Technical Report (Oct. 2015): Region of Interest Coding for Aerial Surveillance Video Using AVC & HEVC

Holger Meuel, Florian Kluger and Jörn Ostermann

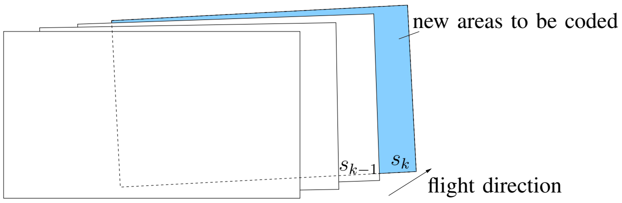
Institut für Informationsverarbeitung
Gottfried Wilhelm Leibniz Universität Hannover, Germany
{meuel, kluger, office}@tnt.uni-hannover.de

***Abstract*—Aerial surveillance from *Unmanned Aerial Vehicles* (UAVs), *i. e.* with moving cameras, is of growing interest for police as well as disaster area monitoring. For more detailed ground images the camera resolutions are steadily increasing. Simultaneously the amount of video data to transmit is increasing significantly, too. To reduce the amount of data, *Region of Interest* (ROI) coding systems were introduced which mainly encode some regions in higher quality at the cost of the remaining image regions. We employ an existing ROI coding system relying on *global motion compensation* to retain full image resolution over the entire image. Different ROI detectors are used to automatically classify a video image on board of the UAV in ROI and non-ROI. We propose to replace the modified *Advanced Video Coding* (AVC) video encoder by a modified *High Efficiency Video Coding* (HEVC) encoder. Without any change of the detection system itself, but by replacing the video coding back-end we are able to improve the coding efficiency by 32 % on average although regular HEVC provides coding gains of 12–30 % only for the same test sequences and similar PSNR compared to regular AVC coding. Since the employed ROI coding mainly relies on intra mode coding of new emerging image areas, gains of HEVC-ROI coding over AVC-ROI coding compared to regular coding of the entire frames including predictive modes (*inter*) depend on sequence characteristics. We present a detailed analysis of bit distribution within the frames to explain the gains. In total we can provide coding data rates of 0.7–1.0 Mbit/s for full HDTV video sequences at 30 fps at reasonable quality of more than 37 dB.**

***Index Terms*—Region of Interest (ROI) Video Coding, HEVC, Global Motion Compensation (GMC), Moving Object Detection, UAV Attached Moving Camera, Aerial Surveillance**

I. INTRODUCTION

In aerial surveillance applications from *Unmanned Aerial Vehicles* (UAVs) a small encoded video data rate is as important as a high quality and resolution of the observed area. *Region of Interest* (ROI) coding is a common solution for reducing the coding bit rate at the cost of certain image areas which are considered to be less important (*i. e.* the background, non-ROI) than others (*i. e.* the foreground, ROI) [1]. One challenge in an UAV mounted system is to classify ROI and non-ROI fully automatically in order to assign quality levels and bit rates for different image areas. Finally, a coding scheme is needed which allows to assign different quality levels within one frame. The video coding system in [2] avoids distinguishing different quality levels retaining full HDTV ground resolution at a data rate of 1–3 Mbit/s. This coding system relies on global motion compensation of the background and encoding and transmission of *New Areas* (NAs) contained in the current frame but not in the previous one. To retain also *local* movement in the decoded video, *Moving Objects* (MOs) are encoded and transmitted additionally. Those two types of ROIs are automatically detected by special ROI detectors, one for NAs and one for MOs. By the modular design of this system it is possible to include additional ROI detectors like shape based detectors [3] or replace the ROI-MO detector *e. g.* with a motion vector based MO detector [4]. The video coder itself consists of an externally controlled *Advanced Video Coding* (AVC [5]) *x264* encoder, which sets any non-ROI area to *skip mode* and thus introduces no extra transmission cost by preserving standard compliance for the bit stream. Alternative ROI coding systems propose the variation of the QP for ROI/non-ROI areas on a macroblock/*Coding Unit* (CU) level, respectively [6], which unintentionally introduces lots of extra



(a) Detection of new areas by global motion compensation (GMC).

