Vision-Aided Navigation Based on Three-View Geometry

Vadim Indelman, Pini Gurfil

Distributed Space Systems Lab,

Aerospace Engineering, Technion

Ehud Rivlin

Computer Science, Technion

Hector Rotstein

RAFAEL – Advanced Defense Systems

February 2010

Contents

- Introduction
- Three-view geometry constraints
- Fusion with navigation system
- Results
- Summary

Introduction

- Pure inertial navigation solution diverges with time due to error integration
- Inertial solution may be compensated to eliminate or at least reduce errors, using e.g.:
  - GPS
  - DTM
  - Vision-based methods (VBM)
- GPS may be\become unavailable or unreliable

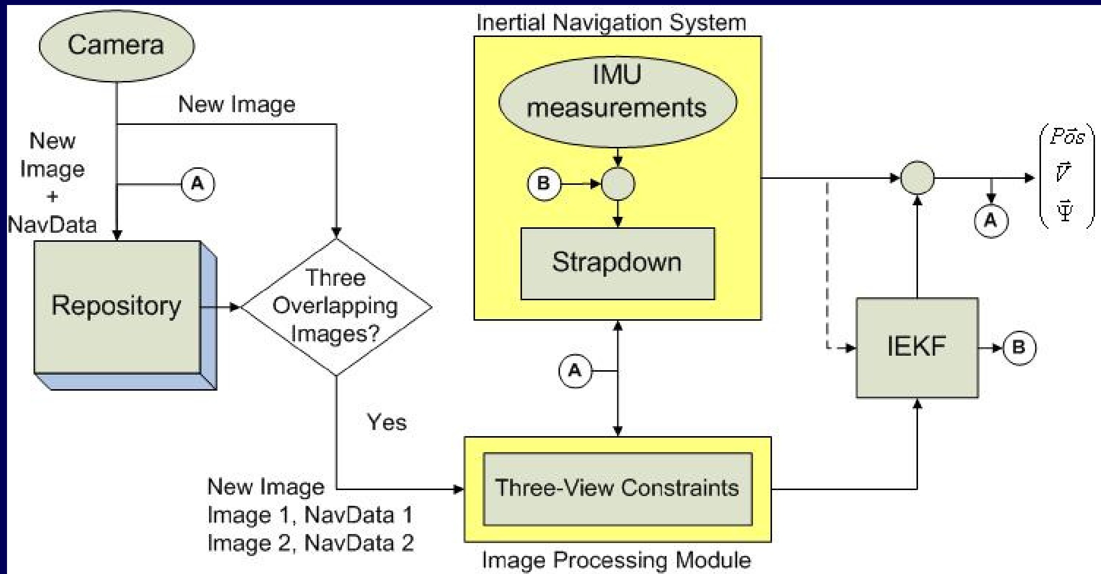
- Objective:
  - Position estimation assuming **<u>only</u>** INS and a <u>single</u> onboard camera
    - Recover scale ambiguity
  - Real-time navigation aiding
    - Handling loop scenarios

