Low Bit Rate ROI Based Video Coding for HDTV Aerial Surveillance Video Sequences

Holger Meuel, Marco Munderloh, Jörn Ostermann

Institut für Informationsverarbeitung (TNT), Leibniz Universität Hannover, Germany

{meuel, munderl, ostermann}@tnt.uni-hannover.de

Abstract

For aerial surveillance systems two key features are important. First they have to provide as much resolution as possible, while they secondly should make the video available at a ground station as soon as possible. Recently so called Unmanned Aerial Vehicles *(UAVs) got in the focus for surveillance operations with operation targets such as environmental and disaster area monitoring as well as military surveillance. Common transmission channels for UAVs are only available with small bandwidths of a few Mbit/s. In this paper we propose a video codec which is able to provide full HDTV (1920 × 1080 pel) resolution with a bit rate of about 1–3 Mbit/s including moving objects (instead of 8–15 Mbit/s when using the standardized AVC codec). The coding system is based on an AVC video codec which is controlled by ROI detectors. Furthermore we make use of additional* Global Motion Compensation *(GMC). In a modular concept different* Region of Interest *(ROI) detectors can be added to adjust the coding system to special operation targets. This paper presents a coding system with two motion-based ROI detectors; one for* new area detection *(ROI-NA) and another for* moving objects *(ROI-MO). Our system preserves more details than an AVC coder at the same bit rate of 1.0 Mbit/s for the entire frame.*

1. Introduction

The Pulse Code Modulation (PCM) data rate of an 8 bit electro-optical (EO) color video image signal with full HDTV (*High Definition Television*) resolution of 1920 by 1080 picture elements (pel) per frame and 25 frames per second amounts a data rate of 622 Mbit/s, already considering the common 4:2:0 chroma subsampling. For data reduction motion compensating hybrid coders have been standardized by ISO/MPEG (ISO: International Standardization Organization, MPEG: Moving Picture Experts Group) [8, 10, 9]. These coders split each frame into blocks which are subsequently coded by a motion compensating transform coding technique. Using the current AVC (Advanced Video Coding) standard [11], the PCM data rate of a color HDTV signal can be cut down to about 8–15 Mbit/s without visible amplitude distortions or loss of spatial resolution. When using a standard compliant AVC coder at a lower bit rate, this would lead to quality scaling or loss of resolution.

For transmission of airborne video sequences from *Unmanned Aerial Vehicles* (UAVs) (or aeroplanes) to a receiving ground station the video data rate has to be further reduced in order to meet the available relatively low channel capacity of a few Mbit/s. Nevertheless the highest possible spatial resolution is necessary for reconnaissance and no degradation of resolution of the transmitted video signal can be accepted. There has been a lot of research about the processing of aerial surveillance video considering *Global Motion Estimation* (GME) [7], *Global Motion Compensation* (GMC) [12] and *Region of Interest* (ROI) detection, whereas the most important applications were mosaicking [16, 17] and moving object detection [13]. In [2] the field of view of a human observer is modeled to predict ROI in aerial surveillance video sequences. The impact on the required video data rate however remains unevaluated. In common video coding for entertainment video or video conferencing applications the term ROI often refers to human faces or moving objects in general. [3] even trains neural networks to recognize ROIs, especially to separate foreground (like faces and bodies) and background. It is apparent, that ROI coding has never been employed to further reduce the bit rate of aerial surveillance video sequences preserving the full resolution *and* preserving moving objects. Therefore, we present an approach based on a modularized *Region of Interest* (ROI) coding scheme in order to achieve additional bit rate reduction while preserving the image quality for the entire frame as well as the movement of moving objects at the receiver side. With the proposed technique multiple ROI detectors can be operated in parallel. Our coding system consists of a global motion estimation and compensation unit, two ROI detectors, one for *new areas* (ROI-NA) and one for *moving objects* (ROI-MO), a ROI coding control unit and a modified AVC encoder.
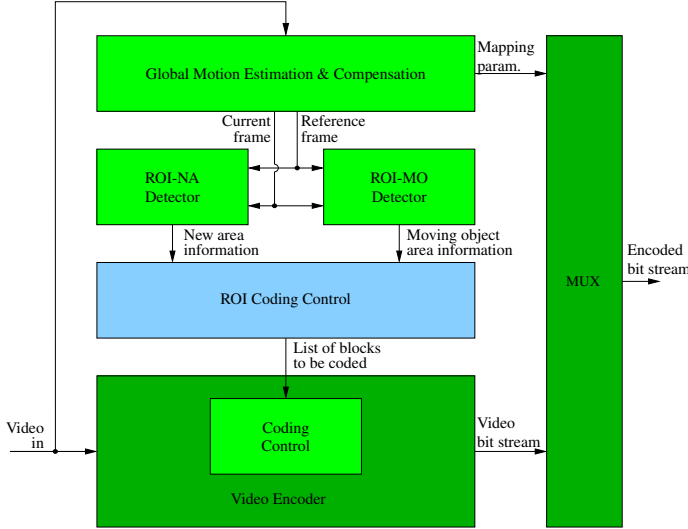
Figure 1. Scheme of ROI based coding control

