Contour-based Multidirectional Intra Coding for HEVC

Thorsten Laude and Jörn Ostermann
Institut für Informationsverarbeitung, Leibniz Universität Hannover, Germany
Email: {laude, office}@tnt.uni-hannover.de

Abstract—**Intra coding is an indispensable part of all video coding systems. However, the intra prediction in the HEVC standard has several limitations with respect to the available amount of information which is exploited for the prediction and the limited number of prediction directions per block. In this paper, we propose a novel contour-based multidirectional intra coding mode which combines computer vision technologies with conventional video coding technologies to overcome these limitations. Structural parts and smooth parts of the video signal are predicted separately. While the structural parts are predicted based on contour detection, parameterization and extrapolation algorithms, the smooth areas are filled by means of low complexity technologies like sample value continuation. Thereby, the coding efficiency is considerably increased, with weighted average BD-rate gains of up to 1.9% compared to HM-16.3. A stand-alone codec implementation achieves average bit rate savings of 29.7% over JPEG and outperforms related work.**

I. INTRODUCTION

During the last decades a tremendous improvement of video coding algorithms could be observed. In January 2013, the Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG finished the technical work for the latest video coding standard, High Efficiency Video Coding (HEVC) [1]. It achieves the same visual quality at half the bit rate compared to the predecessor standard AVC [2].

The prediction tools which led to the prosperous application of these video coding standards can be roughly distinguished into inter and intra coding tools. While intra coding solely relies on information which is contained in the current picture, inter coding uses the redundancy between different pictures to further increase the coding efficiency. Therefore, in general, intra coding requires considerably higher bit rates than inter coding to achieve the same visual quality for typical video signals.

Nevertheless, intra coding is an essential part of all video coding systems: it is required to start a video transmission, for random access into ongoing transmissions and for error concealment. In this paper, we aim at increasing the coding efficiency for intra prediction. Analyzing the intra mode in HEVC [3] reveals multiple areas for improvement hereof. For instance, only one adjacent row/column is used as prediction basis for the current block (which in case of HEVC is referred to as coding unit or CU). Furthermore, in case of angular prediction, only one direction can be applied per CU. Due to these limitations, high bit rates are required for the residuals of intra coded CUs. To overcome these restrictions, we extend the size of the area which is used as prediction basis. Additionally, we combine contour driven technologies with video coding technologies to further improve our prediction.

Our contribution in this paper is a new coding mode which we refer to as *contour-based multidirectional intra coding* or CoMIC. Conceptually, structural parts and smooth parts are predicted separately. For this purpose, contours are detected, parameterized and extrapolated while smooth areas are filled with other algorithms. The philosophy behind our approach is that the entire information for the prediction process is extracted from reconstructed parts of the current picture. Thus, no signaling overhead is needed except for the mode usage itself.

The remainder of this paper is organized as follows: In Section II we analyze related work and discuss the differences to our approach. Our novel coding mode is presented in Section III. Section IV describes the evaluation of our method and Section V concludes the paper.

II. RELATED WORK

The related work in the literature can be categorized into several groups: fast encoder algorithms, error concealment approaches, inpainting algorithms, and coding tools.

In the context of edge-based fast encoder algorithms, Yuan and Sun [4] proposed to calculate an edge map for the current picture and to use this knowledge for intra direction decisions. In contrast to applying edge information for the sake of an accelerated encoding process, we utilize extracted edge information to increase the coding efficiency with a new coding mode. As another difference to our Canny-based edge detection, Yuan et al. apply only the Sobel operator to generate the edge map.

Asheri et al. [5] as well as Au and Chan [6] propose edge-based error concealment algorithms. They utilize edge information to recover from transmission errors. On the contrary to our approach, Asheri et al. use only the Sobel operator for the edge detection. Although Au and Chan rely on Canny edge detection as we do, they not only detect edges in the current picture but also in previously coded pictures. However, this is not desirable for an intra coding mode. Thus, we restrict ourselfs to edges in the current picture. In summary, differently from our approach, their algorithms do not influence the actual coding process but are used as post processing in case of errors.