Previous Work

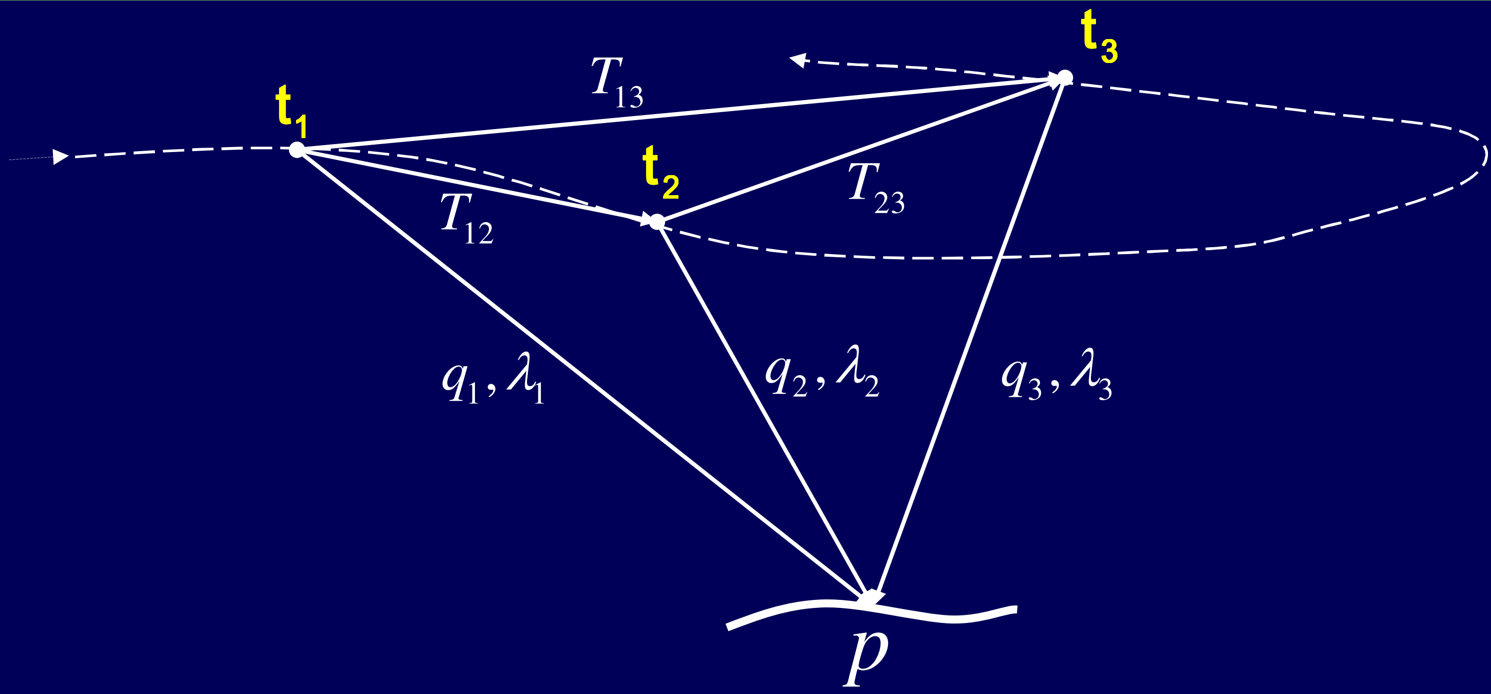
- SLAM: "6DoF SLAM aided GNSS/INS Navigation in GNSS Denied and Unknown Environments", Kim J. and Sukkarieh S., 2005
- Multiple-view + Bundle adjustment: "A Dual-Layer Estimator Architecture for Long-term Localization", Mourikis A.I. and Roumeliotis S.I., 2008
- Trifocal tensor + Motion Estimation: "Recursive Camera-Motion Estimation With the Trifocal Tensor", Yu Y.K., et. al., 2006
- Enhanced-two-view + Constant-size state vector: "Real-Time Mosaic-Aided Aerial Navigation", Indelman V., et. al., 2009

Introduction (Cont.)

- Approach
  - Decouple navigation aiding and environment representation construction (e.g. mosaic) – in contrast to SLAM
    - Constant-size state vector
    - Store and manage captured images (and some nav. data)
  - Three-view geometry
  - Avoid structure reconstruction
- Assumptions
  - INS
  - Camera



