSNR gain obtained for the 'stefan' sequence for different PERs while keeping the bit rate at 600 kbps and 1200 kbps. Table II summarizes the obtained PSNR for different video sequences at 1200 kbps. The performance of the proposed method is compared with two other error resilience schemes (Zhihai [1] and Md+Me [2]) originally proposed for the H.263 and H.264 codecs, respectively. In addition, the performance is compared with respect to standard HEVC encoding assuming motion copy and slice copy error concealments.

The results illustrate that the proposed method has outperformed the state-of-the-art methods. By observing Fig. 2,

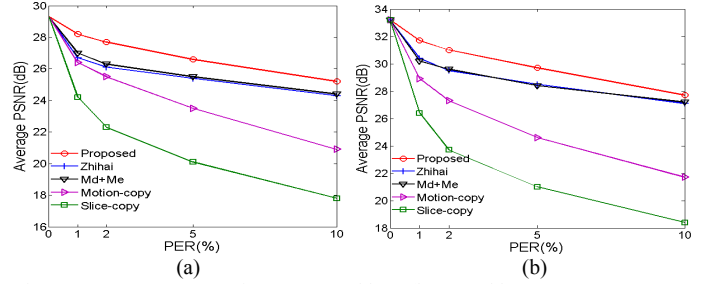


Fig. 2. PSNR vs PER- *Stefan* at (a) 600 kbps, (b) 1200 kbps

it can be deduced that the gain over the traditional methods becomes apparent in low bit rates. Moreover, the PSNR gain for high motion sequences (e.g., *stefan*) is more noticeable than low motion sequences (e.g., *foreman*). For HEVC, both the Zhihai *et al.* [1] and Md+Me [2] methods have demonstrated similar performance, whereas motion copy error concealment outperforms the slice copy concealment.

TABLE II  
PSNR VS PER FOR DIFFERENT VIDEO SEQUENCES

PER	Proposed	Zhihai [1]	Md+Me [2]	Motion-copy	Slice-copy					
						coastguard		stefan		
1	34.4	33.7	33.9	29.7	28.1	31.7	30.4	30.2	28.9	26.4
2	34.0	33.4	33.4	28.6	26.6	31.0	29.5	29.6	27.3	23.7
5	33.2	32.7	32.6	26.2	23.2	29.7	28.5	28.4	24.6	21.0
10	32.0	31.6	31.5	23.7	21.0	27.7	27.1	27.2	21.7	18.4
						foreman		bus		
1	37.6	36.8	36.9	26.8	28.6	31.6	31.1	30.8	30.2	26.2
2	37.0	36.2	36.4	24.2	26.6	31.0	30.3	30.1	28.4	24.7
5	35.7	35.1	35.3	20.8	23.0	29.4	28.9	28.8	24.7	20.7
10	34.1	33.7	33.9	18.7	20.4	27.8	27.3	27.3	23.0	18.3

## IV. CONCLUSION

In this paper, a novel framework for achieving error resilience in HEVC video coding by incorporating the impact of motion vector estimation in past and future coding frames is proposed. The overall performance gain is more significant at low bit rates for high motion sequences, with an average gain of 0.73 dB over the state-of-the-art techniques and 8.8 dB over the standard HEVC encoder. Future work will model the RDO to incorporate the impact from the future coding frames.

## V. REFERENCES

- [1] Z. He, J. Cai, and C. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 511–523, Jun. 2002
- [2] C. Lu and Y. Lin "Robust error resilient H.264/AVC video coding" in *IEEE TENCON*, Xi'an, China, pp. 1-4, Oct. 2013 .
- [3] B. Li, J. Xu, and H. Li, "Parsing robustness in high efficiency video coding-analysis and improvement," in *IEEE international Conf. on Visual Communications and Image Processing*, Tainan, Taiwan, pp. 1–4, Nov. 2011.
- [4] J. Carreira and V. Silva, "Dynamic motion vector refreshing for enhanced error resilience in HEVC," in *Proceedings of the 22nd EUSIPCO*, Lisbon, Portugal, pp. 281–285, Sept. 2014.