Closest to our approach, Liu et al. add an edge-based inpainting coding mode to JPEG [7] as well as to JPEG 2000 and AVC [8]. Similar to our approach, they separate structural information (e.g. contours) and smooth areas and apply different prediction methods to them. The HEVC intra prediction relies on pixel extrapolation. Liu et al. add contour processing, inpainting and side information signaling on top of that. In contrast to that, we only use contour processing and extrapolation. There are two main differences between our contribution and the works of Liu et al. They signal side information related to the contour shape. In contrast to that, we extract all information at the decoder and have no signaled side information for the contour shape. Additionally, they use a high complexity algorithm (PDEs) for the prediction of smooth areas while we aim at low complexity solutions. Our method will be compared to their JPEG-based implementation in [7].

The main difference between our contribution and related inpainting approaches is that we assume a causal encoding system where only already coded blocks (i.e. above of and to the left of the current block) are available for the contour extrapolation. In contrast to that, inpainting approaches use the signal on all sides of the missing block. This knowledge is highly beneficial for the contour shape prediction. Taking into account that this information is not available for the decoder, our method solves a harder problem.

III. CONTOUR-BASED MULTIDIRECTIONAL INTRA CODING

In this section we discuss our novel intra coding mode COMIC. By combining edge detection, parameterization and extrapolation algorithms with conventional video coding algorithms like sample value continuation we design an intra coding mode for the prediction in multiple directions with minor signaling overhead. Conceptually, contour information is extracted from reconstructed samples and employed to extrapolate these contours into the current CU.

Since we propose an intra coding mode, we limit our area of operation to the current CU and adjacent parts of the same picture. Deciding on the number of involved adjacent parts in the prediction process entails a trade-off: larger areas allow the extraction of more information while they simultaneously increase the amount of samples which need to be available and the computational complexity for information extraction. Our experimental results suggest that a good trade-off is to use areas of equal size as the current CU on the top, top left and left side. Thus, the size of the selected reconstructed area is three times the size of the current CU. Fig. 1 illustrates the selection of the reconstructed area with respect to the current CU. It is noteworthy that all operations which are described in the following solely rely on reconstructed samples. Therefore, the whole prediction process can be applied on both sides of the video coding pipeline without signaling any information except for the mode usage itself.

A. Contour detection

As first step of our intra prediction algorithm, contours are detected in the reconstructed area. Our contour detection is based on the well known Canny edge detection algorithm [9]. The Canny edge detection algorithm can be controlled by setting two thresholds for the suppression of weak edges. Taking into account that the characteristics of video signals may be very different, it is imperative to define thresholds which result in good detection results for all video signals. Intuitively, it is unlikely that fixed thresholds will provide the capability to achieve these results. Thus, we chose to calculate signal dependent thresholds.

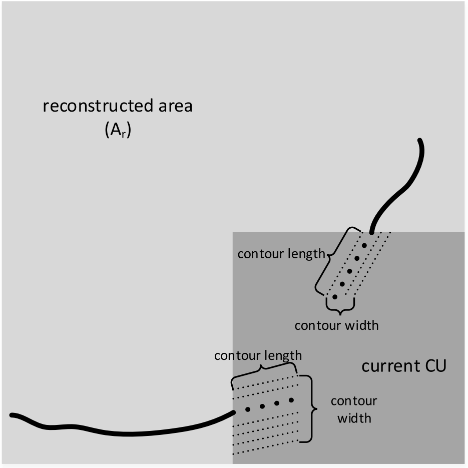

Fig. 1. Illustration of the contour extrapolation process. The extracted contours (solid lines) in the already reconstructed CUs are parameterized and extrapolated (thick, dotted lines) into the current CU.

