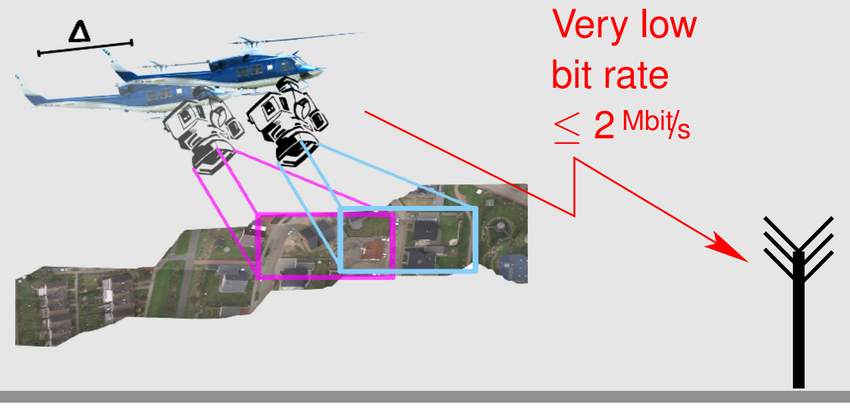
Stereo Mosaicking and 3D-Video for Singleview HDTV Aerial Sequences using a Low Bit Rate ROI Coding Framework

Holger Meuel · Marco Munderloh · Jörn Ostermann

Scenario and Goal

- Transmission system for aerial surveillance and disaster area monitoring
- Fully automatic processing on-board the aerial system
- Only one monocular camera available
- Small bandwidth channel
- High image quality requirement
- Depth information shall be retained



Stereo Region of Interest (ROI)-based ENCODING

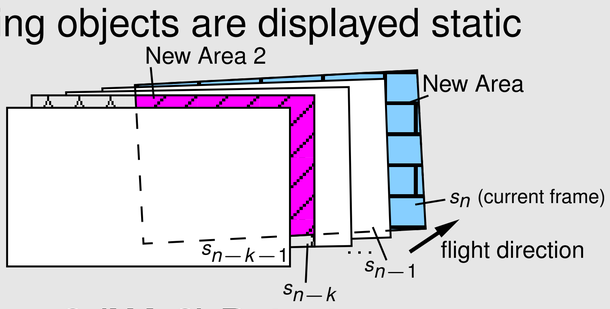
Assumption: Planar ground!

Global Motion Estimation/Compensation (GME/GMC)

- Harris corner detector + KLT feature tracking + RANSAC
- Calculation of a projective transform parameter set

New Area (NA) Detector

- Calculation of new areas based on transform parameters
- Moving objects are displayed static

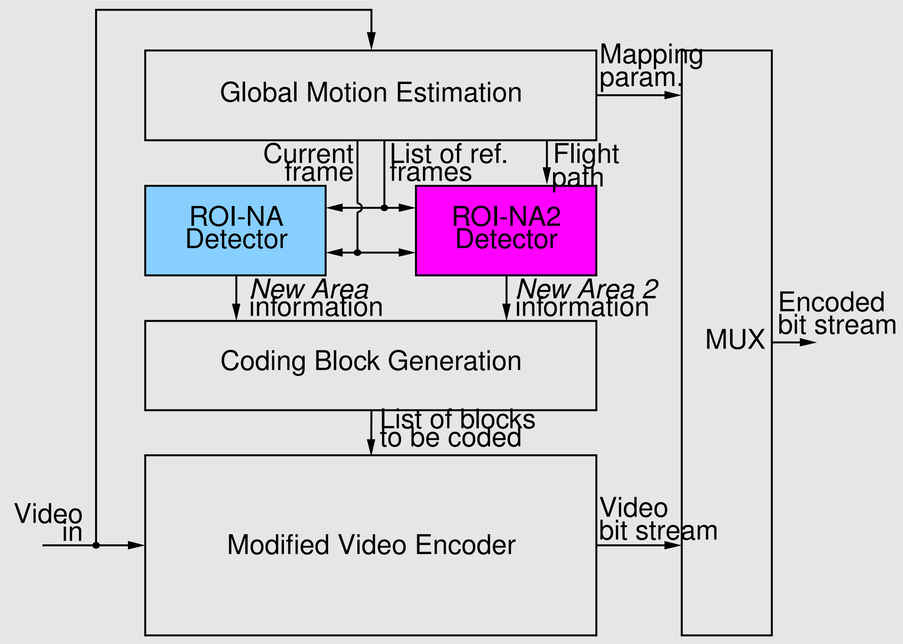


New Area 2 (NA 2) Detector

- Second view from virtual camera, based on flight path:
- Concatenate k homographies until desired displacement Δx is reached: $\mathbf{H_k} = \prod_{i=n-k}^{n} \mathbf{H_i}$
- Baseline distance controllable, i. e. depth impression adjustable
- Check if projected pixel $\mathbf{H_k}\, s_n(x, y)$ lies within frame s_{n-k}
- Check similar if projected pixel does *not* lie in frame s_{n-k-1}

