2009 AIAA Guidance, Navigation and Control Conference

Real-Time Mosaic-Aided Aerial Navigation: I. Motion Estimation

Vadim Indelman, Pini Gurfil
Distributed Space Systems Lab,
Aerospace Engineering, Technion

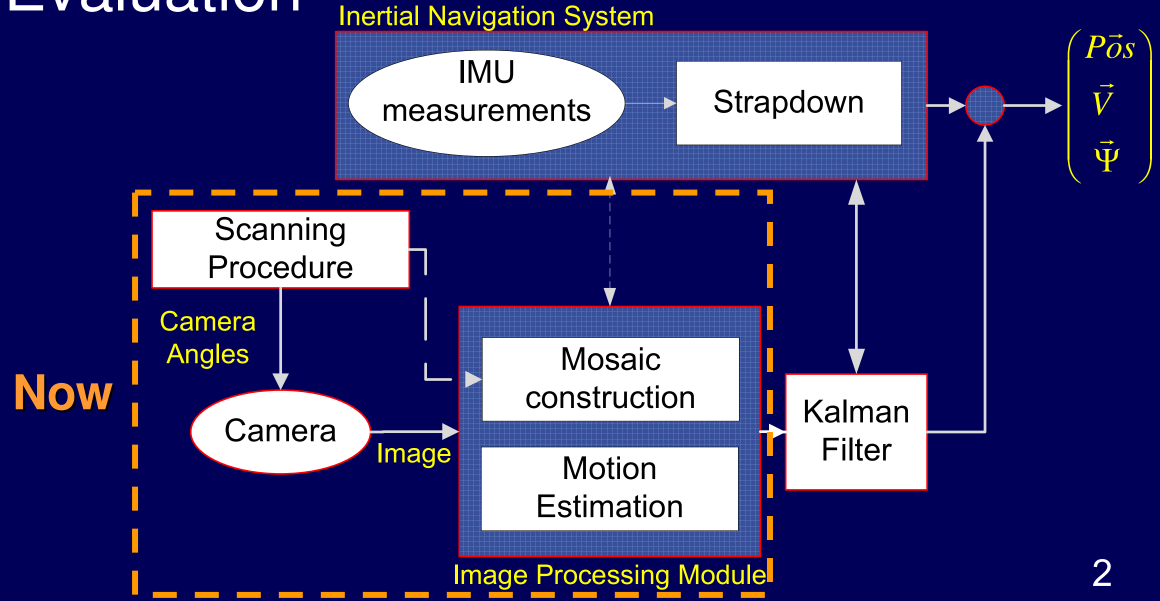
Ehud Rivlin
Computer Science, Technion

Hector Rotstein
RAFAEL

August 2009

Contents

- Introduction
- Camera Scanning Procedure
- Mosaic Construction Method
- Image-Based Motion Estimation
- Performance Evaluation
- Summary