Otsu proposed a method [10] to segment a gray scale image into two classes by maximizing the inter-class variance. In [11], Fang et al. have shown that Otsu's thresholds are also applicable to determine the Canny thresholds. Following their analysis, we set the two Canny thresholds, $t_{\text{Canny, high}}$ and $t_{\text{Canny, low}}$, based on the Otsu threshold t_{Otsu} to:

$$t_{\text{Canny, high}} = t_{\text{Otsu}}, \quad t_{\text{Canny, low}} = 0.5 \cdot t_{\text{Otsu}}. \tag{1}$$

B. Contour parameterization

Once the edges have been detected as depicted above, a binary image is available that describes all edge pixels in the reconstructed area. Since our goal is the extrapolation of distinctive contours, we label individual contours in the binary edge picture according to Suzuki and Be [12]. These contours are processed by our contour parameterization algorithm.

Each contour is described by a polynomial parameterization. The degree of the polynomial needs to be chosen carefully. On the one hand a high degree allows the approximation of irregularly shaped or curved contours. On the other hand, taking into account that many contours only have a length of a few pels, a high polynomial degree runs the risk of overfitting the posed problem. We have tested various approaches to define the degree of the polynomial.

In the end, experimental results suggest that polynomials with two parameters, i.e. lines, form the best trade-off between the above mentioned aspects. Additionally, for contours which hit the left border of the current CU, the horizontal coordinate

x is used as independent variable while the vertical coordinate y is used as independent variable for contours that hit the upper border of the current CU. This differentiation facilitates the parameterization of horizontal and vertical structures. For each kind of contour, an example is illustrated in Fig. 1. Thus, the polynomial $p(a)$, with a being the independent variable (i.e. either x or y) and with coefficients $\beta_i \in \mathbb{R}$ is parameterized as noted in Eq. 2:

$$p(a) = \beta_0 + \beta_1 a. \tag{2}$$

In the following, we describe the polynomial calculation only for x as independent variable. For n given edge pixels (x_i, y_i) with $i = 1, \ldots, n$ we write Eq. 3:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^{n \times 1}. \tag{3}$$

Reformulating Eq. 2, we define the approximation of the given problem as in Eq. 4 where ϵ denotes the approximation error:

$$\begin{aligned} y &= \begin{bmatrix} 1 & \\ \vdots & x \\ 1 & \end{bmatrix} \beta + \epsilon \quad \text{with } \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \epsilon \in \mathbb{R}^{n \times 1} \\ &= X\beta + \epsilon. \end{aligned} \tag{4}$$

This can be seen as a linear regression problem (with β_i as regression parameters) which can be solved with a least squares approach, i.e. by minimizing $f(\beta)$ in Eq. 5:

$$f(\beta) = (y - X\beta)^T (y - X\beta). \tag{5}$$

The estimated regression parameters, denoted as $\hat{\beta}$, are calculated by Eq. 6:

$$\hat{\beta} = (X^T X)^{-1} X^T y. \tag{6}$$

C. Contour extrapolation

On the basis of the previously described contour parameterization in the reconstructed area, the signal is extrapolated into the current CU. The extrapolation process is performed following the illustration in Fig. 1. For each extracted contour (illustrated by thick solid lines), the signal is extrapolated along the approximated polynomial (thick dotted lines).

In contrast to the conventional intra prediction in HEVC which is based on the unaltered continuation of adjacent sample values into the predicted CU, the extrapolation sample value for our approach is modified along the extrapolated contour. The motivation for this altering of the sample value is based on the observation that the accuracy of the extrapolated contours with respect to the original input signal decreases with increasing distance to the reconstructed pixels. To cope with this observation, the sample value s_a of the adjacent pixel is diminished towards the sample value s_m for the mean fill process introduced in the next section. The resulting sample value for the extrapolation for a given pixel on the extrapolated contour is denoted as s_e. Additionally, let (x_a, y_a) and (x_e, y_e) represent the coordinates of the adjacent pixel (i.e. the extrapolation basis) and of the current pixel on the extrapolated contour, respectively. Furthermore, let d_{max} be the distance after which the extrapolated sample value is completely diminished towards s_m and d the distance between the adjacent pixel and the current pixel on the extrapolated contour. With this notation, the diminishing can be described mathematically as noted in Eq. 7.

