H.264 Video Coding for UAV Applications improved by Sensor Fusion

CENTRO ITALIANO RICERCHE AEROSPAZIALI

Angelino C.V., Baraniello V.R., Cicala L., De Mizio M., Leoncini P.
CIRA, the Italian Aerospace Research Center, Italy

Baccaglini E., Gavelli M., Raimondo N., Scopigno R.
Istituto Superiore «Mario Boella», Italy

MORE SENSORS

LESS PROCESSING

ABSTRACT

This work presents a new low-complexity **H.264** encoder for **Unmanned Aerial Vehicles** (UAV) applications.

Standard video coding systems currently employed in UAV applications do not rely on some peculiarities in terms of scene 3D model and correlation among successive frames.

In particular, the observed scene is static, i.e. the camera movement is dominant, and it can often be well approximated with a plane. Moreover, **camera position and orientation** can be obtained from the **navigation system**.

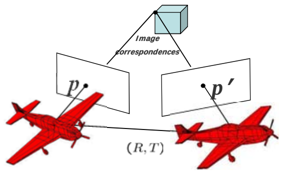
Therefore, correspondent points on two video frames are linked by a simple **homography**. The encoder employs a new motion estimation scheme which make use of the **global motion** information provided by the onboard navigation system. For the aim position and orientation data accuracy is further improved analyzing the optical flow and adopting a **sensor fusion** strategy.

The homography is used in order to **initialize the block matching** algorithm allowing a more robust motion estimation and a **smaller search window**, and hence reducing the complexity.

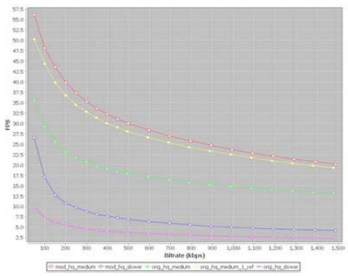
The results are relevant in **low frame rate video coding**, which is a typical scenario in UAV behind line-of-sight (BLOS) missions, in which a **satellite data link** must be used in upload. Experiments open new directions in developing new **sensor aided video coding** standards.

X264 CODE MODIFICATIONS

In the proposed work, as reported in [8], the **open source encoder x264**, has been modified in order to allow to choice the prediction method (inter/intra) for each macroblock and the prediction of motion vectors by substituting the standard x264 predictors with predictors derived from external data. The sintax of the output bitstream is still fully **compliant with the standard H.264**.

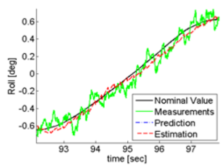


The Global Motion Estimation is performed using external position (Global Navigation Satellite System) and orientation (Attitude & Heading Reference System) sensors. The observed scene is supposed to be approximated with a plane



Encoded fps vs output bitrate for the original x264 library and the proposed sensor aided H.264 encoder

Improving Position and Attitude Estimation using vision



Improved attitude estimation using vision data

PLANAR HOMOGRAPHY ESTIMANTION

The position and attitude of the UAV is further refined using the optical flow. The optical flow is calculated by the **KLT** algorithm and the Planar Homography is estimated using a fast version of **RANSAC** (after discarding data that are very incongruent with the Homography model obtained by sensors), At the end the homography matrix is decomposed into his motion and structure parameters. These parameters are the input of the sensor fusion module.

SENSOR FUSION

The information carried by the camera is then integrated with classical data coming from the IMU (Inertial Masurement Unit) and GPS in a sensor fusion algorithm. An **Unscented Kalman Filter** [9] has been implemented, because dynamic and observation equations are non-linear in their original form.

TESTS IN MOTION IMAGERY CODING SCENARIO

We report the results for **low spatial overlapping sequences** (lower than 60%) as they allow to achieve the most significant benefits in the proposed scenario.

In the reported results, the curves labeled **orig** refer to the original x264 libraries in which the GOP structure is left to default optimization policies. Curves labeled **mod** refer to the proposed libraries which exploit data from sensors to initialize motion fields

Both **UMH** [x264 preset = **slower**] and **HEX** [x264 preset = **medium**] refinement have been tested for the original encoder.

The label **1-ref** refers to the GOP structure [I,P,P,P,...] where a I frame followed by 250 P frames.

ENCODING TIME

Concerning encoding time, it is worth noting that motion estimation is significantly **faster** when it is possible to include in the video compressor system data acquired from sensors.

It is possible to notice that there is a significant reduction in the encoding time. For a total output bitrate of 400 kbps, both for HEX and UMH refinement, a **50% speed-up** in encoding time is achieved and this value slightly when the total output rate increases.