(b) Detection of moving objects (diff. image-based).

(c) Transmission of macroblocks/CTUs containing ROI.

Figure 1: Illustration of ROI detection and coding.

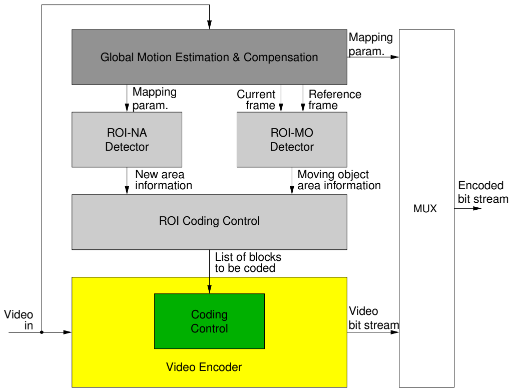


Figure 2: Block diagram of GME/GMC-based ROI coding system. Gray: unmodified (dark: GME/GMC, light: ROI detection), yellow/green: external controlled video encoder (based on [2]).

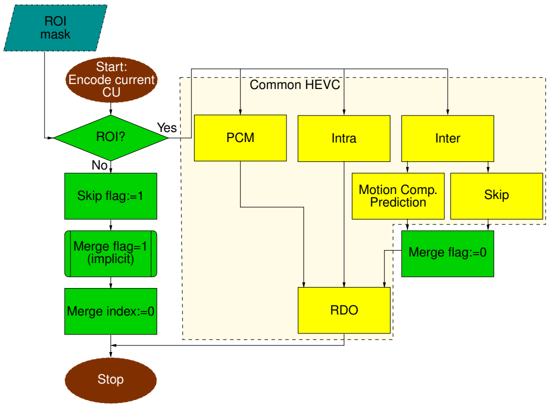


Figure 3: Flowchart of the HEVC-skip coding system. Yellow: common HM, green: proposed modifications, blue/top left: externally provided ROI mask, brown/ellipses: start/stop.

transmission cost for signaling of the QP changes for numerous non-connected ROIs [7]. Other ROI coding schemes replace the *Rate-Distortion Optimization* within the HEVC encoder in order to assign a different amount of bits to ROI and non-ROI [8], [9]. However, when employing a global motion compensation postprocessing, all data from non-ROI area is discarded anyway at the decoder and thus the optimal bit allocation scheme obviously is to spend as much bits as possible on ROI and as few bits as possible on non-ROI areas. This constraint can best be fulfilled by employing the skip mode like in the reference system [2] why we decided for a skip-implementation in the HEVC reference software HM 10.0 similar to the AVC-based coding back-end.

In this paper we propose the replacement of the video encoder by an externally controlled *High Efficiency Video Coding* (HEVC [10]) encoder [11]. We demonstrate an efficient mode control including the *skip mode* and the mandatory HEVC syntax elements *merge flag* and *merge index*. Moreover we present a detailed analysis of the spatial bit distribution.

The remaining paper is organized as follows: Section II summarizes the ROI coding system shortly, and explains the encoding process in detail. The proposed replacement of the coding back-end towards HEVC and implementation details are given in Section III. Experimental results are discussed in Section IV before Section V concludes the paper.

II. ROI-BASED REFERENCE CODING SYSTEM

Although the ROI detection system remains unchanged compared to the reference [2] like afore-mentioned we summarize the system before we focus on the (AVC-based) coding back-end and the proposed upgrade to HEVC in Section III.