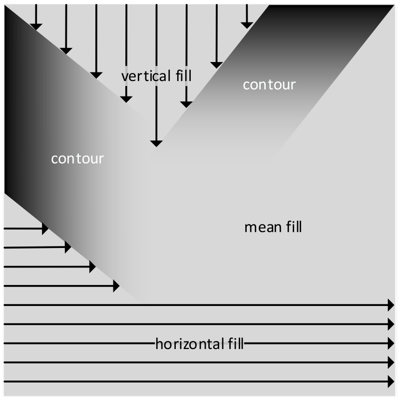

Fig. 2. Smooth areas are filled either by horizontal/vertical sample value continuation or by filling in the mean sample value of the reconstructed area. Furthermore, it can be observed how the contours are diminished towards the mean sample value.

$$s_e = \frac{s_m d + s_a (d_{\text{max}} - d)}{d_{\text{max}}} \tag{7}$$

with

$$d = \sqrt{(x_a - x_e)^2 + (y_a - y_e)^2}. \tag{8}$$

Based on subjective inspections for the first picture of the Basketball Drive sequence, we determined empirically that a suitable value for d_{max} is 40% of the size of the current CU. An example for the diminishing process is given in Fig. 2.

As it can be observed in Fig. 1, the width of the extrapolated contour itself (thick dotted line) is only one pel. However, typically the contours in the coded video signals are wider. Thus, the width of the extrapolated contour is extended as illustrated by the thin dotted lines. As criterion for the extended width, the sample values of pixels on the border between the current CU and the reconstructed area in the neighborhood of the adjacent pixel (with sample value s_a) which is used as extrapolation base are analyzed. If the absolute sample value difference between the adjacent pixel (s_a) and the i-th cohesive pixel ($s_{c,i}$) is below a threshold t, $|s_a - s_{c,i}| < t$, the contour is widened to this cohesive pixel. Experimental results indicate that $t = 30$ is a reasonable value for this threshold. Furthermore, overlapping contours are handled by averaging.

D. Smooth area prediction

Following our approach to apply different algorithms for the prediction of contours and of smooth areas, these smooth areas are predicted subsequently to the contour extrapolation. We distinguish between two different algorithms for the smooth area prediction: the horizontal and vertical continuation of adjacent sample values and the filling of unpredictable areas

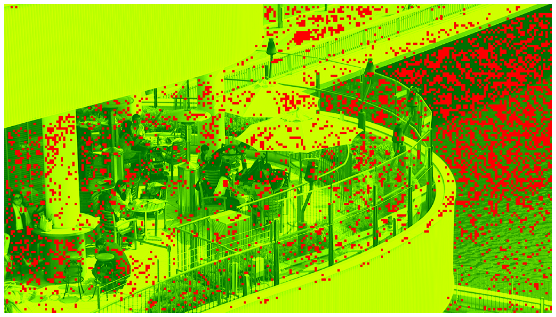

Fig. 3. Example for the usage of our new coding mode for the first picture of the sequence *BQTerrace*: the CUs which are highlighted in red are coded with our mode while the remaining parts of the picture (highlighted in green) are coded with the HEVC intra mode. It can be observed that parts of the picture with clear directional structures such as the sides of the building are coded with the HEVC intra mode. Other parts with more complex structural information (e.g. the rippled water and the terrace) are often coded with our mode.

TABLE I

WEIGHTED AVERAGE BD-RATES: NEGATIVE NUMBERS INDICATE INCREASED CODING EFFICIENCY. THE CODING EFFICIENCY IS NOTICEABLY IMPROVED BY THE NEW CODING MODE. THE USAGE INDICATES THE PERCENTAGE OF PIXELS PREDICTED WITH OUR MODE.