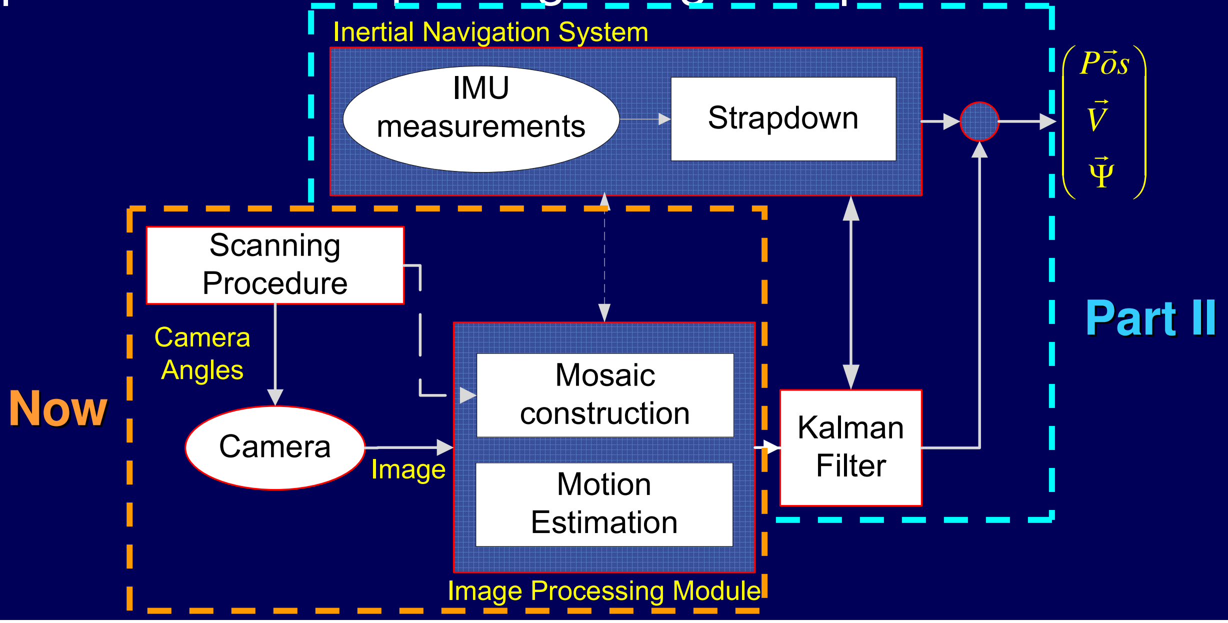


Introduction

- Navigation of airborne platforms is attended
- 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)
- This work (Part I and II)
  - Vision-based navigation aiding independent method
  - Integration of visual information into a consistent set (the mosaic image)
- Connection to Simultaneous Localization and Mapping (SLAM)

Research Idea

- An airborne platform is equipped with a low-quality Inertial Measurement Unit (IMU) and a gimbaled camera
- During the flight
  - The camera captures images according to a scan procedure
  - A mosaic image is constructed from the captured images
- Objective:
  - Utilize the camera scanning and mosaic construction processes for improving navigation performance



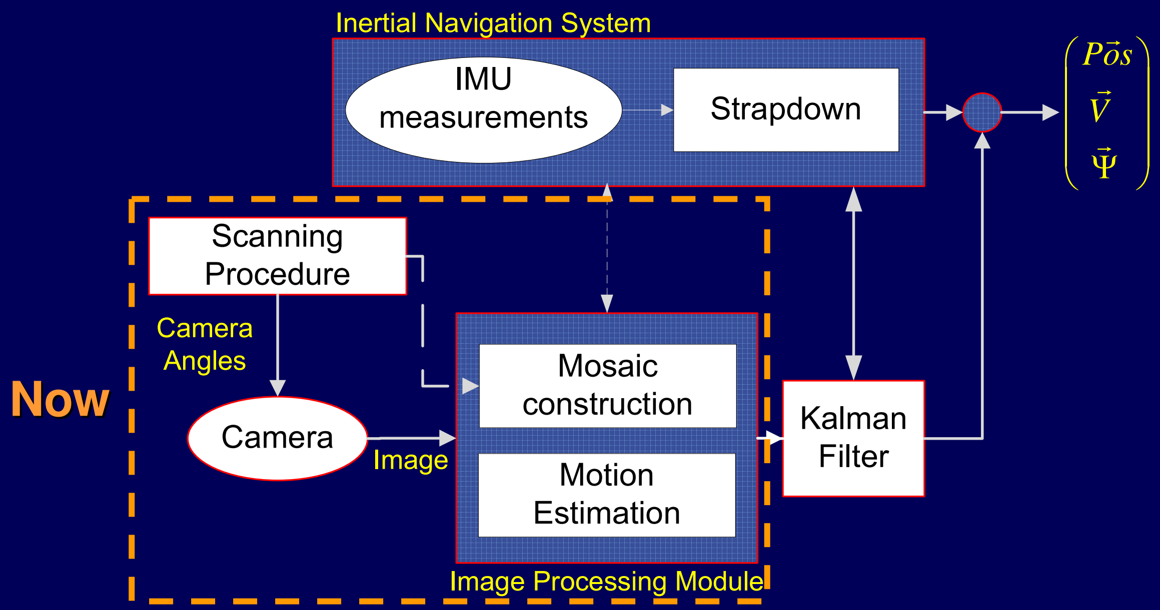
Previous Work

- "Improved Real-Time Video Mosaicking of the Ocean Floor", Fleischer S., et. al., 1995
- "*Results on Underwater Mosaic-based Navigation*", Gracias N. et. al., 2002
- "*Bearing-Only SLAM for an Airborne Vehicle*", Bryson M. and Sukkarieh S., 2005
- "*Improving Vision-based Planar Motion Estimation for Unmanned Aerial Vehicles through Online Mosaicing*", Caballero F., et. al., 2006
- "Real-Time Visual Mosaicking and Navigation on the Seafloor", Richmond K., 2009

Statement of Purpose

- Introduction
- Camera Scanning
- Mosaic Construction
- Motion Estimation
- Performance Evaluation
- Summary

