ESA360 - Early SKIP Mode Decision Algorithm for Fast ERP 360 Video Coding

Iago Storch, Guilherme Correa, Bruno Zatt, Luciano Agostini, Daniel Palomino

Video Technology Research Group (ViTech)
Graduate Program in Computing (PPGC), Federal University of Pelotas (UFPel), Brazil

{icstorch, gcorrea, zatt, agostini, dpalomino}@inf.ufpel.edu.br

***Abstract*—This paper presents ESA360 – Early SKIP mode decision Algorithm for fast ERP 360 video coding. ESA360 performs early SKIP mode selection based on block position and homogeneity to avoid testing the remaining modes and reduce the encoder complexity. The algorithm is developed based on an analysis of the specific behavior of SKIP mode occurrences throughout the video frame when coding equirectangular projected (ERP) 360 videos. This analysis pointed out that due to ERP distortion, the homogeneity of the blocks and the region being coded poses a more significant influence in the selection of SKIP mode for ERP 360 videos when compared to conventional videos. Experimental results show that ESA360 is able to achieve up to 21.44% of complexity reduction with negligible coding efficiency loss, and it is competitive with related works.**

Keywords—360 video, early decision, complexity reduction

I. Introduction

Recent technological advancements turned devices capable of recording and reproducing digital videos affordable to the mass population, and allied to the popularization of internet connection worldwide, digital videos became omnipresent in most people's daily lives. Recent reports point that by 2022, digital videos will be responsible for 82% of total internet traffic [1]. Besides, immersive approaches such as 360 videos are becoming very popular in the last years, and they bring new challenges related to coding, storage, and transmission. For instance, a 360 video needs to be projected into a rectangular shape to be coded by current video coding standards, usually the Equirectangular Projection (ERP) [2].

Although video coding algorithms enable the compression of digital videos, recent video coding standards present a high and ever-increasing computational complexity. Such high complexity is due to the vast number of coding possibilities available. When the High Efficiency Video Coding (HEVC) standard [3] is considered, each basic coding block (Coding Tree Unit – CTU) is partitioned in several Coding Units (CUs), and multiple interframe prediction techniques (integer and fractional motion estimation, special MERGE/SKIP modes) and intraframe prediction modes (35 modes) are applied for each CU aiming to achieve the best coding efficiency. The high complexity of evaluating all these coding modes is even higher when considering ERP 360 videos since they demand higher resolutions to be represented. With a fast decision algorithm able to anticipate the best coding mode, the exhaustive evaluation of all coding possibilities can be avoided, and the coding complexity can be reduced.

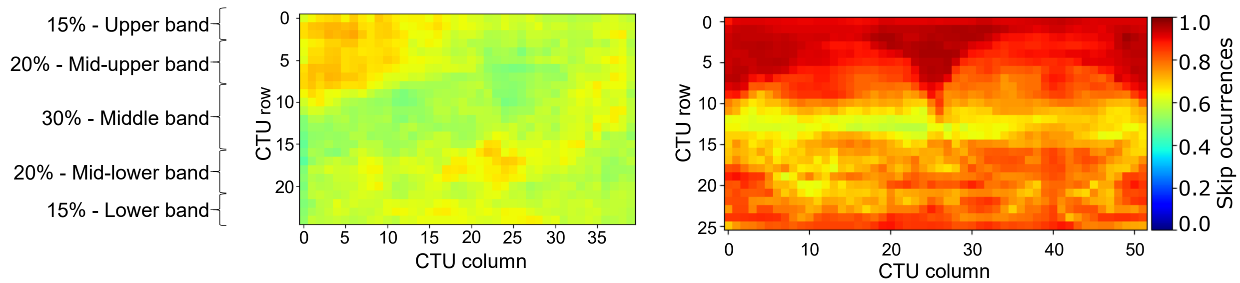
Even though the ERP projection enables 360 videos to be encoded by conventional video coding standards, it creates distortions in the video through horizontal stretching that are more intense in the top/bottom regions of the frame, such as depicted in Fig. 1. Since these texture characteristics are not present in conventional videos, they are not exploited by current video coding standards, such as HEVC. These texture characteristics of ERP 360 videos may lead current coding tools (such as the SKIP mode) to present distinct behavior in specific regions of the frame when compared to conventional videos. When well explored, this behavior can be used to reduce the coding complexity of ERP 360 videos