	AI	LD	RA	**Mean**	Usage (AI)
Bike 2	-0.4%	-0.6%	-0.5%	**-0.5%**	9.5%
Bike 3	-0.3%	-0.5%	-0.5%	**-0.4%**	9.0%
Bike 4	-0.3%	-0.4%	-0.4%	**-0.4%**	7.3%
Bike 5	-0.3%	-0.5%	-0.4%	**-0.3%**	6.5%
Bike 8	-0.2%	-0.4%	-0.4%	**-0.3%**	6.8%
Bike 14	-0.4%	-0.5%	-0.5%	**-0.5%**	8.3%
BVI Ball Under Water	-1.2%	-1.4%	-1.8%	**-1.5%**	12.7%
BVI Bubbles Clear	-0.5%	-1.8%	-1.9%	**-1.4%**	9.2%
BVI Sparkler	-0.5%	-0.7%	-0.8%	**-0.7%**	20.0%
Basketball Drive	-0.3%	-0.1%	-0.2%	**-0.2%**	11.5%
BQTerrace	-0.4%	-0.2%	-0.5%	**-0.3%**	16.6%
Kimono	-0.3%	-0.1%	-0.1%	**-0.2%**	9.7%
Mean	**-0.4%**	**-0.6%**	**-0.7%**	**-0.6%**	10.6%

by the mean sample value of the reconstructed area. Both algorithms are visualized in Fig. 2.

The horizontal and vertical continuation is based on the repetition of the adjacent sample value from the reconstructed area into the current coding unit. Thereby, we exploit that in general there are no discontinuities at the border of the coding unit. The sample values are continued until the first contour is hit, as illustrated in Fig. 2. In summary, the horizontal and vertical continuation is similar to the HEVC intra angular prediction modes 10 and 26, respectively, which denote horizontal and vertical intra prediction.

Unpredictable parts of the current coding unit, i.e. parts which are neither predicted by extrapolated contours nor by horizontal and vertical continuation, are filled by the mean sample value of the reconstructed area. Here it is assumed that the mean brightness of neighboring parts in the picture is similar. The difference to the HEVC intra DC mode is that we use the mean sample value of the entire reconstructed area as defined above instead of just using the mean value of the adjacent sample row/column. This way we can better cope with local fluctuation in the signal.

IV. EXPERIMENTAL RESULTS

In this section, the proposed contour-based multidirectional intra coding mode is evaluated. For this purpose, the algorithm was implemented as an additional coding mode in the HEVC reference software HM-16.3. It is selected by the rate-distortion optimization of the encoder if it provides the lowest rate-distortion costs and its usage is signaled by one binary flag in the CU syntax.

A set of 12 test sequences as listed in Table I was used for the evaluation. To cover a variety of different contents, the sequences originate from several databases. Some sequences (*Basketball Drive*, *BQ Terrace* and *Kimono*) are used in the MPEG standardization. Other sequences (*Ball Under Water*, *Bubbles Clear* and *Sparkler*, spatial resolution 1920 × 1080) were published recently in the *BVI Texture* database [13]. They were chosen because of the rich structural content they represent. Additionally, some self-recorded sequences (*Bike #*, spatial resolution 1280 × 720) were included in the test set. The encoder configurations all intra (AI), low delay (LD) and random access (RA) as defined in the HEVC *common test conditions* [14] were used for the evaluation. We test our method for high bit rates because our algorithm requires contour information which is not available in low quality pictures due to the strong quantization. Therefore, also taking into account that many application scenarios which strongly rely on intra coding (such as production scenarios) require very high quality, low quantization parameters (12, 17, 22, 27) were selected. The Bjøntegaard-Delta (BD)-rate [15] was calculated to evaluate the coding efficiency of the proposed coding tool. Additionally, as suggested in [16], weighted average BD-rates BD_{avg} were calculated with weighting factors of 6/1/1 for the three color components Y/Cb/Cr. The resulting BD-rates are summarized in Table I. It is worth noting that the coding efficiency is increased across all sequences, with a mean weighted average gain of 0.6%. For individual sequences, coding gains as high as 1.9% are observed. The reason why our coding gains for the LD and RA configuration are higher than for the AI configuration is that those blocks which are difficult to predict (i.e. which are intra coded even in inter pictures) require a large amount of the total bit rate. With our algorithm, we save more bits for these difficult blocks than for blocks which are easily predictable by the HEVC inter mode. Thereby, we save the same absolute amount of bits for these difficult blocks in intra and inter pictures. However, taking into account the lower bit rate for inter pictures compared to intra pictures and the mentioned high ratio of difficult blocks in the total bit rate, the percentage gains are higher for inter pictures.