The remainder of this paper is organized as follows. Section 2 presents the developed ROI coding concept in detail, whereas Section 2.1 describes the process of feature extraction and estimation of the camera movement, Section 2.2 explains the new area detector and Section 2.3 deals with the moving object detector. Our method for integrating a new area and moving object detector in a complete coding system is presented in Section 2.4. Section 2.5 contains different possibilities for decoding and visualizing of the ROI encoded and decoded video stream. Section 3 shows experimental results of our work and illustrates the advantages of our proposed concept in terms of coding efficiency. Finally, Section 4 concludes the paper.

2. ROI based coding concept

In order to make the integration of new video codecs or additional ROI detectors easier, a modular approach has been developed (Figure 1). The global motion estimation is a fundamental step in our processing chain because it enables motion-based ROI detectors. From the tracked features, mapping parameters can be derived employing a projective transform. Assuming a planar surface and using the mentioned transform, we can cope with arbitrary translational movements, rotations and scalings. For high 3D objects, like buildings, our assumption is violated. However, since projection failures occur only at the borders of high objects (like buildings), we can neglect the visual influence for reconnaissance activities. The projection parameters can be used by all further modules [1], while global motion parameters are applied to compensate motion.

When comparing the global motion compensated reference frame with the current frame, new areas can easily be

calculated to be encoded and moving objects can be detected [13]. The pel-wise information of new areas and moving objects are mapped to corresponding macroblocks by our ROI coding control unit and the result is passed to a modified AVC encoder.

Global motion estimation/compensation as well as the detection of new areas and moving objects are well known in the literature [15, 14, 16, 17, 13]. However, a concatenation of these different techniques for the improvement of the coding efficiency has not been proposed yet. Therefore, we present a complete system and introduce these well known techniques into a new application area.

2.1. Global Motion Estimation

To estimate the global motion in a frame of a video sequence it is necessary to know about the movement of the camera. The camera is fixed at the aeroplane, therefore both move the same way. For an accurate tracking process the movement of the camera has to be estimated using the video sequence, given the accuracy limitations of GPS and drift problems of INS [16]. We use a tracking approach based on fixed points on the background.

Assuming the ground is planar, it is possible to transform one frame into another using 8 parameters (projective transformation). The projection describes for every pel x and y in frame $k-1$ a matching pel $\vec{p} = (x, y)$ in the succeeding frame k with the mapping parameter set $\vec{a_k}$. The feature tracker [14] detects corners in a frames using the *Harris Corner Detector* [5]. This detector is based on a two-dimensional gradient method which uses the luminance (grey values) within the picture. Features are defined as corners with high gradients in horizontal and vertical direction.

Afterwards, a correspondence analysis is performed, employing the KLT (Kanade-Lucas-Tomasi) feature tracker [15, 14]. Based on a local optical flow method the position of all features from frame $k-1$ are aligned in the consecutive frame k.