UMH refinement requires higher computational time with respect to HEX search.

The modified version of x264 has quite the same processing time than the original version with a the very simple GOP structure (**1-ref**), but gives better quality at the same bitrate.

RD CURVES

Experimental results show that the proposed scheme outperforms standard H.264 in terms of PSNR and throughput.

The proposed algorithm outperforms traditional H.264/AVC coding because the motion estimation fails when the frame overlapping is very low, with **HEX** motion refinement (**real time applications**) the gain can be approximately **1.8 dB** when comparing the approach vs the original x264. In case of **UMH** refinement (**non real-time applications**) the gain reduces to **1.4 dB**.

SUBJECTIVE VISUAL QUALITY

The image quality comparison between the standard ITU H.264 and the proposed method. As one can see, the proposed method better preserves the image details which are very important for UAV surveillance and monitoring missions.

TESTS ON IMAGING SIMULATOR DATA & ON REAL VIDEO SEQUENCES

Recently other tests has been performed on images acquired at low altitude **by drone.** The results on real data confirm that obtained by computer graphic simulations.

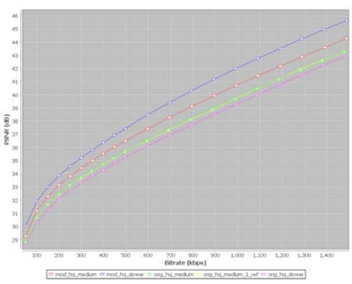
Acquisition platform used for results validation on real video sequences [9]

PROTOTYPE

A full working prototype of sensor aided video encoder has been realized at CIRA. The prototype is fully compliant with the standards **ITU H.264**, **KLV** and **MPEG2-TS,** according to the NATO Standard **STANAG 4609**. It demonstrates that as this novel technology is **short time to market**.

LOWER BITRATE

HIGHER QUALITY



Quality (PSNR) vs output bitrate for the original x264 library and the proposed sensor aided H.264 encoder

Image visual quality comparison between x264 (left) and the proposed sensor aided encoder (right)

REFERENCES

[1] R. Collin, Video Encoding using Motion Data, Nokia, international patent n. WO/2009/156965, 30-12-2009.

[2] M. Stephan, Sensor Assisted Video Compression, Harman Becker Automotive Systems GmbH, european patent n. EP 1 921 867 A1, 14-05-2008.

[3] Y. Wang, L. Zhong, A. Rahmati, G. Hong, Sensor Assisted Motion Estimation for Efficient Video Encoding Rice University, Huston (Texas), USA patent n. US 2010/0079605 A1, 01-04-2010.

[4] X. Chen, Z. Zhao, A. Rahmati, Y. Wang, L. Zhong, Sensor-Assisted Video Encoding for Mobile Devices in Real-World Environments IEEE Trans. on Circuits and Systems for Video Technology, vol. 21, no. 3, March 2011, pp. 335-349.

[5] A. F. Rodr´ıguez, B. B. Ready, and C. N. Taylor, Using telemetry data for video compression on unmanned air vehicles, AIAA Guidance, Navigation, and Control Conference and Exhibit, aug 2006.

[6] J. Gong, C. Zheng, J. Tian, and D. Wu, An image-sequence compressing algorithm based on homography transformation for unmanned aerial vehicle,in International Symposium on Intelligence Information Processing and Trusted Computing (IPTC), oct. 2010, pp. 37-40.

[7] M. Bhaskaranand, J. D. Gibson , Low-complexity Video Encoding for UAV Reconnaissance and Surveillance Proceedings of 2011 Military Communications Conference - Track 4 - Middleware Services and Applications, 2011.

[8] C. V. Angelino, L. Cicala, M. De Mizio, P. Leoncini, E. Baccaglini, M. Gavelli, N. Raimondo, R. Scopigno: "Sensor aided H.264 Video Encoder for UAV applications", Picture Coding Symposium - PCS, pp. 173-176, 2013.

[9] C. V. Angelino, V.R. Baraniello, and L. Cicala. "High altitude UAV navigation using IMU, GPS and camera", IEEE International Conference on Information Fusion - FUSION, 2013.

[10] C. V. Angelino, L. Cicala, G. Persechino, E. Baccaglini, M. Gavelli, N. Raimondo: "A sensor aided H.264 encoder tested on aerial imagery for SFM", submitted to International Conference on Image Processing, ICIP 2014.

International Computer Vision Summer School, ICVSS 2014