The idea of data reduction is to exploit the special characteristic of a *planar* landscape of aerial surveillance videos which is true for high flight altitudes (Fig. 1). Assuming a planar landscape, one frame $k-1$ can be projected into the consecutive frame k employing a projective transform with 8 parameters $\vec{a}_k = (a_{1,k}, a_{2,k}, \dots, a_{8,k})^T$. The pixel coordinates from the preceding frame $\vec{p}_{k-1} = (x_{k-1}, y_{k-1})$ are mapped to the position $\vec{p}_k = (x_k, y_k)$ of the current one with the mapping parameter set $\vec{a}_k$ (1).

$$F(\vec{p}, \vec{a}_k) = \frac{a_{1,k} \cdot x_{k-1} + a_{2,k} \cdot y_{k-1} + a_{3,k}}{a_{7,k} \cdot x_{k-1} + a_{8,k} \cdot y_{k-1} + 1}, \frac{a_{4,k} \cdot x_{k-1} + a_{5,k} \cdot y_{k-1} + a_{6,k}}{a_{7,k} \cdot x_{k-1} + a_{8,k} \cdot y_{k-1} + 1} \quad (1)$$

To determine $\vec{a}_k$, first, a global motion estimation is performed. To do so, *Harris Corners* [12] are used to defined a set of good-to-track feature points in the frames k. A *Kanade-Lucas-Tomasi* (KLT) [13], [14] feature tracker is employed afterwards to relocate the feature positions in the frame $k-1$ and thereby generate a sparse optical flow between the frames. Outliers such as false tracks are removed and the final mapping parameter set $\vec{a}_k$ is determined by *Random Sample Consensus* (RANSAC) [15]. This mapping parameter set is used for the *Global Motion Compensation* (GMC) as the first block in the block diagram of the coding system (Fig. 2) by employing Equation (1). The mapping parameter set $\vec{a}_k$ is further employed to determine the *New Area* (NA) in the current image k by the *ROI-NA Detector*. In order to detect local motion, the pel-wise difference image between the global motion compensated frame $\hat{k}$ and the current frame k is calculated and spots of high energy are considered to be moving objects (Fig. 1b). Both ROIs are passed to the *ROI Coding Control* block which basically assigns the pel-wise ROI to the corresponding macroblocks for AVC coding (Fig. 1c). Any ROI macroblock is AVC encoded as usual whereas any non-ROI macroblock is forced to *skip mode*. Thus, the data rate is significantly reduced while standard compliance of the bit stream is retained. The mapping parameter set has to be transmitted in the data stream as well which could be realized by encapsulating the 8 parameters per frame in *Supplemental Enhancement Information* (SEI) messages. However, after decoding of the bit stream a postprocessing is necessary in order to align ROIs from the current frame within the reconstructed background from the previous frames [1].

III. PROPOSED VIDEO ENCODER IMPLEMENTATION

To incorporate the increased coding performance of HEVC compared to AVC [16] we transfer an external skip mode control similar to the AVC implementation ("AVC-skip") into HEVC ("HEVC-skip") and replace the video encoder in the ROI coding system (Fig. 2) [11]. We distinguish two cases again: ROI and non-ROI. Since we are not interested in any content of non-ROI CUs as explained in the last section, we force to use *skip mode* regardless of any *Rate-Distortion Optimization* (RDO) assuming that there cannot be any other mode (*i. e.* PCM/intra/inter prediction) which saves more bits than skip mode. By contrast, *Coding Units* (CU) containing ROI are encoded as usual by HM. Since the skip mode in HEVC implies the *merge mode* as mandatory, allowing the inheritance of motion vectors from spatially or temporally neighboring prediction units [17], the merge mode has to be controlled as well. It has two syntax elements: the binary *merge flag* and an integer *merge index* indicating the rate-distortion optimized best motion predictor from a candidate list for the current CU. The merge flag only has to be transmitted for non-skip modes whereas the merge index has to be transmitted for every skipped (and merged) block. To retain standard compliance of the bit stream while minimizing the coding cost for a skipped CU we force the merge mode a constant value (zero) for non-ROI blocks (Fig. 3, left/green column) in order to reduce the bit rate after CABAC encoding. For ROI blocks we perform a normal RDO with the only difference that for skip mode the merge mode is disabled completely (merge flag set to zero) to prevent prediction with a non-ROI CU. The flow diagram is depicted in Fig. 3. Since the ROI mask relies on 16×16 pel macroblocks in AVC-skip, we propose to restrict the reference software HM to *Coding Tree Units* (CTUs, formerly *Largest Coding Units*, LCUs) of 16×16 pels and a maximum partition depth of 2 resulting in smallest CUs of 4×4 pels. Coding results for bigger CTUs and higher partition depths (down to 4×4-CUs) are additionally presented for HEVC/HEVC-skip for comparison.