The alignment is used to estimate the global motion of the camera. Therefore, an overdetermined linear equation system is set up for estimation of the 8 parameters of the projective transform. To improve the accuracy of its error and to reduce the movement estimation error, RANSAC (Random Sample Consensus) [4] is used. By optimizing the Mean Squared Error (MSE) it is detecting outliers. The resulting projection parameters are used to align two frames.

The mapping parameters attained from the tracking process are used for global motion compensation. The motion compensated reference frame as well as the current frame are passed to the following ROI detectors.

2.2. ROI-NA detector for new areas

The new area detector uses the characteristic of airborne video sequences that the movement of the camera is
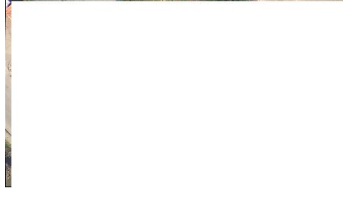
Figure 2. *New area coding*: Frame content which has to be transmitted for one HDTV frame (test sequence from 500 m flight height)
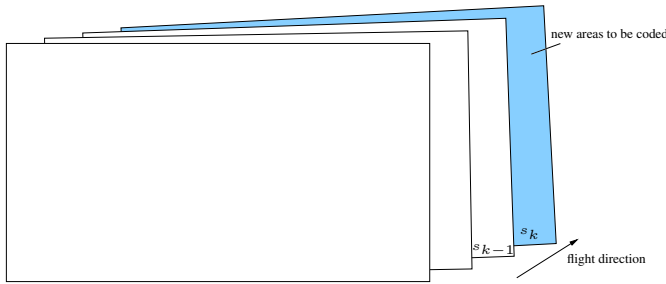

Figure 3. Detection of new areas by global motion compensation (GMC)

very small while taking two consecutive images at typical sampling rates (*e.g.* 25 Hz) and typical velocities of UAVs ($\approx 100\,\mathrm{km/h}$). For a flight height of 350 m it amounts up to 48 pel. Thus, most of the content of one frame is contained in the previous frame as well and does not have to be transmitted again. Figure 2 illustrates the small amount of data which has to be transmitted for one HDTV frame.

To determine new areas in a frame, two consecutive frames of a video sequence are aligned (Figure 3). Therefore, the ROI-NA detector gets the current frame and the motion compensated [12] reference frame as input (Figure 1).

Only picture elements from the current frame s_k, which are projected outside the previous frame s_{k-1}, are detected as new area (ROI-NA) and need to be encoded, transmitted and decoded. However, a ROI-NA is not feasible to present any moving object at decoder side.

2.3. ROI-MO detector for moving objects

The ROI-MO detector is used to enable reconstruction of moving objects as such. Therefore moving objects and their uncovered background are detected on board the UAV as well. The ROI-MO detector also gets the current frame and the motion compensated [12] reference frame as input (cf. Figure 1). A difference picture between these two frames is derived in order to detect moving objects and uncovered background. To reduce noise the average of a 3×3 block is calculated and values below a predefined threshold t_1 are considered as noise. If a region is larger than a predefined minimum m, a moving object is registered.

It is obvious that the ROI-NA as well as the ROI-MO de-

tector provide useful information. Thus, it is desirable to combine both detector results for reconnaissance.

2.4. ROI Coding Control

As shown above ROI-NA coding can efficiently encode static aerial video sequences and ROI-MO based coding is able to preserve moving objects. To maintain the image quality in the entire frame *and* to be able to present moving objects at the receiver side, we propose a ROI coding control unit. This unit manages which blocks are encoded by the video encoder. It is based on information from the ROI-NA and ROI-MO detectors (cf. Figure 1). Therefore, the entire frame is divided into macroblocks of size 16×16 pel. The first frame of a video sequence has to be encoded completely as a reference frame. In every succeeding frame only macroblocks containing pels of new areas *or* of moving objects are passed from the ROI coding control unit to the video encoder. Macroblocks which are not marked as ROI for encoding will be encoded by the video encoder in *skip mode*. This special mode causes a macroblock to remain unchanged until new information for this macroblock is encoded [11].