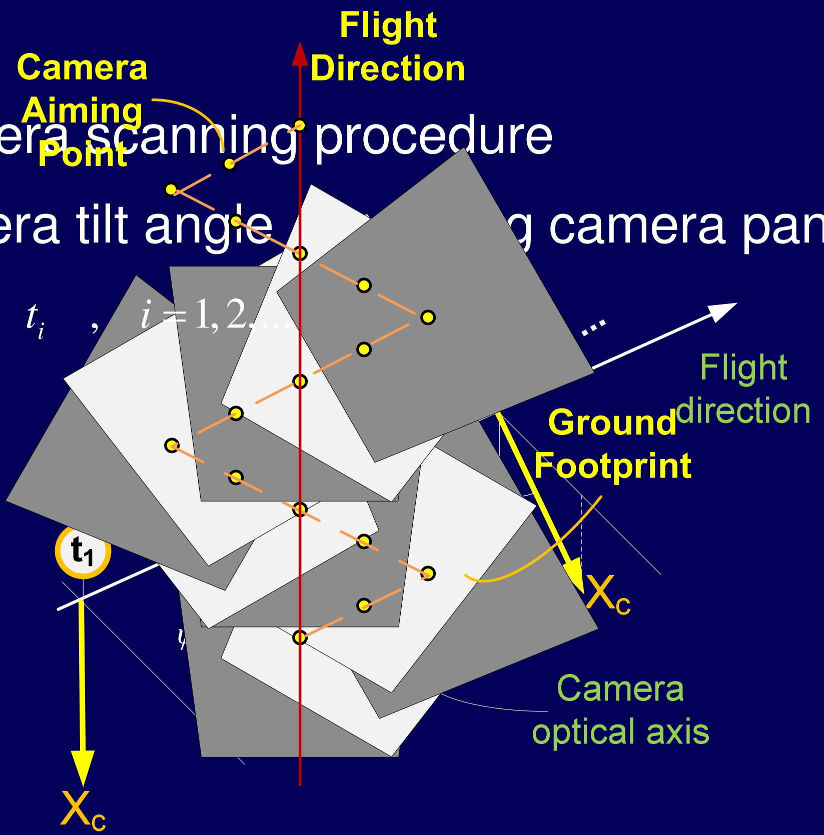
- Image-based motion estimation and vision-aided navigation in challenging scenarios
  - Narrow FOV camera
  - Low-texture scene
  - Poor weather conditions
  - (No GPS)



Camera Scanning

- Camera Scanning - Requirements
  - An overlapping area between adjacent images
  - Desired: Overlapping regions between a new image and images from previous scan stripe

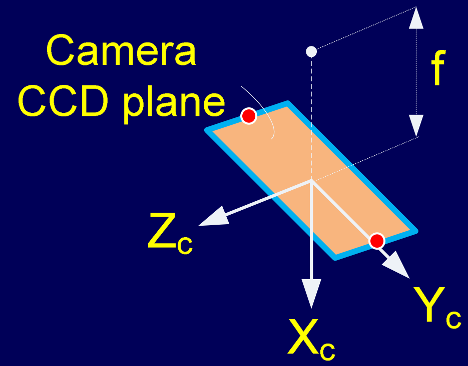
- Implemented camera scanning procedure
  - Constant camera tilt angle ... g camera pan angle ψ^C
  - Discrete time: t_i , $i = 1, 2, \ldots$



Camera Scanning – In Formulas

- Assumption: Sufficient overlap along the flight heading (V/h vs. image capture frequency)

- New line of sight (LOS) unit vector



$$\hat{r}^{C}_{NEW} = \frac{\left[f \quad \pm CCD_{Y_C}/2 \quad 0\right]^T}{\|\cdot\|}$$

$$\hat{r}^{B}_{NEW} = T^{C}_{B}\left(\psi^{C}\right)\hat{r}^{C}_{NEW}$$

- Camera angles calculation

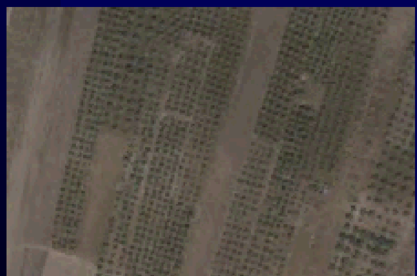
$$\hat{r}^{B}_{NEW} = T^{C}_{B}\left(\psi^{C}_{NEW}\right)\left[1 \quad 0 \quad 0\right]^T = \left[0 \quad \sin\psi^{C}_{NEW} \quad \cos\psi^{C}_{NEW}\right]^T$$

$$\mapsto \psi^{C}_{NEW} = \tan^{-1}\left(\frac{\hat{r}^{B}_{NEW}(2)}{\hat{r}^{B}_{NEW}(3)}\right)$$