Several works propose complexity reduction techniques for the coding of 360 videos. In [4], the authors propose a complexity reduction technique based on early terminating the block partitioning. With this technique, an average of 32% of complexity reduction is achieved. However, the coding efficiency is significantly harmed, and it may render the technique inadequate in some coding scenarios. The authors of [5] propose a complexity reduction technique based on adjusting the precision of fractional motion estimation (FME) and limiting the minimum width of coding blocks. By employing such limitations, Ray *et al.* can achieve 15% of complexity reduction, on average. However, these techniques consist in simply limiting some coding tools linearly in pre-defined regions of the frame disregarding the content in the region, for instance, limiting the minimum width of *all blocks* in the top and bottom region of the frame. In [6], the authors propose a fast sample adaptive offset (SAO) algorithm that uses a fuzzy logic control to accelerate the SAO mode decision. Although it is able to achieve more than 70% of complexity reduction in the SAO step with negligible coding efficiency variations, in-loop filters such as SAO are responsible for a marginal fraction of the coding complexity [7]. *Most related works do not perform a consistent evaluation of the specific behavior of coding ERP 360 videos. Moreover, they rely on performing early termination of the block size decision. They do not exploit strategies that could improve the coding complexity reduction for each block size, such as early mode selection. Considering this, there is a need for fast mode decision algorithms that use the specific behavior of coding ERP 360 videos to improve complexity reduction.*

This paper proposes the ESA360: an **E**arly **S**KIP mode decision **A**lgorithm for fast ERP 360 video coding. The proposed algorithm is based on the specific behavior of the SKIP mode occurrence throughout the frame when coding ERP 360 videos with HEVC. An analysis of the behavior of the SKIP mode throughout the frame is performed, and the proposed algorithm uses the relation between SKIP occurrences and block homogeneity to accelerate the ERP 360 video coding by performing early SKIP decisions.

Fig. 1. Videos in the ERP projection.

The authors thank FAPERGS, CNPq, and CAPES (Finance Code 001) Brazilian research support agencies that financed this investigation.



(a) SKIP occurrences in conventional videos (b) SKIP occurrences in 360 videos

Fig. 2. Spatial evaluation of SKIP mode occurrences throughout the video frame.

II. SKIP MODE BEHAVIOR IN ERP 360 VIDEOS

In the HEVC coding process, each frame of the video is divided into CTUs, which is the basic coding block and usually has dimensions of 64×64 pixels. Each one of these CTUs is partitioned recursively into CUs, which have dimensions from 8×8 up to 64×64. Finally, in the case of interframe prediction, these CUs can comprise either a single square-shaped Prediction Unit (PU) or two rectangular PUs. Although block-based video coding presents many advantages, in some cases, multiple blocks share movement information, and the SKIP mode was developed to exploit this situation. The SKIP mode consists in allowing a block to inherit motion information from an already coded block. To achieve this, a list of motion candidates is generated based on previously encoded blocks, in a way that the motion information for the current block can be signaled by the index of the candidate from which the motion information is inherited. This is possible because the list of motion candidates can be generated identically in both encoder and decoder sides. In addition to that, the SKIP mode is only available for square-shaped PUs, and when it is employed, the prediction error is discarded. Considering these properties, the SKIP mode is mainly employed in the coding of homogeneous regions, which tend to present similar motion information throughout its blocks. Finally, if it were possible to predict when the SKIP mode is chosen to encode a given region, it would result in a drastic complexity reduction since the remaining of interframe prediction steps could be avoided.