The resulting bit stream first contains all macroblocks of the reference frame. Further intra-coded reference frames can be inserted in pre-defined intervals to allow for random access and error resilience. As long as they are not updated either with ROI-NA or with ROI-MO information, every macroblock keeps its original information from the reference frame. Due to the missing global motion compensation, the video decoded by an AVC standard compliant player is not useful. For example, if an UAV is moving with constant velocity, new areas are detected in every frame at the same position. Thus, only macroblocks at the appropriate border(s) are updated with new area information at the decoder. Any other macroblocks remain like they were in the reference frame. If also moving objects have been detected, these moving objects containing macroblocks are inserted at the detected positions. Since the motion of the background cannot be compensated by a standard AVC player, the macroblocks containing moving objects replace macroblocks from the reference frame, *e.g.* cars are driving where no street has been in the reference frame. Hence, a ROI encoded video decoded by an AVC standard compliant player needs further processing to become useful.

2.5. Decoding and visualization of a ROI controlled video sequence

With our presented approach, a bit stream is created which is compatible to the AVC standard. Since global motion compensation is not a feature of AVC, the movement of macroblocks representing the background is not considered. However, to decode ROI encoded data properly, a special ROI decoder had to be developed, which is able to deal correctly

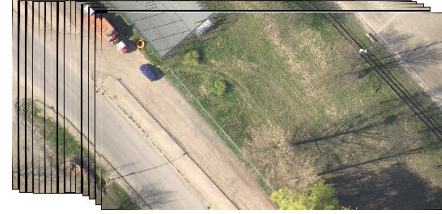Figure 4. Generation of the decoded video sequence out of *new area* (sequence from 500 m flight height).

Figure 5. Principle of creating a mosaic. Black marker lines between single frames are for illustration purpose only (test sequence recorded at 500 m flight height).

with global motion compensation. In principal a video or a mosaic is created from the ROI-NA encoded data and afterwards ROI-MO blocks are inserted in the resulting video sequence or the mosaic at appropriate positions. There are several possibilities to visualize a ROI encoded and decoded video sequence at the receiver.

1. Generation of a mosaic from ROI-NA data

2. Generation of a video sequence from ROI-NA data

3. Generation of a video sequence from combined ROI-NA and ROI-MO data

In every case it is necessary to transmit the mapping parameters as side-information to the receiver in order to apply a GMC also at the decoder. This can be done without modification of the standardized bit-stream when encapsulating the mapping parameters as SEI (*Supplemental Enhancement Information*) in the data stream [11]. Also information about the position of moving objects containing macroblocks has to be transmitted.

2.5.1 Generation of a mosaic from ROI-NA data

A mosaic of the landscape can be generated out of ROI-NA data at the receiver side. An observer gets a general overview of the surveillance area and is also provided with detailed information. Therefore, the decoder must be provided with the mapping parameters which have to be transmitted as side-information. To create a complete mosaic, every macroblock has to be registered to its final position in memory in the mosaic according to the mapping parameters. Thus, only one filtering is needed per stripe of new area and multiple filtering as in a frame-by-frame based motion compensation can be avoided. The principle of memory management can be found in [6]. To determine the global position for a macroblock in the mosaic we need to concatenate global motion parameters back to the last reference frame. Since the final dimensions of the mosaic are not known a priori for streaming applications, the memory to store the mosaic in has to be extended dynamically with every stripe of new area. Figure 4 shows the principle of stitching stripes of ROI-NA together.

Figure 5 gives a closer look at the growing process of the mosaic. It shows some succeeding stripes of new area

which are stitched beside the reference frame. The black marker lines between the single frames are only for illustration purposes.

A complete mosaic is shown in Figure 6, with the real level of details shown in magnification. The original mosaic has 21104×4500 pel, which corresponds to about 30 seconds of flight in a flight height of 350 m.

Moving objects are displayed statically at the position of their first occurrence in the new area, since no additional information about their movement is available. Thus, they are processed like any other pixel of static objects. Also the noise remains frozen.