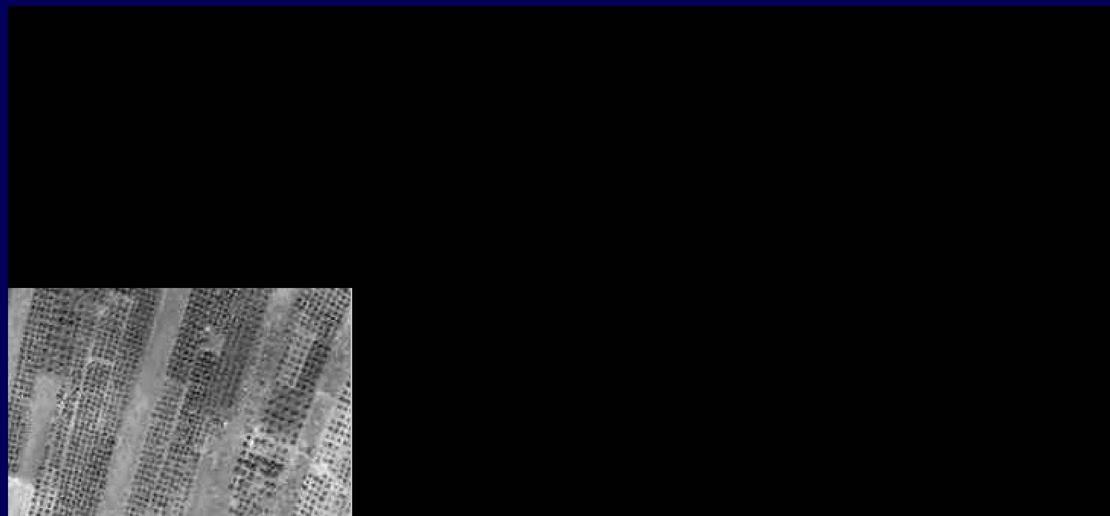
Mosaic Construction

- Concept
  - A single mosaic image
  - Homography matrix estimation between a new image and previous mosaic image
  - Image warping based on H
  - Integration of the two images into an updated mosaic image
- Google Earth is used as a substitute for complicated experiments

Camera Scanning image sequence
(from Google Earth)



Mosaic image construction



Homography Estimation

Introduction

Camera Scanning

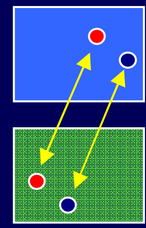
Mosaic Construction

Motion Estimation

Performance Evaluation

Summary

- Input:
  - Two images with an overlapping region
- Denote
  - $(\vec{x}, \vec{x}\,')$- a pair of corresponding points in the two images
    - Homogeneous coordinates $\vec{x} = (x \quad y \quad 1)^T$



- General scene – Fundamental matrix connection

$$(\vec{x}\,')^T \cdot F \cdot \vec{x} = 0 \quad , \quad F \in \Re^{3\times3}$$

- Planar scene - Homography connection

$$\vec{x}\,' \cong H\vec{x} \quad , \quad H \in \Re^{3\times3}$$

equality up to a scale

Homography Estimation (Cont.)

- Homography\Fundamental matrix Estimation Process:
  - Features extraction from input images (SIFT)
  - Computation of matching features
  - Outliers rejection using RANSAC algorithm ⟶ $\{(\vec{x}_i, \vec{x}_i')\}_{i=1}^{N}$

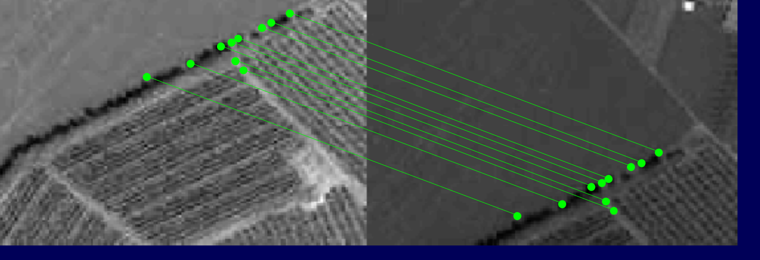
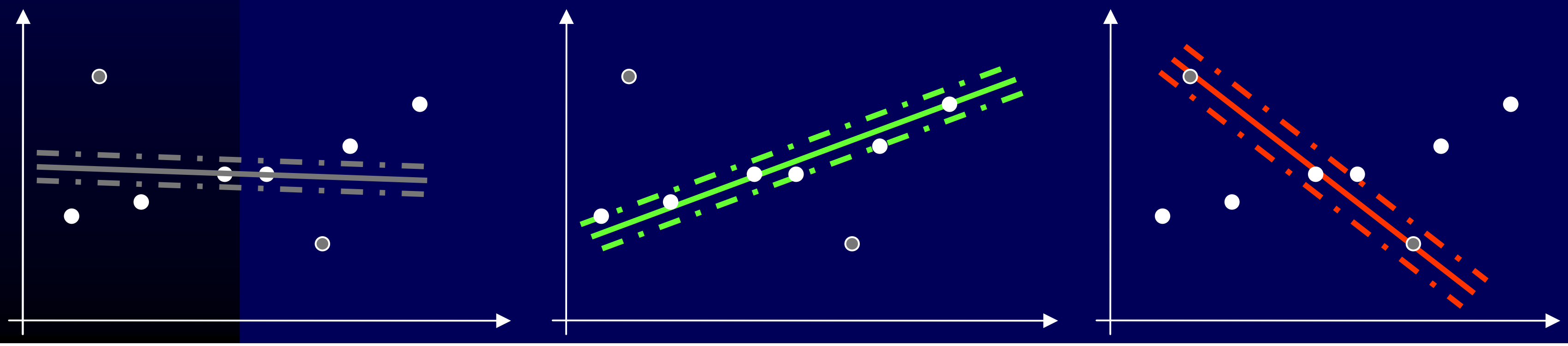