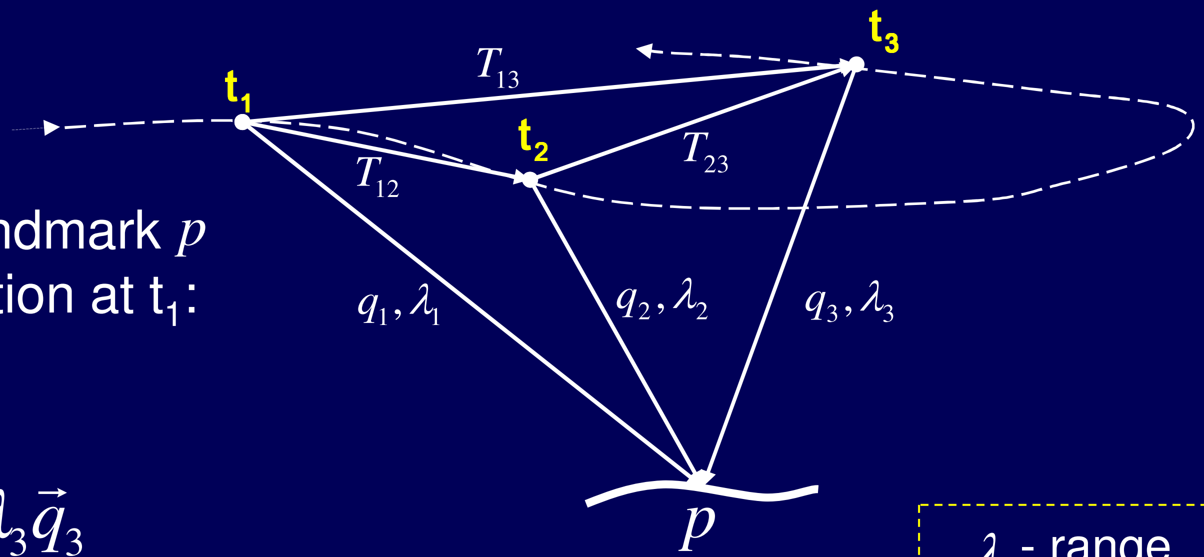
Three View Geometry



- p - ground landmark
- λ - range
- q - line of sight (LOS)
- T_{ij} - translation from i to j

Three View Geometry (cont.)

- Coordinate systems
  - L - Local Level Local North (LLLN)
  - C - Camera



- Position of a ground landmark p relative to camera position at t_1:

$$\lambda_1 \vec{q}_1 = \vec{T}_{12} + \lambda_2 \vec{q}_2$$

$$\lambda_1 \vec{q}_1 = \vec{T}_{12} + \vec{T}_{23} + \lambda_3 \vec{q}_3$$

λ - range
q - LOS

- Expressing in LLLN system at t_2:

$$\lambda_1 C_{L_2}^{C_1} \vec{q}_1^{C_1} = C_{L_2}^{C_1} \vec{T}_{12}^{C_1} + \lambda_2 C_{L_2}^{C_2} \vec{q}_2^{C_2}$$

$$\lambda_1 C_{L_2}^{C_1} \vec{q}_1^{C_1} = C_{L_2}^{C_1} \vec{T}_{12}^{C_1} + C_{L_2}^{C_2} \vec{T}_{23}^{C_2} + \lambda_3 C_{L_2}^{C_3} \vec{q}_3^{C_3}$$

Three View Geometry (cont.)

- Matrix formulation

$$\underbrace{\begin{bmatrix} \vec{q}_1 & -\vec{q}_2 & \vec{0}_{3\times1} & -\vec{T}_{12} \\ \vec{0}_{3\times1} & \vec{q}_2 & -\vec{q}_3 & -\vec{T}_{23} \end{bmatrix}_{6\times4}}_{A} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ 1 \end{bmatrix}_{4\times1} = \vec{0}_{6\times1}$$

- Note:
  - Range parameters are unknown
  - LOS and translation vectors are expressed in LLLN of t_2

$$\rightarrow \quad rank(A) < 4$$

- Denote

  $\tilde{A}$ - a matrix comprised of **<u>any</u>** 4 rows of A

$$\rightarrow \quad \det(\tilde{A}) = 0$$

Three View Geometry (cont.)

- The following constraints were obtained for a single ground landmark:

$$\vec{q}_1^{\,T} \cdot \left(\vec{T}_{12} \times \vec{q}_2 \right) = 0$$

$$\vec{q}_2^{\,T} \cdot \left(\vec{T}_{23} \times \vec{q}_3 \right) = 0$$

$$\left(\vec{q}_2 \times \vec{q}_1 \right)^T \cdot \left(\vec{q}_3 \times \vec{T}_{23} \right) = \left(\vec{q}_1 \times \vec{T}_{12} \right)^T \cdot \left(\vec{q}_3 \times \vec{q}_2 \right)$$

- First two equations – epipolar constraints
- Third equation – relates between the magnitudes of $\vec{T}_{12}$ and $\vec{T}_{23}$

$$\left[\vec{m}^T \right]_{1\times 3} \vec{T}_{12} = 0$$

$$\left[\vec{d}^T \right]_{1\times 3} \vec{T}_{23} = 0$$

$$\left[\vec{b}^T \right]_{1\times 3} \vec{T}_{23} = \left[\vec{c}^T \right]_{1\times 3} \vec{T}_{12}$$

where

$$\vec{b} = \vec{b}\left(\vec{q}_1, \vec{q}_2, \vec{q}_3 \right)$$

$$\vec{c} = \vec{c}\left(\vec{q}_1, \vec{q}_2, \vec{q}_3 \right)$$

$$\vec{d} = \vec{d}\left(\vec{q}_2, \vec{q}_3 \right)$$

$$\vec{m} = \vec{m}\left(\vec{q}_1, \vec{q}_2 \right)$$

Three View Geometry (cont.)