IV. EXPERIMENTAL RESULTS

The same detection results like for the AVC-skip encoder are provided as input for the HEVC-skip encoder ("ROI mask" in Fig. 3) and the coding performance of the AVC-skip (modified *x264*, v0.78 [20]) and the HEVC-skip video encoder (based on HM 10.0 [21], modifications according to Section III, *low delay* (LD) based configuration [18] with modified CTU size/maximum partition depth) are compared directly. As a reference also the unmodified HEVC encoder (HM 10.0) is compared to the unmodified x264 (v0.78) AVC encoder Table I. We used 4 self-recorded HDTV (1920×1080, 30 fps, consumer camera with global shutter) aerial video sequences from suburban areas from different flight heights (350 m, 500, 1000, 1500 m, Fig. 4) resulting in corresponding ground resolutions of 43, 30, 15, 10 pel/m (*TNT Aerial Video Testset* (TAVT), [19]). The test sequences have different characteristics such as varying amount of ROIs due to various sizes of ROI-NA and changing numbers of moving objects like pedestrians and cars. For the highest flight altitude the video is relatively noisy due to growing dusk.

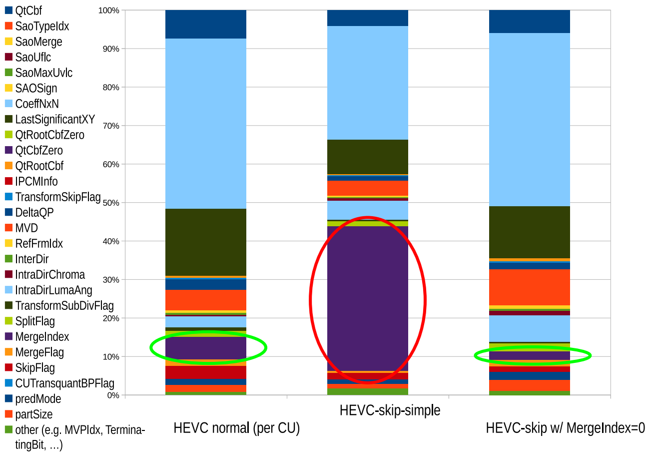


Figure 5: Comparison of relative data rate consumption of syntax elements in the HEVC bit streams for normal HEVC/HEVC-skip-simple/HEVC-skip w/ merge index handling (per CU).

For a reliable data rate comparison we measured the (luminance) PSNR only for ROI blocks for all "skip-implementations" assuming that the background quality stays constant due to the postprocessing (including GMC). The QP for the HEVC implementations were altered to match the bit rate of the AVC-skip implementation with $\text{QP} = 25$ as close as possible.

Coding results for the different test sequences are provided in Table I. It is obvious that the average coding gain of 32 % (and also 38 % for CTUs of size 64×64) is lower than literature references [16], since only small parts of each frame (typically 5–20 %) are actually encoded in non-skip modes (all ROI areas) and consequently are available for inter prediction. Additionally the coding efficiency is limited by forcing smaller block sizes than allowed by the standard [10]. Coding results for bigger block sizes (32×32 and 64×64) are also presented in Table I for comparison. *NS-CTU* or *non-splitted CTUs* means that CTUs containing any ROI-CU are not splitted but entirely encoded in non-skip modes, leading to unnecessary encoded (non-ROI) areas. As a consequence the coding performance is decreased compared to *HM-subskip* (CTUs containing ROI may be further splitted in skipped/non-skipped CUs). Consequently for *NS-CTU* implementations the smallest (external enforced) skip block is equal to the CTU size whereas it is 16×16 for the *HM-subskip* implementation. We also tested a predictive encoder configuration based on the HEVC *Random Access* (RA) configuration with hierarchical coding which performs similar to the LD configuration. For an *All Intra* (AI) configuration the relative gain is fairly constant at approximately 25 % but of course at a notable higher total bit rate. It is salient that the coding gain of HEVC-skip (16×16 CTUs) over AVC-skip is also constant (about 32 %, Table I, bold numbers) whereas the gains of unmodified HEVC over AVC vary in a wide range from 11.9–30.1 %, which can be assumed as typical considering different sequence characteristics (*e. g.* noise) [22] and the reduced CTU