2.5.2 Generation of a video sequence from ROI-NA data

The receiver can generate a decoded video sequence from ROI-NA encoded data only. The obvious way to generate a video sequence out of stripes of new area is to compensate the global motion frame by frame and to stitch the new area to the appropriate position according to the mapping parameters.

Since this generation process of the video implies multiple filtering, the generation of a video sequence is more difficult. A mosaic has to be created in memory like described above. From this mosaic an outtake with the dimensions of the transmitted video can be cut out frame by frame and concatenated as a video sequence. The correct position for the outtake also can be determined employing the mapping parameters.

Moving objects are displayed statically at the position of their first occurrence.

2.5.3 Generation of a video sequence from combined ROI-NA and ROI-MO data

Previous ROI codecs [16, 13] provide high quality Regions of Interest while in non-ROI areas image degradation and loss of resolution can be observed. With our proposed combined approach of ROI-NA and ROI-MO this insufficiency can be corrected.

In principal, it is possible to play the decoded video sequence with a standard compliant AVC video player. How-

Figure 6. Mosaic and magnification (sequence from 350m flight height)

ever, because the movement of macroblocks representing the background is not considered due to missing global motion compensation in AVC, it is not useful and will lead to a video sequence where only ROI-NA and ROI-MO macroblocks are updated. Thus, the frames of the resulting video are the input to the decoder GMC process.

It is obvious that there is need of a special ROI decoder which is capable of global motion compensation. A complete ROI decoder is depicted in Figure 7. In its center the intelligent memory for the mosaic – which can produce the frame cut from the mosaic in memory required for the current frame – can be found. The mapping parameters are employed to calculate the correct frame cut. Macroblocks containing either *new area* or *moving objects* are inserted in the combination unit from the current frame from the video decoder. The positions of the macroblocks containing moving objects has to be available at the decoder as additional side-information. For the succeeding frame the memory is updated using the next mapping parameters and the current new area. The result is displayed as a video sequence where moving objects are displayed as moving objects. Since the frame-rate remains the same as in the recorded video sequence, no additional filtering for smoothing the movement is necessary.

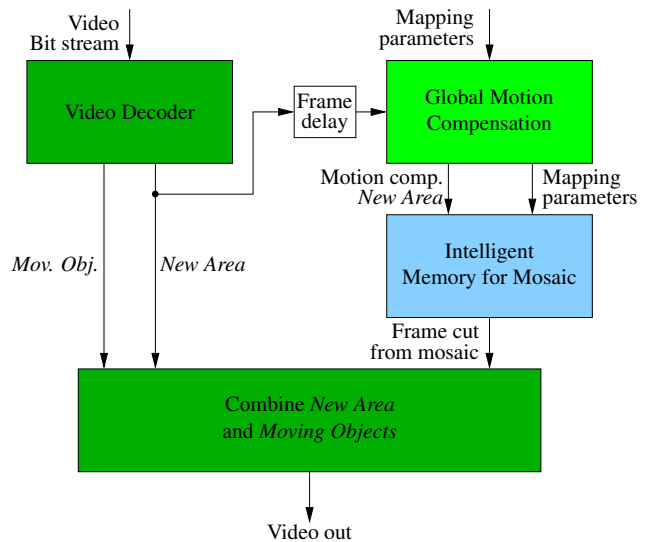A decoded frame from the ROI decoder is shown in Figure 11a.

Figure 7. Proposed ROI decoder

3. Experimental Results

Both of the introduced ROI based codings (new area coding as well as moving object based coding) lead to a bit rate of 0.8–2.5 Mbit/s each which is much less than the bit rate needed for detail preserving regular AVC video coding for HDTV signals. However, it is not possible to use a PSNR (*peak-signal-to-noise ratio*) based quality comparison directly because the output of the ROI coder is not pel-wise the same as the input especially at object borders due to image
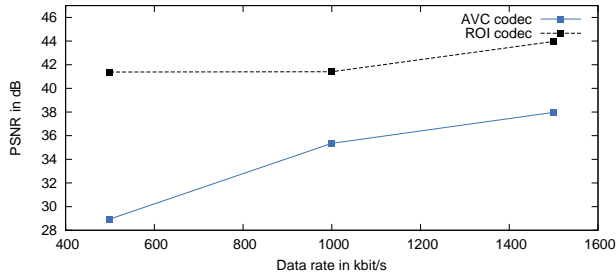
Figure 8. Rate distortion diagram of AVC coding in comparison with ROI-NA and ROI-MO areas at low bit rates.