- Multiple features
  - Matching pairs between 1st and 2nd view $\left\{ \vec{q}_{1_i}^{\,C_1}, \vec{q}_{2_i}^{\,C_2} \right\}_{i=1}^{N_{12}}$
  - Matching pairs between 2nd and 3rd view $\left\{ \vec{q}_{2_i}^{\,C_2}, \vec{q}_{3_i}^{\,C_3} \right\}_{i=1}^{N_{23}}$
  - Matching triplets between the three views $\left\{ \vec{q}_{1_i}^{\,C_1}, \vec{q}_{2_i}^{\,C_2}, \vec{q}_{3_i}^{\,C_3} \right\}_{i=1}^{N_{123}}$

$$\left[\vec{b}_i^T \right]_{1\times 3} \vec{T}_{23} = \left[\vec{c}_i^T \right]_{1\times 3} \vec{T}_{12} \qquad i = 1, \ldots, N_{123}$$

$$\left[\vec{d}_j^T \right]_{1\times 3} \vec{T}_{23} = 0 \qquad j = 1, \ldots, N_{23}$$

$$\left[\vec{m}_r^T \right]_{1\times 3} \vec{T}_{12} = 0 \qquad r = 1, \ldots, N_{12}$$

$$\Rightarrow \boxed{\begin{bmatrix} B \\ D \\ 0 \end{bmatrix}_{N\times 3} \vec{T}_{23} = \begin{bmatrix} C \\ 0 \\ M \end{bmatrix}_{N\times 3} \vec{T}_{12} \qquad N = N_{123} + N_{12} + N_{23}}$$

Fusion with Navigation

- Implicit Extended Kalman Filter (IEKF) formulation

$$\vec{z} \equiv \begin{bmatrix} B \\ D \\ 0 \end{bmatrix}_{N\times 3} \vec{T}_{23} - \begin{bmatrix} C \\ 0 \\ M \end{bmatrix}_{N\times 3} \vec{T}_{12}$$

- Recall
  - All original LOS vectors are expressed in camera system of the appropriate view
  - $\vec{T}_{23}, \vec{T}_{12}$ are functions of $Pos(t_3), Pos(t_2), Pos(t_1)$
  - t_3 is the current time

$$\vec{z} = h\left(Pos(t_3), \Psi(t_3), Pos(t_2), \Psi(t_2), Pos(t_1), \Psi(t_1), \left\{ \vec{q}_{1_i}^{C_1}, \vec{q}_{2_i}^{C_2}, \vec{q}_{3_i}^{C_3} \right\} \right)$$

Fusion with Navigation (cont.)

- State vector definition

$$\vec{X} = \begin{bmatrix} \Delta\vec{P}^T & \Delta\vec{V}^T & \Delta\vec{\Psi}^T & \vec{d}^T & \vec{b}^T \end{bmatrix}^T \in \Re^{15\times1}$$

- Continuous system matrix

$$\Phi_c = \begin{bmatrix} 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & A_s & 0_{3\times3} & C_L^B \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & -C_L^B & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \end{bmatrix} \in \Re^{15\times15}$$

  - A_s - a skew-matrix constructed based on accelerometer sensors readings
  - C_L^B - DCM from Body to Local Level Local North systems

Fusion with Navigation (cont.)