Image 1 Image 2

  - Least Squares (LS) estimation of $\{H, F\}$

RANSAC - Robust Model Estimation

- Robust estimation of a model from a data set that contains outliers
- Example – robust line estimation
  - Goal: LS line estimation from **valid** points (inliers) of a given data set
- Algorithm steps
  - Draw 2 points from the given points set. Establish a line from the chosen points.
  - Compute number of points (support) that are within a proximity envelope.
  - Repeat the process and choose the pair of points that maximize the support.
    - The points that are outside the proximity envelope are considered to be outliers.
  - Perform LS line estimation based only on valid points (inliers)



Homography Estimation (Cont.)

- For each matching pair $\left(\vec{x}_i, \vec{x}_i^{'}\right)$
  - Denote:
    - $\vec{x}_i = \left(x_i \quad y_i \quad w_i\right)^T \quad , \quad \vec{x}_i^{'} = \left(x_i^{'} \quad y_i^{'} \quad w_i^{'}\right)^T$
    - $\vec{h}$ - a vector made up of entries of H
  - Homography relation $\vec{x}_i^{'} = H\vec{x}_i$ may be written as

$$A_i\vec{h} = \vec{0} \qquad A_i = \begin{bmatrix} \vec{0}^T & -w_i^{'}\vec{x}_i^T & y_i^{'}\vec{x}_i^T \\ w_i^{'}\vec{x}_i^T & \vec{0}^T & -x_i^{'}\vec{x}_i^T \end{bmatrix}$$

- Taking into account all matching pairs: $A\vec{h} = \vec{0}$
- H may be found using standard techniques

- Fundamental matrix estimation – $\left(\vec{x}_i^{'}\right)^T \cdot F \cdot \vec{x}_i = 0$

$$\tilde{A}_i\vec{f} = 0 \quad , \quad \tilde{A}_i = \left(x_i^{'}x_i \quad x_i^{'}y_i \quad x_i^{'} \quad y_i^{'}x_i \quad y_i^{'}y_i \quad y_i^{'} \quad x_i \quad y_i \quad 1\right)$$

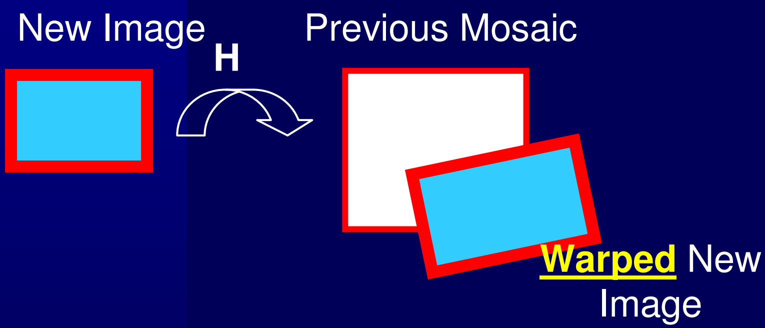
$$\tilde{A}\vec{f} = \vec{0} \quad \rightarrow F$$

Mosaic Construction (Cont.)

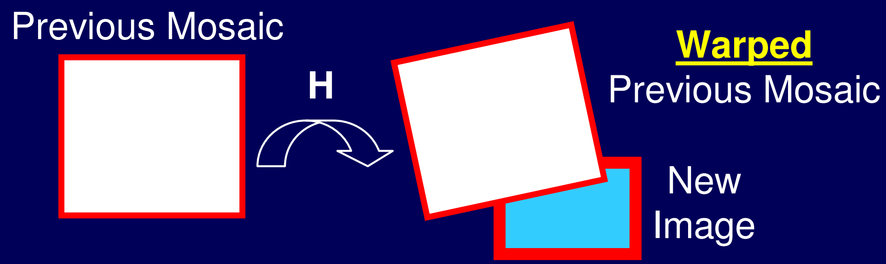
- Homography estimation based on two images √
- Next
  - Mosaic construction logic
  - Incremental homography estimation

Mosaic Construction (Cont.)