degradations which are immanent for a new area ROI coding. Even though the full spatial resolution can be preserved there are remaining degradations in image quality due to model violations. To estimate the resulting PSNR in ROI processed video sequence nevertheless, we compare the PSNR of those macroblocks containing ROI pels with the recorded sequence (Figure 8). It can be noticed that our approach outperforms a standard AVC coder by far in terms of better PSNR at any bit rate, since – apart from header and signaling bits for encoding *skip mode* – more bits are available to encode a very small amount of macroblocks (compared to the entire frame, cf. Figure 2). Only macroblocks containing *new area* or *moving objects* have to be encoded with a (modified) AVC coder. Additional bit rate reduction is achieved because noise remains frozen in the the new area stripes since only translational movement is applied in succeeding frames. However, since image degradations remain, the rate distortion diagram can be seen only as a coarse estimation of the true image quality.

The projective transform assumes a planar landscape for transforming one frame into another. This assumption fits for most of the observed landscape. However, for objects with high total height like buildings, this assumption is not valid any more. These limitations are *only* affecting objects with high total height like buildings or trees, hence they are negligible for reconnaissance. Figure 9 illustrates this effect (especially at the borderlines of the rooftop). Since a frame in full HDTV resolution is reconstructed at the receiver side out of small ROI-NA areas, there will remain some perspective distortions at borders of high objects.

A property of any 2D mosaic occurs when flying over a three dimensional object. The observer gets different views of the same object from different angles, see Figure 10. During the ROI detection and encoding process on-board the aeroplane, only the first view is considered. In the reconstructed video sequence at the receiver, there are systematic differences because of these different views. Hence, an observer is not capable of viewing things *behind* high objects. When employing a standardized video codec, the view of maximal three different sides of a static object (one

Figure 9. Systematic errors in the reconstruction due to the influence of different angles of view during the flight over the house. The *new area* in this outtake can be found above the white marker line. It is recorded at another time and from another angle of view than the remaining frame. The red circles highlight the offset at the borderlines of the rooftop (sequence from 350m flight height).

Figure 10. Same building from different angles of view.

while approaching, one while passing by and one while the object is directly below/behind the observation camera onboard an UAV) are available at the receiver side, while with our proposed coding scheme only two sides of the same object are visible, since non-moving objects are global motion compensated. However, for this loss of comfort the gain in terms of coding efficiency is high.

In our experiments, a value of $t_1 = 255$ for the noise reduction filter lead to good results. The resulting image quality after ROI coding and decoding is shown in Figure 11. With fixed $QP = 30$ (*Quality Parameter*) we got a resulting bit rate of about 1000 kbit/s and an estimated PSNR based on the ROI areas of about 41.4 dB. To achieve a similar PSNR with a standard AVC coder, a bit rate of about 15.5 Mbit/s would be needed, Figure 11a shows a complete frame which was ROI encoded with our proposed coder and decoded at the receiver side.

Figure 11b is a magnified outtake of the original frame as it was recorded by the camcorder. A standardized AVC coding at a bit rate of slightly greater than 1000 kbit/s leads to the loss of details like shown in the outtake in Figure 11c. Our proposed ROI codec is able to provide a higher grade of details in the entire frame, as can be seen in the magnified outtake in Figure 11d. The effect is clearly visible *e.g.* in the tree and at the man holes on the right. The AVC encoder of the ROI codec (Figure 11d) is operated with the same parameters as the reference AVC coder (Figure 11c) except for the quantizer.