Our premise that the new coding mode is beneficial for video signals with complex structural information is confirmed by the results shown in Fig. 3. In this figure, the usage of the new coding mode is highlighted in red for a representative picture of the *BQTerrace* sequence. It can be observed that the new coding mode is heavily used in areas with complex structural information, e.g. for the rippled water and on the terrace. Due to the diminishing it is not used for the continuous structure on the parapet. Furthermore, we analyzed the origin

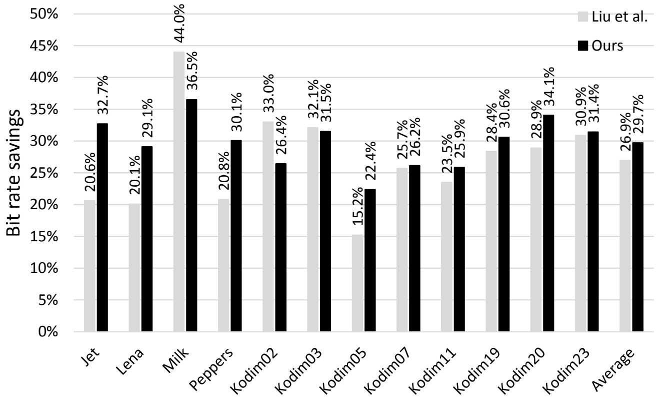

Fig. 4. Bit rate savings of our method and of [7] compared to JPEG

of our gain by determining the previous intra direction applied for all those pixels, which are coded with our method now. It can be observed that the major part (∼60%) of the gain comes from the planar and DC intra modes. This is an additional indication that clear angular structures can be predicted with the normal angular modes while the prediction of complex structures benefits from the new coding mode. No visual artifacts are visible between structural parts and smooth parts due to the signaled residual. We analyzed the complexity of our method by measuring the encoder and decoder runtime for our unoptimized code. The encoder and decoder runtime is increased in average by 49% and 116%, respectively.

Liu et al. propose their works in the context of JPEG and AVC. Thereby, our results are not directly comparable since we propose our method for HEVC. Nevertheless, we believe that a comparision with the state-of-the-art works closest to our method is highly desirable. For this purpose, we provide a stand-alone codec implementation of our method in addition to our HEVC-based implementation. Basically, we add our prediction method on top of JPEG (8×8 DCT, quantization, entropy coding). In consequence, this enables the comparison with their JPEG-based work [7]. For a fair comparison, we selected the same images and encoder configuration that Liu et al. chose for their work. The results are presented in Fig. 4. While they achieve average bit rate savings of 26.9%, our method has 29.7%. Thereby, it can be concluded that our method provides higher coding efficiency. For two images with large homogeneous areas without contours and smooth color gradients (*Milk*) and contours within structured areas (*Kodim02*), one can observe the advantages of the PDEs that are used by Liu et al. For these two images, our bit rate savings are lower than those of the state-of-the-art because the PDEs perform better for this content. However, these higher results of Liu et al. are achieved at the expense of high complexity for solving PDE systems. Otherwise, our method considerably outperforms Liu et al.'s work for some images (e.g. Jet, Lena, Peppers and Kodim05). These images include clear contours within areas of smooth background for which our contour extrapolation and low complexity smooth area prediction algorithms provide supperior results.