- Simplified idea
  - **1st Alternative**



  - **2nd Alternative**

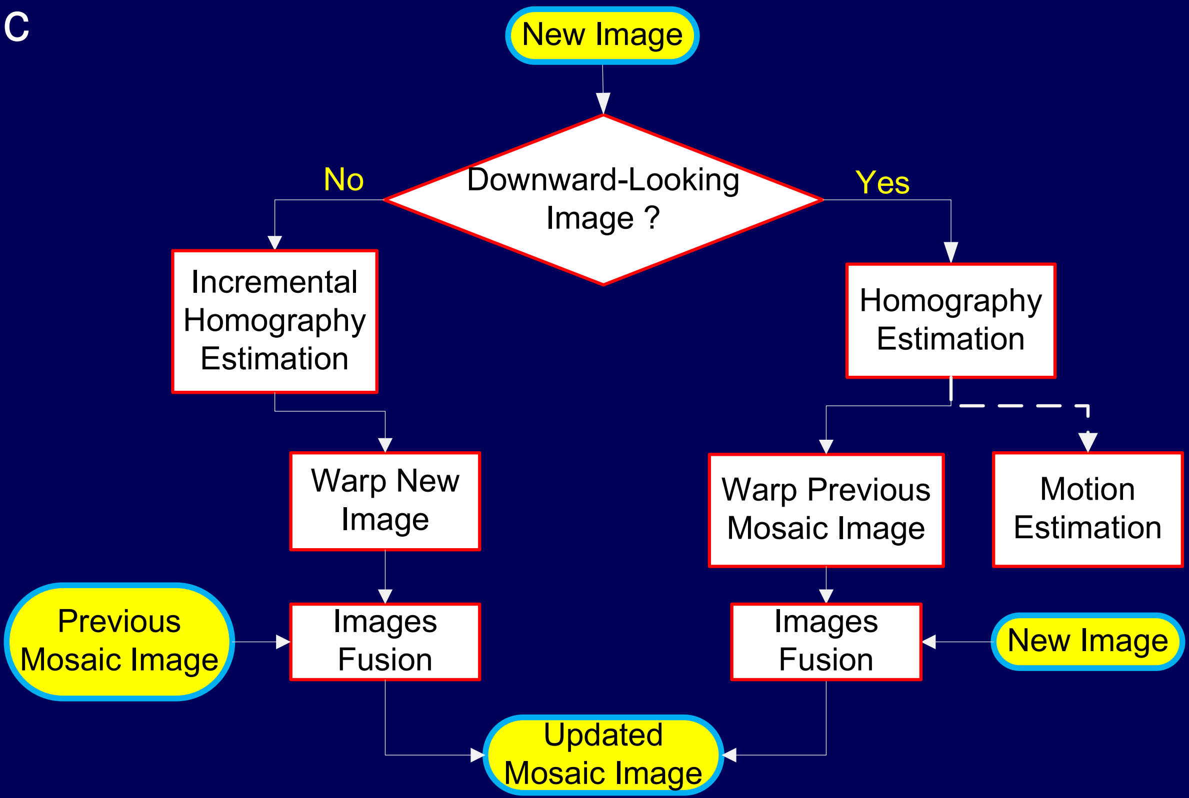


  - Final step: Images fusion into updated mosaic image

- Notes
  - Mosaic warping is required when motion estimation is performed
  - Computational aspects

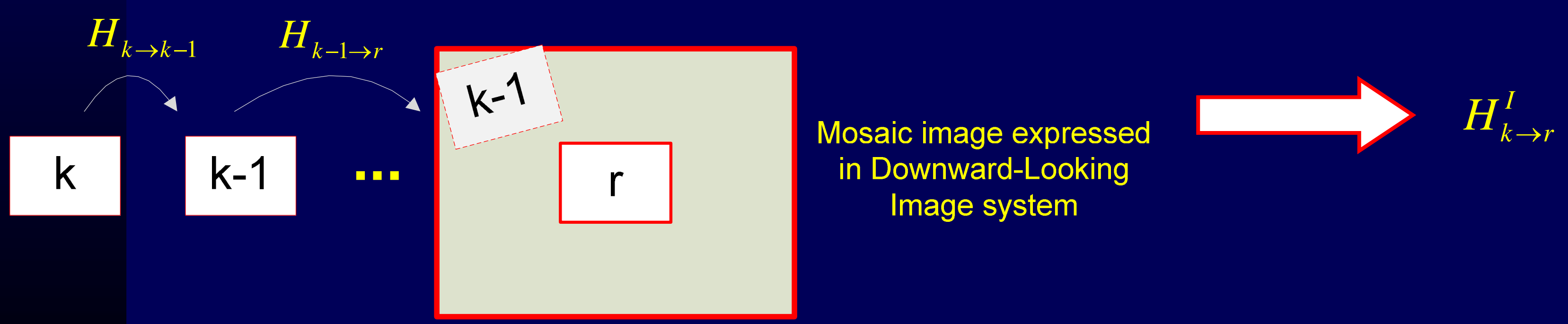
Mosaic Construction (Cont.)