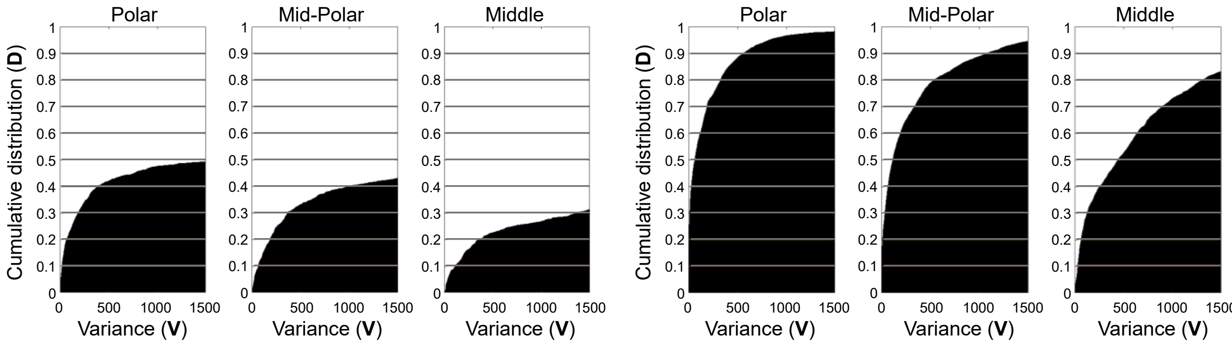
When using the ERP projection, each parallel of the sphere is mapped into a row of the frame. As the equator presents the largest radius among all parallels, its circumference corresponds precisely to the frame width. However, the remaining parallels present a smaller radius and must be stretched to fill the frame width. The closer to the north and south poles, the smaller is the radius of such parallels, and more severe is the stretching observed in the projected frame. This horizontal stretching creates highly homogeneous regions in the top and bottom regions of the frame that can influence the behavior of the SKIP mode occurrence throughout the frame coding.

Considering this, two evaluations are performed to understand how the distortions caused by ERP projection influence in the occurrences of the SKIP mode. The evaluation is performed throughout the frame during the coding of 360 videos, and it is compared to conventional videos. In Section II-A, the behavior of the SKIP mode occurrences throughout the video frame is presented, whereas in Section II-B, a homogeneity-related evaluation of SKIP mode is presented. In both cases, the coding is performed using the HEVC reference software HM-16.16 [8] allied to the library 360Lib-5.0 [2]. A modified HM is employed in order to trace the coordinates and dimensions of PUs coded with **any interframe** prediction mode (including SKIP mode) and with **skip** mode specifically. In addition to that, the experiments are performed according to the Common Test Conditions for 360 Video (CTCs) [9], which specify that 4k videos should be encoded in 3328×1664 resolution, and 6k and 8k videos should be encoded in 4096×2048 resolution. Also, only the first 100 frames of four (*AerialCity*, *Broadway*, *PoleVault,* and *SkateboardInLot*) test sequences are evaluated. Finally, four conventional sequences of resolution 2560×1600 (*NebutaFestival*, *PeopleOnStreet*, *SteamLocomotiveTrain,* and *Traffic*) are coded following the same procedure to serve as a reference.

A. Occurrence of SKIP mode for ERP 360 videos

Based on the coordinates and dimensions of the PUs coded with SKIP mode, it is possible to evaluate if some regions of the frame are more likely to be coded with SKIP than others. Fig. 2 presents the average occurrence of the SKIP mode throughout the video frame after coding. The vertical axis represents the CTU rows, whereas the horizontal axis represents the CTU columns. As the conventional videos have a smaller resolution, they are composed of 40×25 CTUs, whereas the 360 videos have 52×26 CTUs. The color represents the average occurrence rate of SKIP modes per interframe predicted PU, that is, an occurrence rate of 1 means that all PUs coded with interframe prediction are coded with SKIP mode. The average SKIP mode occurrence is calculated as follows: first, all occurrences (**any inter** or **skip**) are normalized according to the dimensions of their PU size (8×8 PUs have unitary weight, occurrences in 16×16 PUs have weight four, and so on). Then these normalized occurrences are grouped according to the CTUs in which they belong, and the occurrence rate of **skip** PUs per **any inter** PUs is calculated for each CTU. Finally, once the SKIP mode occurrence rate in each CTU is calculated to the 100 frames of each video coded with QPs 22, 27, 32, and 37, the results for all co-located CTUs are averaged.

When Fig. 2 is analyzed, it is visible that the SKIP mode behaves differently when comparing conventional videos with ERP 360 videos. When the distribution for conventional videos in Fig. 2 (a) is evaluated, it is visible that the occurrence rate of SKIP mode is around 60% and rarely reaches 70% in some regions. Although there are variations in different regions of the frame, these are marginal and most likely related to the contents of the videos. However, when the distribution for 360 videos of Fig. 2 (b) is evaluated it is visible that in the top/bottom regions the SKIP mode occurs very frequently (85% on average), whereas in the very center of the frame the SKIP mode behaves similarly to conventional videos. These evaluations show that whereas for conventional videos, the occurrences of SKIP mode are content-related, for



(a) Conventional videos

(b) 360 videos

Fig. 3. Cumulative skip distribution according to block variane.