V. CONCLUSIONS

In this paper, we present a new coding mode (the *contour-based multidirectional intra coding* or COMIC) which is based on the detection, parameterization and extrapolation of contours and the filling of smooth areas with different algorithms. The entire prediction process is based on information which is extracted from reconstructed samples. Thus, no signaling of side information except for the mode usage itself is required. By means of this new coding mode, the coding efficiency is considerably increased, with weighted BD-rate gains of up to 1.9% compared to HM-16.3. Furthermore, a stand-alone codec implementation of our method outperforms the state-of-the-art work with bit rate savings of 29.7% over JPEG.

REFERENCES

[1] "ITU-T Recommendation H.265/ ISO/IEC 23008-2:2013 MPEG-H Part 2: High Efficiency Video Coding (HEVC)," 2013.
[2] Philippe Hanhart, Martin Rerabek, Francesca De Simone, and Touradj Ebrahimi, "Subjective quality evaluation of the upcoming HEVC video compression standard," in *SPIE Optical Engineering + Applications*, oct 2012, p. 84990V.
[3] Jani Lainema, Frank Bossen, Woo-Jin Han, Junghye Min, and Kemal Ugur, "Intra Coding of the HEVC Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1792–1801, dec 2012.
[4] Yule Yuan and Xiaohang Sun, "Edge Information Based Effective Intra Mode Decision Algorithm," in *2012 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC)*. aug 2012, pp. 628–633, IEEE.
[5] Hadi Asheri, Hamid Rabiee, Nima Pourdamghani, and Mohammad Ghanbari, "Multi-directional Spatial Error Concealment using Adaptive Edge Thresholding," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 880–885, aug 2012.
[6] O.C. Au and S.-H.G. Chan, "Edge-Directed Error Concealment," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 3, pp. 382–395, mar 2010.
[7] Dong Liu, Xiaoyan Sun, Feng Wu, Shipeng Li, and Ya-Qin Zhang, "Image Compression With Edge-Based Inpainting," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 10, pp. 1273–1287, oct 2007.
[8] Dong Liu, Xiaoyan Sun, Feng Wu, and Ya-Qin Zhang, "Edge-oriented Uniform Intra Prediction.," *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1827–36, oct 2008.
[9] John Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 8, no. 6, pp. 679–698, nov 1986.
[10] Nobuyuki Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
[11] Mei Fang, Guang Xue Yue, and Qing Cang Yu, "The Study on an Application of Otsu Method in Canny Operator," in *Proceedings of the 2009 International Symposium on Information Processing (ISIP)*, 2009.
[12] Satoshi Suzuki and Keiichia Be, "Topological Structural Analysis of Digitized Binary Images by Border Following," *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32–46, apr 1985.
[13] Miltiadis Alexios Papadopoulos, Fan Zhang, Dimitris Agrafiotis, and David Bull, "A Video Texture Database for Perceptual Compression and Quality Assessment," in *International Conference on Image Processing (ICIP)*, Quebec, Canada, 2015.
[14] Frank Bossen, "JCT-VC L1100: Common HM test conditions and software reference configurations. 12th Meeting of the Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11. Geneva, CH," 2013.
[15] Gisle Bjøntegaard, "VCEG-AI11: Improvements of the BD-PSNR model. ITU-T Study Group 16 Question 6. 35th Meeting, Berlin, Germany," 2008.
[16] Gary J. Sullivan and Jens-Rainer Ohm, "Meeting Report of the Fourth Meeting of the Joint Collaborative Team on Video Coding," *ITU-T/ISO/IEC JCT-VC Document JCTVC-D500*, 2011.