Table I: Coding gains (negative numbers) of proposed HEVC-based over AVC-based ROI coding system compared to the reference (*Ref.*) as marked in the table column by column. AVC and HEVC coding data rates without ROI coding are additionally given (LD configuration based [18] with modified CTU size/maximum partition depth). Coding results for bigger block sizes are given for HEVC. *NS*: *non-splitted* CTUs: CTUs containing any ROI are always *entirely* encoded in a non-skip mode, *HM-subskip*: only those (small) CUs containing ROI are encoded in a non-skip mode, the remaining CUs containing non-ROI are encoded in the highly efficient skip mode.

		350 m sequence 43 pel/m, 821 frames PSNR ≈ 38.9 dB			**500 m sequence 30 pel/m, 1121 frames PSNR ≈ 37.2 dB**			**1000 m sequence 15 pel/m, 1166 frames PSNR ≈ 37.7 dB**			**1500 m sequence 10 pel/m, 1571 frames PSNR ≈ 37.6 dB**		
Coder	CTU in pel	Data rate in kbit/s	Diff. in %	Diff. in %	Data rate in kbit/s	Diff. in %	Diff. in %	Data rate in kbit/s	Diff. in %	Diff. in %	Data rate in kbit/s	Diff. in %	Diff. in %
AVC (x264)	16 × 16	9287	*Ref.*	—	11491	*Ref.*	—	9420	*Ref.*	—	13560	*Ref.*	—
HEVC (LD)	16 × 16	6489	−30.1	—	8973	−21.9	—	7243	−23.1	—	11942	−11.9	—
HEVC (LD)	64 × 64	5568	−40.0	—	7947	−30.8	—	5849	−37.9	—	11901	−12.2	—
AVC-skip	16 × 16	943	−89.8	*Ref.*	1423	−87.6	*Ref.*	1153	−87.8	*Ref.*	967	−92.9	*Ref.*
HEVC-skip	16 × 16	**634**	**−93.2**	**−32.8**	**938**	**−91.8**	**−34.1**	**797**	**−91.5**	**−30.9**	**664**	**−95.1**	**−31.3**
HEVC-skip (NS)	32 × 32	659	−92.9	−30.1	987	−91.4	−30.6	872	−24.4	−90.7	836	−93.8	−13.6
HEVC-skip (NS)	64 × 64	829	−91.1	−12.1	1338	−88.4	+42.6	1172	−87.6	+1.7	1335	−90.2	+38.1
HEVC-subskip	64 × 64	559	−94.0	−40.7	853	−92.6	−40.1	743	−92.1	−35.6	616	−95.5	−36.2

(a) *350 m sequence*, 43 pel/m.

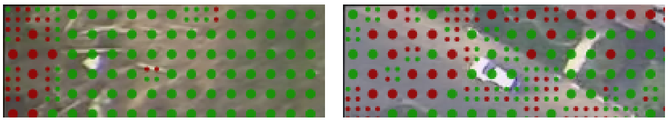
(b) *500 m sequence*, 30 pel/m.

(c) *1000 m sequence*, 15 pel/m.

(d) *1500 m sequence*, 10 pel/m.