360 videos, this distribution is directly related to the distortions caused by ERP projection. In the most stretched regions, the textures are more homogeneous, and the SKIP mode occurs very frequently, whereas in regions with little or no stretching, the SKIP mode does not occur so frequently and behaves similarly to conventional videos.

B. *Homogeneity influence in SKIP mode selection for ERP 360 videos*

In this section, we present an evaluation of the homogeneity influence of SKIP mode selection in ERP 360 videos. This evaluation aims to find the homogeneity distributions when the SKIP mode is more likely to be selected in both conventional and 360 videos. The homogeneity of the block is evaluated as the variance of its samples [10][11]. To perform such evaluation, only the occurrences of **SKIP** modes are considered. This evaluation also considers that the video is divided into five non-overlapping bands [5][12] to model regions of distinct distortion intensities. Fig. 2 presents the position and shares of each band.

First, the SKIP occurrences are grouped into three categories according to the band they belong to. Occurrences in upper/lower bands are grouped into **polar**, occurrences in mid-upper/mid-lower bands are grouped into **mid-polar**, whereas occurrences in the middle band compose the **middle** group. In the next step, the variance of the blocks coded with SKIP mode is calculated, and such variances are divided into three groups (**polar**, **mid-polar,** and **middle**) according to their bands. Finally, each group has the variance of all blocks coded with SKIP mode in their respective bands. Note that this evaluation is performed individually for each PU size.

Then, these variances are used to calculate the cumulative distribution of SKIPs per block variance in each group. Fig. 3 (a) and (b) present these distributions for PUs 64×64 in conventional videos and 360 videos, respectively, where the vertical axis represents the cumulative distribution and the horizontal axis represents the variance from 0 up to 1500 (the variances can reach 16000). These plots can be read as follows: each variance $\boldsymbol{V}$ is associated with a cumulative distribution $\boldsymbol{D}$, therefore $\boldsymbol{D}$**%** of all SKIPs occur in blocks with a variance smaller than $\boldsymbol{V}$. In other words, a cumulative distribution $\boldsymbol{D}$ represents the share of SKIP modes occurring in blocks with a variance smaller than a threshold $\boldsymbol{V}$.

When the results from Fig. 3 are evaluated, it is possible to see that the distribution for conventional and 360 videos differ significantly. For instance, when a variance of 500 is fixed, 42% of SKIPs in polar bands of conventional videos occur in blocks with variance smaller than 500, whereas in 360 videos 88% of all SKIPs in the polar bands occur in blocks with variance smaller than 500. Evaluating the same variance in the mid-polar bands, it follows that 33% of the SKIPs occur in blocks with variance smaller than 500 in conventional videos, whereas in 360 videos, 79% of the SKIPs occur in blocks with variance smaller than 500. Finally, 23% of the SKIPs in the middle band of conventional videos occur in blocks with variance smaller than 500, whereas for 360 videos 52% of the SKIPs occur in blocks with variance smaller than 500. These values demonstrate that (1) the block homogeneity is a better indicator for the SKIP mode in ERP 360 videos when compared to conventional videos, (2) the SKIPs are more clustered in lower variances for 360 videos when compared to conventional videos, and (3) the concentration of SKIP occurrences in the same variance range varies depending on the region of ERP 360 videos.

In summary, the behavior of the SKIP mode occurrence throughout the frame for ERP 360 video coding is very different when compared with conventional videos. From our analysis, one can learn that the *SKIP mode is more frequently used in the distorted regions of ERP 360 videos and that the homogeneity is decisive for the SKIP mode selection*. These observations are crucial for developing an early SKIP mode decision algorithm for ERP 360 videos that can reduce the complexity with low coding efficiency degradation.

III. Proposed ESA360

This paper proposes the ESA360: an **E**arly **S**KIP mode decision **A**lgorithm for fast ERP **360** video coding. The ESA360 evaluates both the region and the homogeneity of the block being coded to decide whether an early SKIP mode selection can be performed. The block variance is used as a metric for homogeneity, and the variance value is compared with a variance threshold that defines when an early SKIP occurs. The variance thresholds are generated considering a set of coding parameters. Whenever an early SKIP selection is performed, the remaining interframe prediction modes are not evaluated, and therefore the coding is accelerated.