In conclusion, a ROI based coding is able to reduce the data rate of an electro-optical aerial HDTV color video sequences from 622 Mbit/s to about 0.8–2.5 Mbit/s for typical scenarios. The compression gain of the combined ROI coding is smaller because additional data for moving objects has to
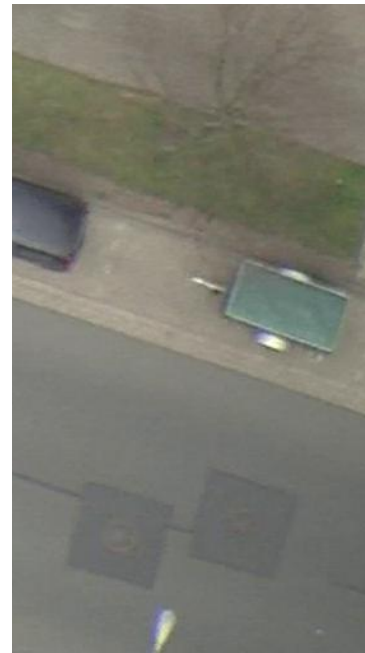
(a) Complete ROI encoded and decoded frame
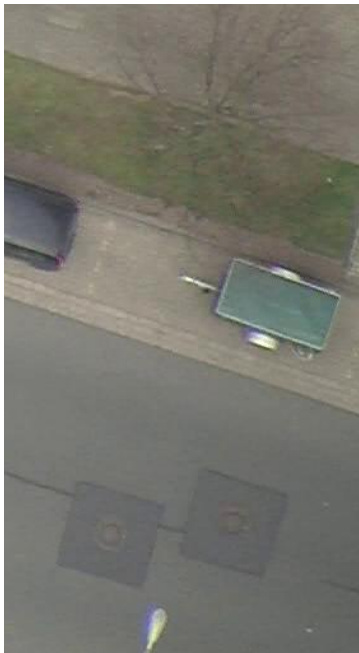
(b) Outtake in original resolution (like recorded).
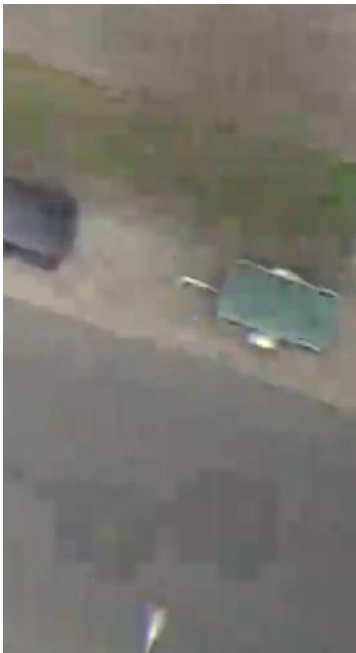
(c) Outtake: H.264 encoded (> 1000 kbit/s). Clearly visible are block artifacts. Due to insufficient bit rate the low-pass characteristic from the interpolation filter disturbs details like tree-structures and the man wholes.

(d) Outtake: ROI encoded and decoded frame at a bit rate of about 1000 kbit/s.

Figure 11. Visualization of ROI encoded and decoded data. One ROI processed full frame is shown in Figure 11a. Figures 11b–11d show magnified outtakes for quality comparison, whereas Figure 11b shows the original quality from the camcorder, Figure 11c was encoded with a standardized H.264 coding at slightly more than 1000 kbit/s and Figure 11d shows the same outtake processed with our proposed ROI based coding at a bit rate of about 1000 kbit/s (test sequence from 350 m flight height).

be transmitted. The exact bit rate (at a fixed quality) needed for a ROI coded video sequence depends on the number of moving objects in a frame. In a worst case scenario every macroblock contains moving objects and has to be encoded. The bit rate then equals the AVC bit rate of 8–15 Mbit/s since a (modified) AVC coder has been used. However, for typical scenarios a bit rate of 1–3 Mbit/s can be achieved at high image quality with full resolution.