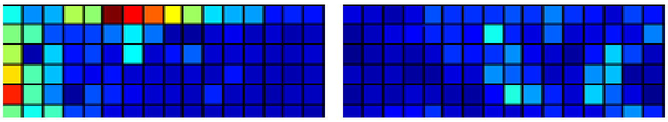
Figure 4: Example frames from the test sequences with flight height and ground resolution [19].

size. Whereas the coding gains of the unmodified HEVC are up to 30 % for sequences containing very little noise (*e. g.* in the 350 m sequence) we only gain about 12 % for a noisy and highly textured sequence (1500 m sequence). The ROI areas mainly contain new content (NA is located on the left side for the test frame from the 350 m sequence, and on the left and top side for the 1500 m sequence) which is predominantly intra coded anyway (Fig. 6, note also the high amount of intra coded blocks (red dots) in non-ROI for the 1500 m sequence in Fig. 6b). Those ROI areas consume a high amount of bits which can be seen in the bit distribution maps in Fig. 7, especially for the 350 m sequence. Blue colors within these “heat maps” correspond to low bit usage for an CTU whereas red colors indicate high bit usage. The gain of HEVC-skip over AVC-skip is much higher than the gain of HEVC over AVC for the 1500 m sequence than for the 350 m sequence. In order to predict the coding efficiency gain of aerial video sequences, we analyze the sequence characteristics. Therefore we define the cost of coding individual blocks. With the *ROI-bit-ratio* C (2) and the *ROI-area-ratio* A (3) we define the *bit-distribution-ratio*



(a) *350 m sequence, ROI left.* (b) *1500 m sequence*, ROI left and top.

Figure 6: Prediction modes of HEVC (red dots: intra, green: inter, outtakes, ROI-NA left in (a) and left/top in (b)).



(a) *350 m sequence*, $R = 3$. (b) *1500 m sequence*, $R = 1.5$.

Figure 7: Bit usage of example frames ("heat map", outtakes).

R according to (4).

$$C = \frac{\text{ROI bit cost}}{\text{total bit cost}} \tag{2}$$

$$A = \frac{\text{ROI area}}{\text{frame area}} \tag{3}$$

$$R = \frac{C}{A} \tag{4}$$

The difference in relative coding gains from HEVC over AVC compared to HEVC-skip over AVC-skip depends on very diverse ratios of R for different sequences meaning that the bit usage for ROIs drastically differs from the corresponding *areas* covered by those ROIs.

If R is ≈ 1, the ROI bit cost is proportional to the area covered by ROI (*e. g.* for the 1500 m sequence with $R = 1.5$, Fig. 7b). If $R \gg 1$ the ROI bit costs are unproportional high for ROI areas, *i. e.* a huge amount of bits consumed by one frame is used to encode only a small part of the frame (which is true for the other sequences with $3 < R < 4$). When such a frame is encoded with HEVC-skip, the gain is much higher compared to the gain of HEVC over AVC like for this test set. Consequently we can use the bit-distribution-ratio R as an indicator for the HEVC-skip coding gain relative to the unmodified HEVC gain. It is noteworthy that the encoding runtime decreases approximately linear with the number of blocks to be coded. Thus, the encoding time of HEVC-skip is decreased by 80–95 % compared to unmodified HM for our test set. Despite additional processing time needed for global motion estimation and ROI detection the entire detection & coding system is much faster than HM.

V. CONCLUSIONS

In this paper we propose to replace the AVC video encoder by HEVC in a *Region of Interest* (ROI)-based coding system for aerial surveillance videos with a moving camera, *e. g.* attached to an UAV. The coding system relies on an external control of the video encoder by ROI detectors. Only ROI areas are regularly encoded whereas non-ROI areas are forced to *skip* mode. We present an efficient mode control for HEVC and can gain 32 % on average over an AVC-skip implementation at similar coding block size and up to 38 % for bigger coding block sizes (CTU size of 64×64) which corresponds to coding data rates of 0.7–1.0 Mbit/s at more than 37 dB (ROI-PSNR) for full HDTV (30 fps) aerial surveillance video. We provide a detailed analysis of spatial bit distribution of inter frames for the HEVC encoder HM and derive a bit-distribution-ratio as an indicator for the achievable coding gains of the proposed HEVC-skip video encoder. Results show highest relative gains of HEVC-skip over AVC-skip compared with HEVC over AVC for noisy and highly textured sequences.