A. *Variance thresholds generation*

The variance thresholds are derived from the cumulative distribution of the SKIP mode, such as presented in Fig. 3. To account for the different intensities of distortions throughout the frame, ESA360 employs an approach where the video frame is divided into five bands, as presented in Fig. 2. In order to have variance thresholds fit for every possible coding scenario, the SKIP decisions when coding the sequences *AerialCity*, *Broadway*, *PoleVault* and *SkateboardInLot* are used to generate multiple cumulative distributions, one for

TABLE I. VARIANCE THRESHOLDS FOR DIFFERENT CU SIZES AND QP VALUES, CONSIDERING POLAR BAND AND RESOLUTION 3328×1664

QP	CU size			
	64×64	32×32	16×16	8×8
22	8	4	2	2
27	20	7	4	2
32	25	9	4	2
37	32	10	4	2

each coding scenario. In the context of ESA360, a coding scenario is composed of one instance of each one of following attributes: **(1) band position**, which can be either polar or mid-polar; **(2) CU size**, which can be from 8×8 up to 64×64; **(3) video resolution**, which can be either 4096×2048 or 3328×1664; and **(4) QP value**, which can be 22, 27, 32, or 37.

It is necessary to generate one cumulative distribution – and thus, one set of variance thresholds – for each coding scenario because the behavior of SKIP mode occurrences throughout the frame is related to these attributes in different manners and intensities. Considering the band position, in polar bands the SKIP mode occurs more frequently, as visible in Fig. 2. ESA360 is not applied in the middle band since it presents small projection distortion, therefore, its distribution is not calculated. As for different CU sizes, smaller CUs tend to comprise the same object more frequently; therefore, they are usually more homogeneous than larger CUs. The frame resolution also influences in the SKIP mode occurrences. For higher resolutions, the same block size represents less of a scene; therefore, they are more likely to be more homogeneous. Finally, when using larger QPs, the SKIP mode is used more frequently, therefore it is essential to have distinct cumulative distributions – and their respective variance thresholds – for distinct QP values.

In sequence to that, the variance associated with multiple distribution ***D*** values were extracted for each cumulative distribution. More precisely, ***D*** assumed values between 0.05 and 0.90, with steps of 0.05, totalizing 18 variance values for each cumulative distribution. Then a linear regression was performed with these variances and ESA360 (described in Section III-B) to find the variance threshold that results in the best complexity reduction – coding efficiency tradeoff. Finally, the linear regression concluded that using the variance values associated with cumulative distribution ***D*** equal to 0.35 represents the best variance threshold for the polar bands, whereas for mid-polar bands, the best variance threshold is associated with cumulative distribution ***D*** equal to 0.15.

Considering this, it is visible that the control scheme of ESA360 is based on a 4D data structure that holds the variance threshold to be employed during the early decision for different coding scenarios. This structure is indexed as `var[band][CUsize][resolution][QP]`, where **band**, **CU size**, **resolution,** and **QP** correspond to the current coding scenario, that is, the characteristics of current block/video and

```
ESA360 - Early SKIP mode decision algorithm
Input: resolution, QP
1 allTh = initialize(resolution, QP)
2 for each CU in video do:
3   var ← calcVariance(CU_SAMPLES)
4   currTh = getTh(allTh, CU_POSITION, CU_SIZE)
5   if var < currTh do:
6     Early SKIP decision
7   else:
8     Perform entire interframe prediction
```

Fig. 4. Pseudocode of ESA360 algorithm.

the active QP. Table I presents the variance threshold for multiple block sizes and QPs, supposing that the blocks are on polar bands, and the video resolution is 3328×1664. When analyzing Table I, it is visible that for blocks belonging to the same band and same video, significantly different variance thresholds are employed depending on the coding scenario since such thresholds vary from 2 up to 32. When these variance thresholds are compared to the ones of different band and video resolution (not presented in this paper), these values differ significantly as well. Therefore, Table I serves as evidence that ESA360 is based on an adaptive variance threshold scheme to control in which cases an early SKIP decision will be performed.