4. Conclusion

In this paper we present an approach to reduce the data rate of electro-optical aerial HDTV color video sequences from 622 Mbit/s PCM data rate to about 1–3 Mbit/s at subjectively good quality for reconnaissance purposes. We achieve a PSNR of more than 40 dB for all ROI areas while an AVC coder at the same bit rate achieves a PSNR of less than 34.5 dB. Global motion estimation is employed and mapping parameters are derived frame-by-frame. We suggest to operate different ROI codings like a *new area* detection and a *moving object* detection in parallel on-board of an UAV. Both detectors make use of the mapping parameters. Further, we introduce a ROI coding control unit which controls an AVC encoder. We obtain an AVC compliant bit stream which is created on-board the UAV. For proper decoding of such ROI encoded streams we implemented an ROI decoder which is able to handle the necessary global motion compensation. It is based on the creation of a mosaic in memory to avoid multiple filtering.

References

[1] H. Broszio. *Schätzung von Brennweite und Rotation einer Kamera mit Radialverzerrung und Rotationsachsenversatz aus Bildsequenzen*. PhD thesis, Universität Hannover, Jan. 2007. Dissertation. 2

[2] H. Cheng and J. Wus. Adaptive region of interest estimation for aerial surveillance video. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 3, pages III – 860–3, sept. 2005. 1

[3] N. Doulamis, A. Doulamis, D. Kalogeras, and S. Kollias. Low bit-rate coding of image sequences using adaptive regions of interest. *Circuits and Systems for Video Technology, IEEE Transactions on*, 8(8):928 –934, dec 1998. 1

[4] 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*, 24(6):381–395, June 1981. 2

[5] C. Harris and M. Stephens. A Combined Corner and Edge Detection. In *Proceedings of The Fourth Alvey Vision Conference*, pages 147–151, 1988. 2

[6] M. Hötter. *Objektoriente Analyse-Synthese-Codierung basierend auf dem Modell bewegter, zweidimensionaler Objekte*. PhD thesis, Universität Hannover, Jun. 1992. Dissertation. 4

[7] M. Hötter and R. Thoma. Image segmentation based on object oriented mapping parameter estimation. *Signal Processing*, 15(3):315 – 334, 1988. Multidimensional Signal Processing. 1

[8] ISO/IEC. *ISO/IEC 11172-2 (MPEG-1 Part 2): Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 2: Video*. Aug. 1993. 1

[9] ISO/IEC. *ISO/IEC 14496:2000-2: Information technology - Coding of Audio-Visual Objects - Part 2: Visual*. Dec. 2000. 1

[10] ISO/IEC and ITU-T. *Recommendation ITU-T H.263 and ISO/IEC 13818-2 (MPEG-2 Part 2): Information technology - Generic coding of moving pictures and associated audio information: Video*. Mar. 1995. 1

[11] ISO/IEC and ITU-T. *Recommendation ITU-T H.264 and ISO/IEC 14496-10 (MPEG-4 Part 10): Advanced Video Coding (AVC) - 3rd Edition*. Geneva, Switzerland, July 2004. 1, 3, 4

[12] A. Krutz, A. Glantz, and T. Sikora. Recent advances in video coding using static background models. In *Picture Coding Symposium (PCS), 2010*, pages 462 –465, dec. 2010. 1, 3

[13] R. Mech and M. Wollborn. A noise robust method for 2d shape estimation of moving objects in video sequences considering a moving camera. *Signal Processing*, 66(2):203 – 217, 1998. 1, 2, 4

[14] J. Shi and C. Tomasi. Good Features to Track. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, Seattle, June 1994. 2

[15] C. Tomasi and T. Kanade. Detection and Tracking of Point Features. Technical Report CMU-CS-91-132, Carnegie Mellon University, April 1991. 2

[16] S. Yahyanejad, D. Wischounig-Strucl, M. Quaritsch, and B. Rinner. Incremental mosaicking of images from autonomous, small-scale uavs. In *Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on*, pages 329 –336, 29 2010-sept. 1 2010. 1, 2, 4

[17] Z. Zhu and H. Tang. Content-based dynamic 3d mosaics. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on*, page 169, june 2006. 1, 2