REFERENCES

[1] H. Meuel, J. Schmidt, M. Munderloh, and J. Ostermann, *Adv. Vid. Cod. for Next-Generation Multimed. Services – Chpt. 3: ROI Coding for Aerial Video Seq. Using Landscape Models*. Intech, Jan. 2013. [Online]. Available: http://tinyurl.com/ntx7u29

[2] H. Meuel, M. Munderloh, and J. Ostermann, "Low Bit Rate ROI Based Video Coding for HDTV Aerial Surveillance Video Sequences," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition - Workshops (CVPRW)*, June 2011, pp. 13 –20.

[3] S. Wang, Y. Fu, K. Xing, and X. Han, "A Method of Target Recognition from Remote Sensing Images," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ Internat. Conf. on*, Oct 2009, pp. 3665–3670.

[4] H. Sabirin and M. Kim, "Mov. Object Detect. and Tracking Using a Spatio-Temporal Graph in H.264/AVC Bitstreams f. Video Surveillance," *IEEE Transact.on Multimedia*, vol. 14, no. 3, pp. 657–668, June 2012.

[5] AVC, *Rec. ITU-T H.264 & ISO/IEC 14496-10 MPEG-4 Part 10: Adv.Video Coding (AVC)-3rd Ed.* Geneva, Switzerland: ISO/IEC&ITU-T, Jul. 2004.

[6] Y. Liu, Z. G. Li, and Y. C. Soh, "Region-of-Interest Based Resource Allocation for Conversational Video Communication of H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 1, pp. 134–139, Jan 2008.

[7] H. Meuel, J. Schmidt, M. Munderloh, and J. Ostermann, "Analysis of Coding Tools and Improvement of Text Readability for Screen Content," in *Picture Coding Symposium (PCS)*, May 2012, pp. 469–472.

[8] X. Deng, M. Xu, and Z. Wang, "A ROI-based Bit Allocation Scheme for HEVC Towards Perceptual Conversational Video Coding," in *Advanced Computational Intelligence(ICACI),6.Int.Conf.*, Oct 2013, pp. 206–211.

[9] M. Xu, X. Deng, S. Li, and Z. Wang, "Region-of-Interest Based Conversational HEVC Coding w/ Hierarch. Perception Model of Face," *Sel.Top.in Signal Proc., IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2014.

[10] HEVC, "ITU-T Recommendation H.265/ ISO/IEC 23008-2:2013 MPEG-H Part 2/: High Efficiency Video Coding (HEVC)," 2013.

[11] H. Meuel, M. Munderloh, M. Reso, and J. Ostermann, "Mesh-based Piecewise Planar Motion Compensation and Optical Flow Clustering for ROI Coding," in *APSIPA Transactions on Signal and Information Processing*, 2015.

[12] C. Harris and M. Stephens, "A Combined Corner and Edge Detection," in *Proceedings of The Fourth Alvey Vision Conf.*, 1988, pp. 147–151.

[13] C. Tomasi and T. Kanade, "Detection and Tracking of Point Features," Carnegie Mellon Univ., Tech. Rep. CMU-CS-91-132, April 1991.

[14] J. Shi and C. Tomasi, "Good Features to Track," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Seattle, June 1994.

[15] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, June 1981.

[16] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards - Including High Efficiency Video Coding (HEVC)," *Circ. and Systems f. Video Technology, IEEE Transact. on*, vol. 22, no. 12, pp. 1669 –1684, Dec. 2012.

[17] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.

[18] F. Bossen, “L1100: Common HM Test Conditions and Software Reference Configurations. JCT-VC of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11. Geneva, CH, 14-23 Jan,” 2013.

[19] Institut für Informationsverarbeitung (TNT), Leibniz Universität Hannover. (2010–2014) TNT Aerial Video Testset (TAVT). [Online]. Available: https://www.tnt.uni-hannover.de/project/TNT_Aerial_Video_Testset/

[20] VideoLAN Organization. (2009) x264. [Online]. Available: http://www.videolan.org/developers/x264.html

[21] I.-K. Kim, K. McCann, K. Sugimoto, B. Bross, and W.-J. Han, “High Effic. Video Coding (HEVC) Test Model 10 (HM10) Encoder Description,” in *JCT-VC Doc. JCTVC-L1002*, Geneva, Switzerland, Jan. 2013.

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