B. ESA360 Algorithm

Based on this set of variance thresholds, ESA360 is defined according to the algorithm present in Fig. 4. ESA360 receives as input from the encoder itself the video **resolution** and **QP** values being used. When analyzing Fig. 4, it is visible that in **line 1**, the algorithm loads into **allTh** all the thresholds that may be used during the current coding based on the video resolution and QP being used. The **for loop** in **line 2** represents the main coding loop, in which all **CUs** are coded. Therefore, for each **CU** being coded, **line 3** calculates the variance for such **CU** based on its samples. In sequence, in **line 4**, the threshold for the current CU (**currTh**) is loaded from **allTh** based on the current **CU** vertical **position** in the frame and the **CU size**. The **band** to which the CU belongs is derived from its vertical position in the frame. The conditional in **line 5** evaluates if the variance of the current **CU** (**var**) is smaller than the current threshold (**currTh**). If true, then an **early SKIP decision** is carried for the current CU, and other interframe prediction steps are bypassed. Otherwise, the entire interframe prediction is conducted for such **CU**. In the next loop iteration, **CU** may present a different size or vertical position (and, therefore, belong to a different band), and the threshold **currTh** will be adapted accordingly. It is important to emphasize that since the middle band comprehends a region of the frame with little or no distortion, ESA360 does not perform early decisions during the coding of such band, therefore only the coding of 70% of the frame area is accelerated (in polar and mid-polar bands).

IV. EXPERIMENTAL RESULTS

The proposed ESA360 algorithm is implemented in the HM-16.16, together with 360Lib-5.0. The complexity reduction is evaluated as the amount of processing time saved when using ESA360 in comparison with the original encoder. The coding efficiency is evaluated according to the Bjøntegaard Delta Rate (BD-BR) [13]. The BD-BR is a metric to compare the coding efficiency of a target encoder in relation to a reference encoder, and it measures the bitrate variation for the target encoder to achieve the same visual quality of the reference. In addition, the BD-BR is calculated based on the Weighted-to-spherically uniform PSNR (WS-PSNR) [2] that is used instead of the conventional PSNR. The WS-PSNR evaluates the quality of ERP 360 videos more properly [2]. The experiments were conducted following the CTCs [9], using all the 12 available test sequences and QPs 22, 27, 32, and 37. Besides, the experiments were conducted in a server equipped with an Intel Xeon Gold-5118 @ 2.3 GHz and 56GiB of RAM. The experimental results obtained with this setup are depicted in Table II. In this table, column ΔT represents the complexity reduction, while column BD-BR represents the coding efficiency of the proposed ESA360.

TABLE II. RESULTS OF PROPOSED ESA360

Resolution	Video	ΔT [%]	BD-BR [%]
4096×2048 (original) 3328×1664 (coding)	*AerialCity*	9.99	0.16
	PoleVault	10.09	0.03
6144×3072 (original) 4096×2048 (coding)	*Balboa*	21.44	0.30
	BranCastle	2.67	0.34
	Broadway	20.61	0.66
	Landing	12.49	0.53
8192×4096 (original) 4096×2048 (coding)	*ChairliftRide*	5.43	0.09
	GasLamp	19.61	0.00
	Harbor	21.16	0.06
	KiteFlite	7.58	-0.01
	SkateboardInLot	13.47	0.45
	Trolley	18.58	0.00
	Average	**13.59%**	**0.24%**

When analyzing the results from Table II, it is visible that the proposed ESA360 algorithm is able to achieve 13.59% of complexity reduction, on average, whereas incurring an average BD-BR of only 0.24%. The variations in ΔT results between different videos are directly related to the homogeneity of the video sequences. For instance, sequences such as Harbor have the polar and mid-polar bands very homogeneous and ESA360 is able to perform early SKIP decisions more frequently. On the other hand, sequences such as BranCastle have the polar and mid-polar bands not so homogeneous and ESA360 performs less early SKIP decisions. The variations in BD-BR between different videos are explained by the precision of the decisions performed by ESA360. In some cases, the encoder would normally employ the SKIP mode and ESA360 is able to predict it correctly, therefore, other interframe prediction steps are avoided whereas the coding efficiency is not harmed. In other cases, the reference encoder would normally not employ the SKIP mode, and the decisions of ESA360 cause an impact in the coding efficiency incurring a higher BD-BR.

Table III presents a comparison of ESA360 with the average results of two related works. For comparing the complexity reduction and coding efficiency tradeoff between solutions, the RCR metric is employed, which represents the time reduction (ΔT) divided by coding efficiency (BD-BR). When analyzing the results from Table III it is visible that Liu [4] achieves an average complexity reduction of 32% and an average BD-BR of 1.3%, whereas Ray [5] achieves 15% of average complexity reduction with an average BD-BR of -0.18%. When comparing the results of ESA360 with related works, it is visible that ESA360 achieves a smaller complexity reduction. However, since ESA360 presents CRC equal to 56.63 while Liu [4] presents RCR equal to 24.62, it is visible that ESA360 presents a complexity reduction – coding efficiency tradeoff better than Liu [4]. When comparing the proposed ESA360 algorithm with Ray [5], it is visible that ESA360 performed worse in both complexity reduction and coding efficiency. However, while ESA360 is adaptive to the properties of the frame and block being coded, the work [5] consists of simplifying coding tools linearly in pre-defined regions of the frame with no concern to the differences between distinct video sequences.