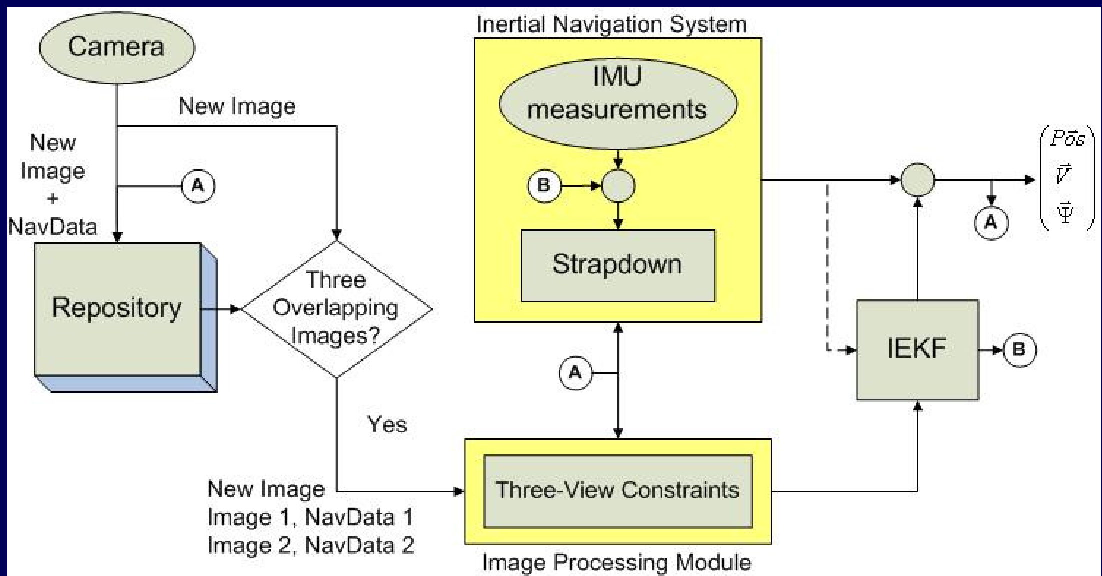
- Standard equations of IEKF update step

$$K_{k+1} = P_{k+1|k} L_{k+1}^T \left[L_{k+1} P_{k+1|k} L_{k+1}^T + D_{k+1} R_{k+1} D_{k+1}^T \right]^{-1}$$

$$\hat{X}_{k+1|k+1} = \hat{X}_{k+1|k} + K_{k+1} \left(\vec{z}_{k+1} - L_{k+1} \hat{X}_{k+1|k} \right)$$

$$P_{k+1|k+1} = \left[I - K_{k+1} L_{k+1} \right] P_{k+1|k} \left[. \right]^T + K_{k+1} D_{k+1} R_{k+1} D_{k+1}^T K_{k+1}^T$$