- Definitions
  - Downward-Looking image: Camera points downwards
  - Non-Downward-Looking image: Otherwise
- Logic



Incremental Homography Estimation

- Why? Alignment errors are accumulated as more images are added into mosaic
- Solution – Scan-Coupled Global Alignment
  (based on: [Negahdaripour2002], [Caballero2006])
  - **k**: New image **r**: Previous Downward-Looking image
  - **k-1**: Previous image
  - Estimate homography between new image and previous image: $H_{k \to k-1}$
  - Retrieve homography between previous image and mosaic: $H_{k-1 \to r}$
  - Homography between new image and mosaic: $H^{I}_{k \to r} = H_{k-1 \to r} H_{k \to k-1}$



Incremental Homography Estimation (Cont.)

- Why? Alignment errors are accumulated as more images are added into mosaic
- Solution – Scan-Coupled Global Alignment
  (based on: [Negahdaripour2002], [Caballero2006])
  - **k**: New image index   **r**: Previous Downward-Looking image
  - **k-1**: Previous image
  - Correction of $H^{I}_{k \to r}$:
    - Use $H^{I}_{k \to r}$ to warp new image into $\tilde{I}^{r}_{k}$
    - Calculate correction homography H_{corr} between $\tilde{I}^{r}_{k}$ and the mosaic
    - Finally: $H_{k \to r} = H_{corr} H^{I}_{k \to r}$

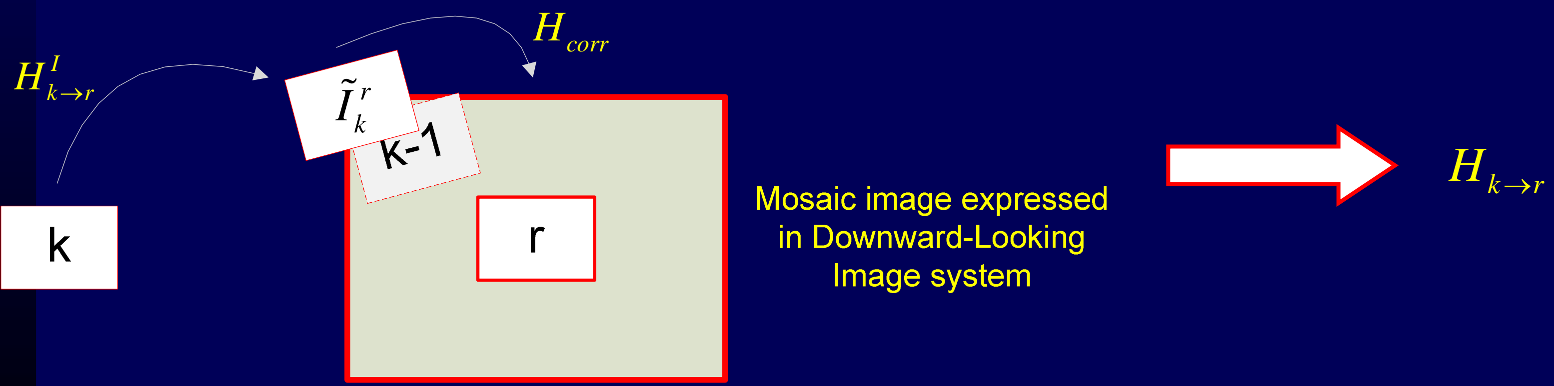


Image-Based Motion Estimation

- Recall
  - Homography matrix (planar scene) $\vec{x}' \cong H\vec{x}$
  - Fundamental matrix (general scene) $(\vec{x}')^T \cdot F \cdot \vec{x} = 0$

- The motion parameters are encoded in each of these matrices

$$H = K'\left[R - \frac{\vec{t}}{z}\vec{n}^T\right]K^{-1} \qquad F = K'^{-T}\left[\vec{t}\right]_\times RK^{-1}$$

  - $\vec{t}$ - **translation motion, can be estimated only up to a scale**
  - R - **rotation motion**
  - z - scene depth (unknown)
  - K, K' - calibration matrices (assumed to be known)
  - $\vec{n}$ - normal vector to the scene plane

Image-Based Motion Estimation (Cont.)