Finally, it is important to mention that our proposed implementation of ESA360 is on for only 70% of the frame area since the middle band is encoded with the regular HEVC encoder. It is possible to associate ESA360 with most generic complexity reduction techniques to improve the overall performance by performing complexity reduction in the middle band as well. Also, it is possible to employ more band divisions to reduce the middle band area and use thresholds more specialized for each band.

TABLE III. ESA360 COMPARED TO RELATED WORKS

Work	ΔT [%]	BD-BR [%]	RCR
ESA360	13.59	0.24	56.63
Liu [4]	32.00	1.30	24.62
Ray [5]	15.00	-0.18	-83.3

V. CONCLUSION

This work presented ESA360: an **E**arly **S**KIP mode decision **A**lgorithm for fast ERP **360** video coding. ESA360 is based on evaluations that demonstrate the difference between coding conventional videos and ERP 360 videos in the SKIP mode occurrences. Whereas for conventional videos the SKIP occurrences are directly related to the video content, for ERP 360 videos the SKIP occurrences are highly related to the frame region and intensity of distortion incurred in such region. In addition, these evaluations show that the homogeneity of the block poses great influence in the occurrences of SKIP mode for ERP 360 videos when compared to conventional videos. Finally, experimental results obtained from ESA360 demonstrate that it is possible to design complexity reduction algorithms for coding ERP 360 videos when the coding tools present a specific behavior due to the projection distortions.

REFERENCES

[1] Cisco, “Cisco visual networking index: forecast and methodology 2017-2022”, 2019.

[2] Y. Ye, E. Alshina, J. Boyce, “Algorithm descriptions of projection format conversion and video quality metrics in 360Lib,” Doc. JVET-E1003, Geneva, 2017.

[3] High Efficiency Video Coding, ITU-T Rec. H.265 and ISO/IEC 23008-2, Feb. 2018.

[4] Z. Liu, P. Song, and M. Zhang, "A CU Split Early Termination Algorithm Based KNN for 360-Degree Video," 2019 Data Compression Conference (DCC), Snowbird, UT, USA, 2019.

[5] B. Ray, J. Jung, and M. Larabi, "A Low-Complexity Video Encoder for Equirectangular Projected 360 Video Content," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, 2018, pp. 1723-1727.

[6] Y. Zhou, Z. Chen, S. Liu, “Fast sample adaptive offset algorithm for 360-degree video coding,” in Signal Processing: Image Communication, vol. 80, pp. 1 – 11, 2020.

[7] J. Vanne, M. Viitanen, T. D. Hamalainen and A. Hallapuro, "Comparative Rate-Distortion-Complexity Analysis of HEVC and AVC Video Codecs," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1885-1898, Dec. 2012.

[8] High Efficiency Video Coding Test Model 16.16, available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.16/.

[9] J. Boyce, E. Alshina, A. Abbas, Y. Ye, “Common test conditions and evaluation procedures for 360° video coding,” Doc. JVET-D1030, Chengdu, 2016.

[10] D. Palomino, M. Shafique, A. Susin and J. Henkel, "TONE: Adaptive temperature optimization for the next generation video encoders," 2014 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED), La Jolla, CA, 2014, pp. 33-38

[11] B. Zatt, M. Shafique, F. Sampaio, L. Agostini, S. Bampi and J. Henkel, "Run-time adaptive energy-aware Motion and Disparity Estimation in Multiview Video Coding," 2011 48th ACM/EDAC/IEEE Design Automation Conference (DAC), New York, NY, 2011, pp. 1026-1031.

[12] I. Storch, B. Zatt, L. Agostini, L. A. da Silva Cruz and D. Palomino, "FastIntra360: A Fast Intra-Prediction Technique for 360-Degrees Video Coding," 2019 Data Compression Conference (DCC), Snowbird, UT, USA, 2019, pp. 605-605.

[13] G. Bjøntergaard, “Calculation of average PSNR differences between RD curves,” Doc. VCEG-M33, Tech. Rep., 2001.