- Simplified vision-aided navigation scheme

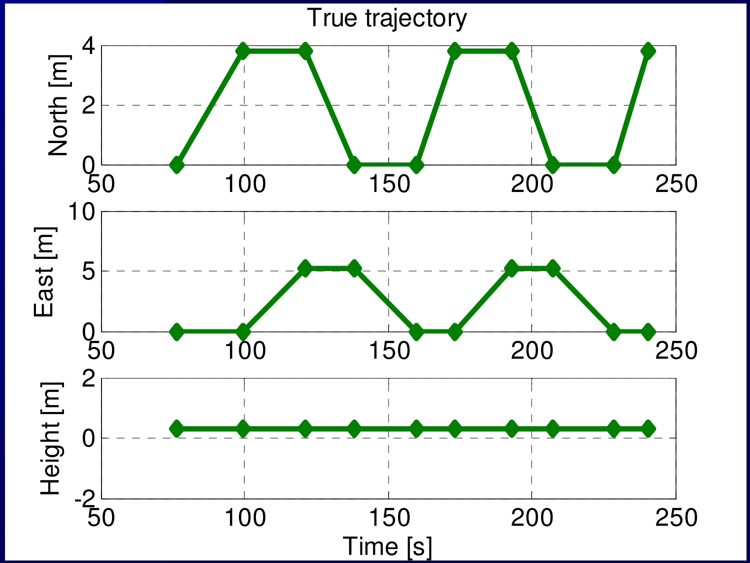


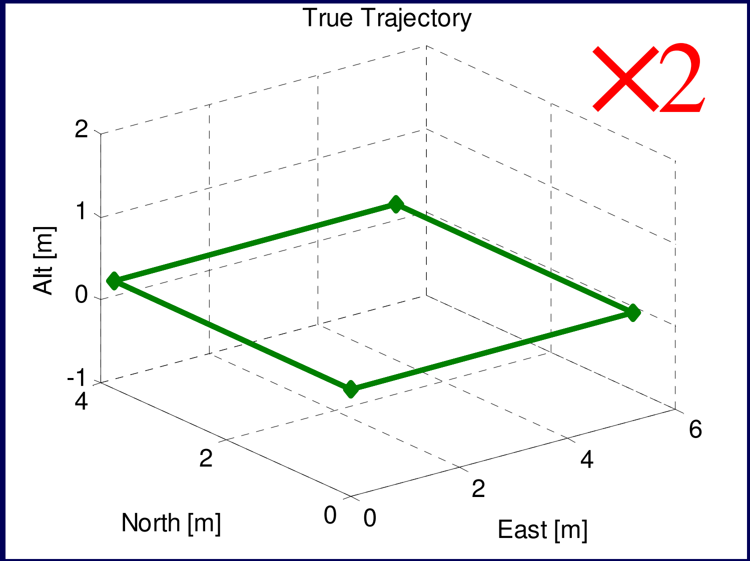
Experiment Results

- Experiment Setup
  - An IMU and a camera were mounted on top of a ground vehicle
  - IMU\INS: Xsens MTi-G
  - Camera: Axis 207MW
- IMU data and captured images were stored and synchronized
  - IMU data @ 100Hz
  - Imagery data @ 15Hz
- The method was applied in two scenarios
  - Sequential updates
  - Loop updates

Experiment Results (Cont.)

True trajectory





Recorded imagery

Experiment Results (Cont.)

Example

Image 1

Image 2

Image 3

- Implementation details (simplified)

Matching Triplets

over the Fundamental matrix model

Image 1 Image 2

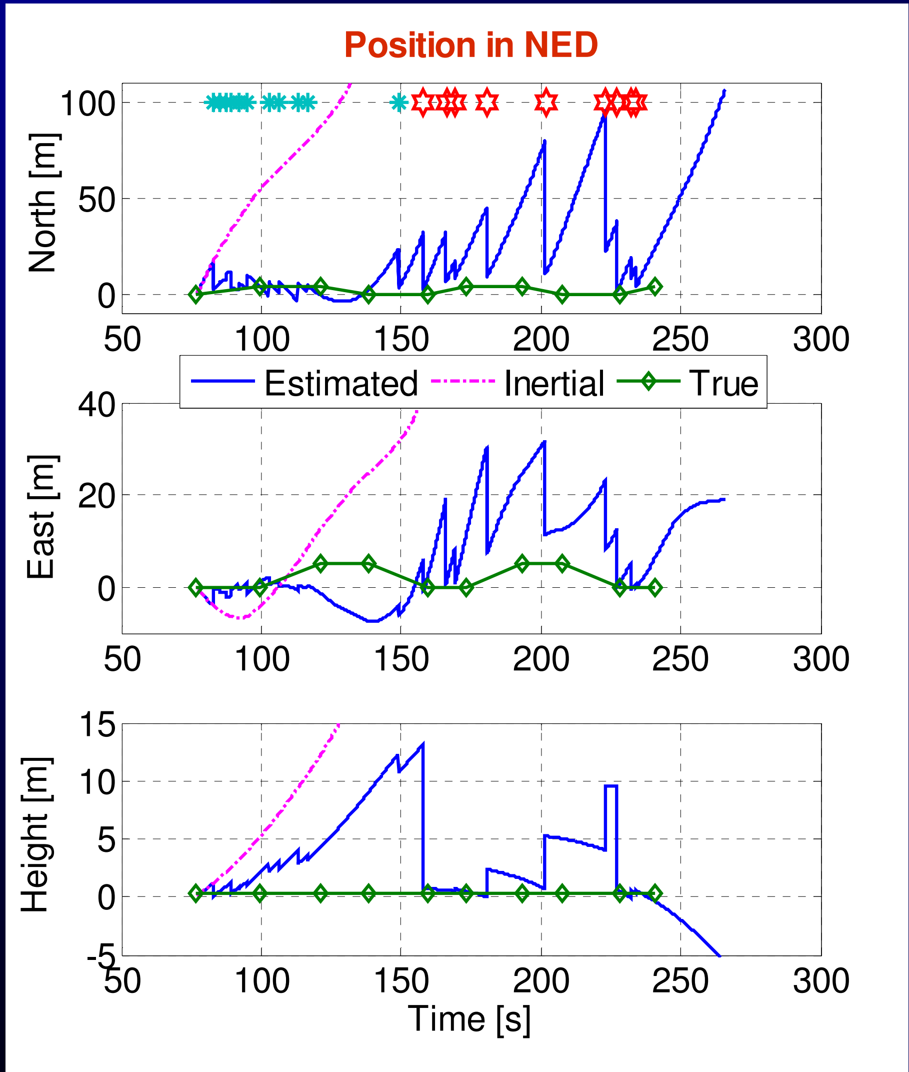
- The Fundamental matrix is not required

