In-Loop Radial Distortion Compensation for Long-Term Mosaicing of Aerial Videos

Holger Meuel · Stephan Ferenz · Marco Munderloh · Hanno Ackermann · Jörn Ostermann

Scenario and Goal

Goal: Fully automatic panorama image generation from aerial videos

- Scenario: Planar, vertical aerial video from UAV affected by (radial) lens distortions
- Feature point-based homography estimation from video frames
- Automatic lens distortion correction challenging in typical approaches
- Lens distortion correction implicitely modeled
- Extension for large automatically generated panorama images

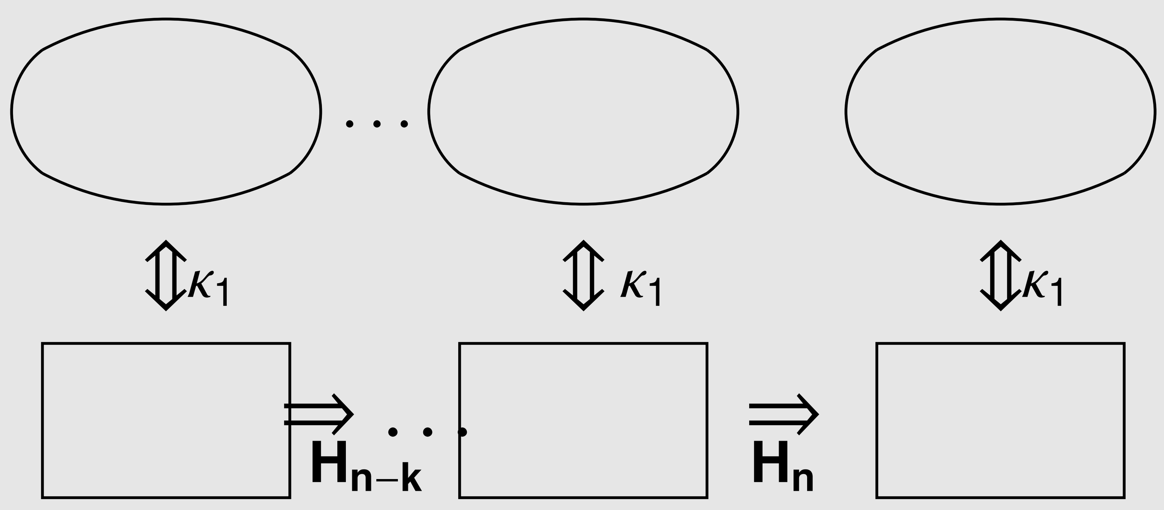
Idea: Joint estimation of homographies and radial distortion

State-of-the-art: Mosaicing fails after 330 frames for non-distortion corrected videos

Joint Homographies and Radial Distortion Estimation

- Model radial distortion as Taylor series, consider only 1st order:
  $x_u = x_d\left(1 + r_d^2\kappa_1\right) = x_d + x_d^3\kappa_1 + x_d y_d^2\kappa_1$
  $y_u = y_d\left(1 + r_d^2\kappa_1\right) = y_d + y_d^3\kappa_1 + y_d x_d^2\kappa_1$
  with radius from center of radial distortion $r_d = \sqrt{\left(x_d^2 + y_d^2\right)}$
  and κ_1 radial distortion parameter (1st order)
- Projective transform (homography) of a point $\vec{x_u} = (x_u, y_u, 1)^\top$ from one frame to a second frame in homogeneous coordinates: $\vec{x_u}' = \mathbf{H} \cdot \vec{x_u}$
- Concatenation of several homographies: $\vec{x_u}'^n = (\mathbf{H_n} \cdot \ldots \cdot \mathbf{H_1}) \cdot \vec{x_u}$
- Quasi-Newton method to numerically determine corresponding radially distorted point $\vec{x_d}'$ in dependence of undistorted point $\vec{x_u}'$

Joint estimation of $\mathbf{H_{n-k}} \ldots \mathbf{H_n}$ and unknown constant κ_1:

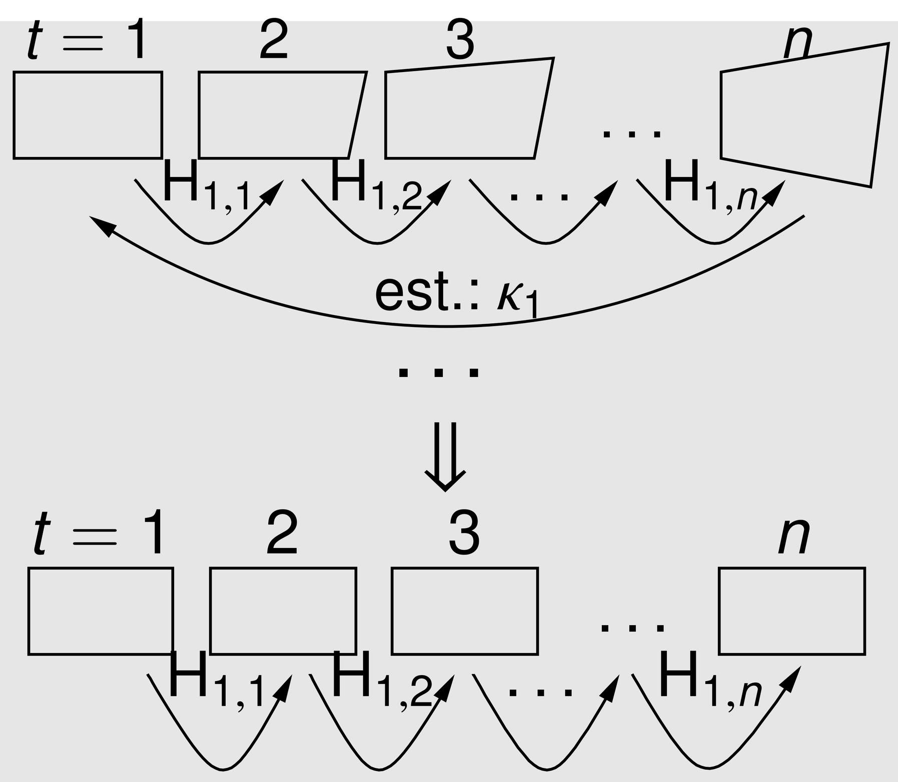


Exponential run time increase for joint estimation
⇒ Approximative solution

Fast approximative solution: In-Loop Radial Distortion Compensation

Idea: Regularize change in size and shape of projection

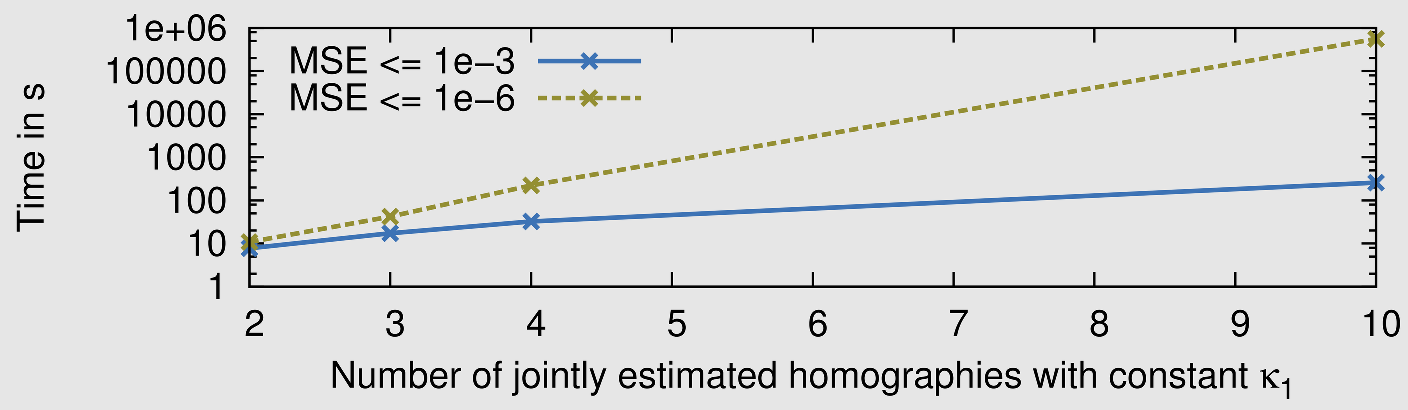
- Decompose $\mathbf{H}$ into rotation $\mathbf{R}$, translation $\vec{t}$, camera matrices $\mathbf{K}$, $\mathbf{K}'$, surface normal $\vec{n}/d$, distance between camera↔surface d: $\mathbf{H} = \mathbf{K}'\left(\mathbf{R} - \frac{\vec{t}\,\vec{n}^\top}{d}\right)\mathbf{K}^{-1}$
- Rotation matrix $\mathbf{R}$, with skew-symmetric matrices $\mathbf{W_x}$, $\mathbf{W_y}$, $\mathbf{W_z}$ induced by rotation around x- (roll), y- (pitch), z-axis (yaw): $\mathbf{R} = \exp\left(\theta_x(t)\mathbf{W_x}\right) \cdot \exp\left(\theta_y(t)\mathbf{W_y}\right) \cdot \exp\left(\theta_z(t)\mathbf{W_z}\right)$
- **Assumption**: Change in size (c_{size}) and shape (c_{shape}) of the projection of the camera target onto mosaic only depends on $\theta_x(t)$ and $\theta_y(t)$
- Iteratively optimize radial distortion parameter κ_1 in a gradient descent (over an entire picture group (PG)) to physically possible rotation changes so that: $\left|\frac{d}{dt}\theta_x(t)\right| < c_x$, $\left|\frac{d}{dt}\theta_y(t)\right| < c_y$



Experimental Results

Joint homographies & radial distortion estimation

- Noise-free, artificially generated point clouds, N=1000 points each
- Randomly sampled S=30 points for homography estimation
- Exponential run time increase
  ⇒ **Impractical for real systems**



Runtimes and MSE of Matlabs Quasi-Newton numerical solver

In-loop radial distortion compensation

- 14 iterations, 60 frames/picture group (PG)
- Limit rotations by **geometrical constraints**
- Max. change in size and shape: $c_{\text{size,max}} = 20\,\%$ and $c_{\text{shape,max}} = 10\,\%$ per PG
- 1000 ms/frame if iterating, 200 ms/frame if not (not optimized)
- 0.0044 pel/frame drift compared to Google Earth

Panorama image of (non preproc., full HDTV) TAVT data set *1000 m sequence* with 1166 frames
(TAVT available at: `https://www.tnt.uni-hannover.de/project/TNT_Aerial_Video_Testset/`)

Summary

- Joint model for estimation of several homograhies and unknown constant radial distortion
- Fast approximative solution due to exponentially increased run time for higher number of jointly estimated homographies
- **Regularization of projection for jointly estimated picture group based on geometrical constraints**
- Fully automatic mosaicing of unpreprocessed video frames
- Panorama images generated from more than 1500 uncalibrated, not preprocessed video frames

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

Leibniz Universität Hannover