Coding Block Generation

- Assign NA and NA 2 to coding units for encoding



Block diagram of a stereo ROI-based encoder (gray: unmodified compared to [1])

Modified Video Encoder

- External control of skip-mode in video encoder
- ROI=normal coding, non-ROI=forced skip mode

Calculate Motion Parallax

- Parallax expressed by flight height h, frame displacement Δx and object height C_z: $p = -\Delta x \frac{h}{(C_z - h)}$
- Parallax and subjective optimal baseline distances:

Sequence	*350 m*	*500 m*	*1000 m*	*1500 m*
max. parallax p	40	28	14	9
Frame offset k	10	15	20	30

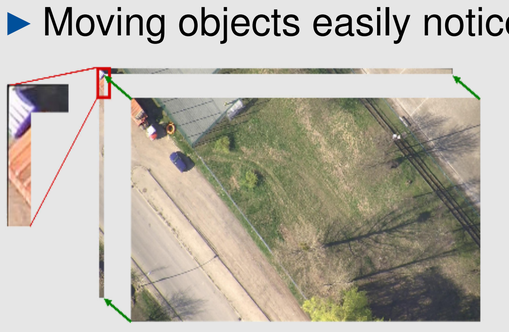
ROI-based DECODING and Stereo Video Generation

Generate Stereo Panorama Images

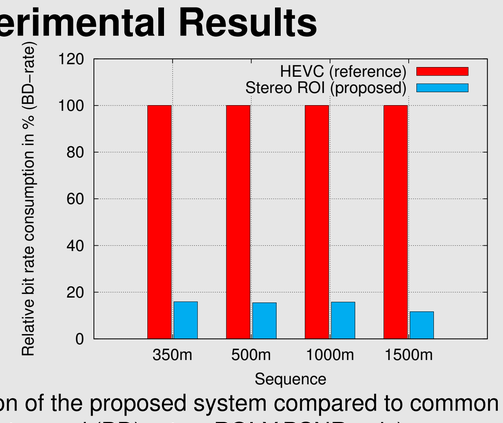
- Usage of transmitted ROI blocks only
- Two panoramas created by mosaicking, one for each view from corresponding NA
- Full ground resolution preserved

Generate a Stereo ("3D") Video

- Extract frames from mosaics, one for each view
- Combine views in "side-by-side" frames
- Depth information derivable from stereo
- Moving objects easily noticeable as parallax errors

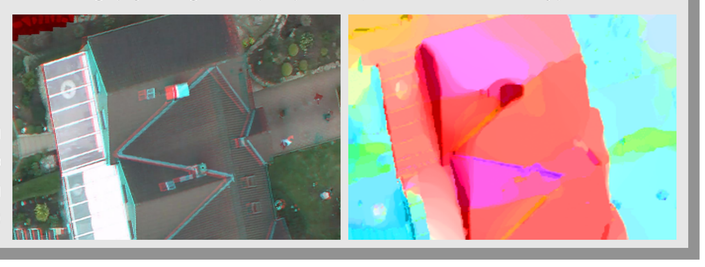


Experimental Results



Bit rate consumption of the proposed system compared to common HEVC coding (Bjøntegaard (BD)-rates, ROI-Y-PSNR only)

Red-cyan anaglyph decoded stereo image (left) and depth map (right)



Summary

- Fully automatic stereo coding system for aerial video
- Stereo/"3D" information retained
- Factor 6 data reduction to 1-2 Mbit/s for full HDTV stereo signals
- Adjustable depth impression
- Depth map generation and subsequent CV tasks possible
- Easy subsequent detection of moving objects possible

References: [1] H. Meuel, M. Munderloh, J. Ostermann: Low Bit Rate ROI Based Video Coding for HDTV Aerial Surveillance Video Sequences, IEEE CVPRW, 1st Workshop of Aerial Video Processing (WAVP), 2011

Institut für Informationsverarbeitung
Holger Meuel
meuel@tnt.uni-hannover.de
http://www.tnt.uni-hannover.de/project/roi_coding

Leibniz Universität Hannover