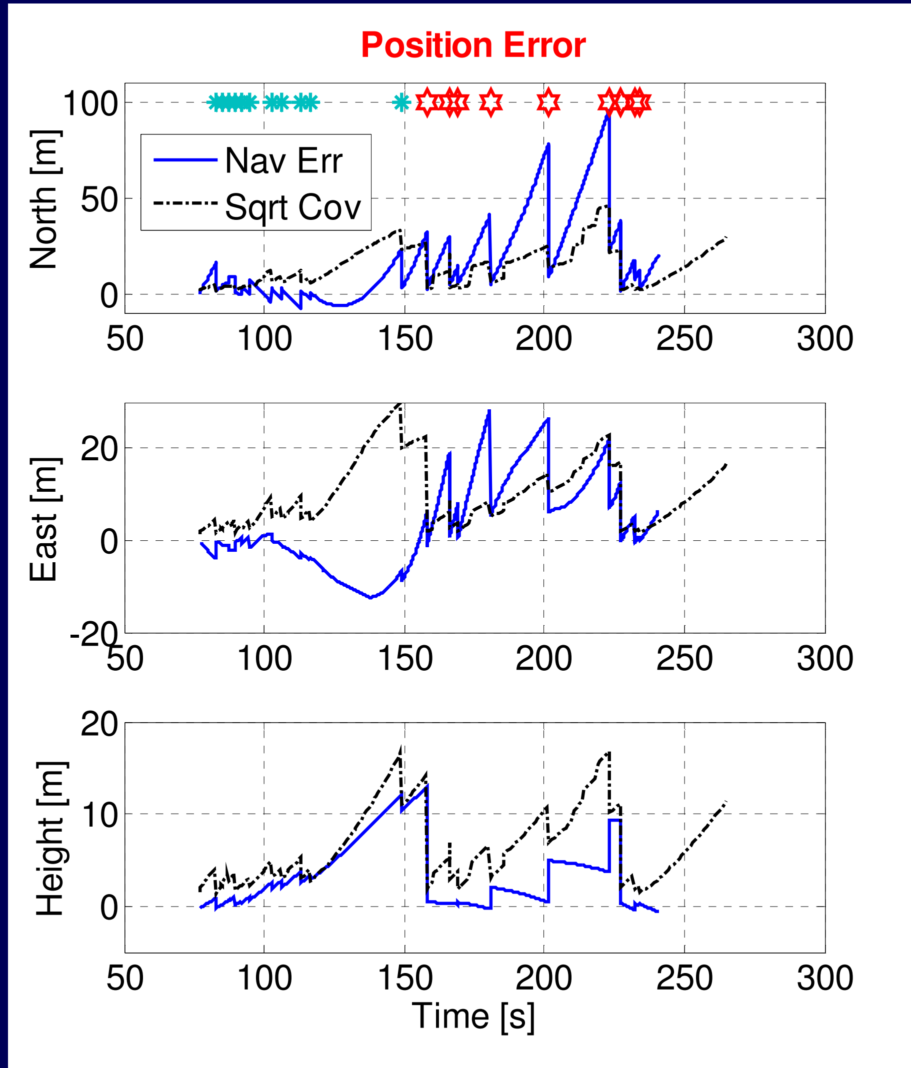
Image 2 Image 3

$$\left\{\vec{q}_{1_i}^{\,C_1}, \vec{q}_{2_i}^{\,C_2}, \vec{q}_{3_i}^{\,C_3}\right\}_{i=1}^{N_{123}}, \left\{\vec{q}_{1_i}^{\,C_1}, \vec{q}_{2_i}^{\,C_2}\right\}_{i=1}^{N_{12}}, \left\{\vec{q}_{2_i}^{\,C_2}, \vec{q}_{3_i}^{\,C_3}\right\}_{i=1}^{N_{23}}$$

Experiment Results (Cont.)

Sequential upd
Loop upd





Potential Applications

- Navigation aiding in loop scenarios
  - Additional scenarios with overlapping three images
- Satellite navigation
- Formation flying navigation

- The developed constraints may be used in SLAM framework

Summary

- A new method for vision-aided navigation based on three-view geometry was presented
  - The method does not require any a-priori information or external sensors, apart from the camera sensor
- The associated navigation data for each of the three images allowed to determine the scale ambiguity
- The method allows to reduce position errors in all axes to the levels of errors present while the first two images were captured
- Reduced computational requirements for vision-aided navigation phase
  - Environment representation construction (e.g. mosaic) may be executed in a background process
- Various potential applications