- The motion parameters $(\vec{t}, R)$ may be extracted from H or F using standard techniques
- From Homography [Tsai 1982]: $H = K'\left[R - \frac{\vec{t}}{z}\vec{n}^T\right]K^{-1}$

  - Calibrated homography $H^C = (K')^{-1} HK$
  - SVD $H^C = UDV^T$
  - Rotation $R = U\begin{bmatrix} \alpha & 0 & \beta \\ 0 & 1 & 0 \\ -s\beta & 0 & s\alpha \end{bmatrix}V^T$
  - Translation (up to scale) $\vec{t} = \frac{1}{w}\left[-\beta\vec{u}_1 + \left(\frac{\lambda_3}{\lambda_2} - s\alpha\right)\vec{u}_3\right]$
  - Normal to scene plane $\vec{n} = w(\delta\vec{v}_1 + \vec{v}_3)$

$$\lambda_i = D(i,i) \quad i = 1,2,3$$
$$\alpha, \beta, \delta = f(\lambda_1, \lambda_2, \lambda_3)$$
$$s = \det(U)\det(V)$$

w - scale factor

Motion Estimation - Performance Evaluation

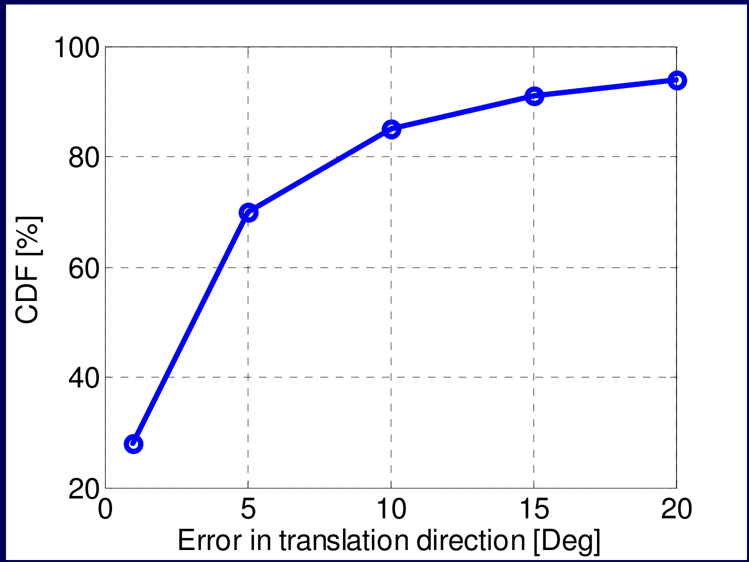
- Real images from Google Earth
- Mosaic-based motion estimation – performed each time a new Downward-Looking image is captured

- Examined scenarios
  - Wide FOV camera, general (non-planar) scene
    - Two-view motion estimation - based on each two images (No Mosaicking)
  - Mosaic-based vs. Two-view motion estimation
    - Narrow FOV camera
    - High-texture and low-texture scenes

Two-View Motion Estimation

- Motion estimation based on consecutive camera-captured images
  - Without mosaic image construction
  - Non-planar scene →Motion estimation via Fundamental Matrix
- **Wide** field-of-view camera

CDF of translation motion estimation error



Mosaic-Based vs. Two-View Motion Estimation

Introduction

Camera Scanning

Mosaic Construction

Motion Estimation

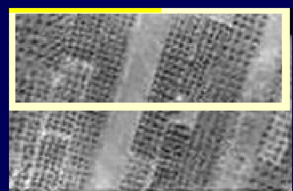
Performance Evaluation

Summary

- Mosaic construction assumes <u>planar</u> scene
- The proposed scanning and mosaic construction processes provide an enlarged overlapping area
  - The additional area is with lower quality than the original
  - Trade-off: More features, BUT some are with lower quality
  - Improves motion estimation in challenging scenarios:
    - Narrow FOV camera , Low-texture type scene

Original Overlapping Area

Additional Overlapping Area



Previous image / Previous <u>mosaic</u> image

Next image

Mosaic-Based vs. Two-View Motion Estimation (Cont.)

- Narrow FOV camera
  - Common in many airborne applications
  - Epipolar geometry is ill-conditioned
  - Planar scene assumption

- Examined Effects:
  - Changing the camera FOV: 7^0x4^0 vs. 5^0x3^0
  - Scene type: High-texture vs. Low-texture

High-texture scene

Low-texture scene

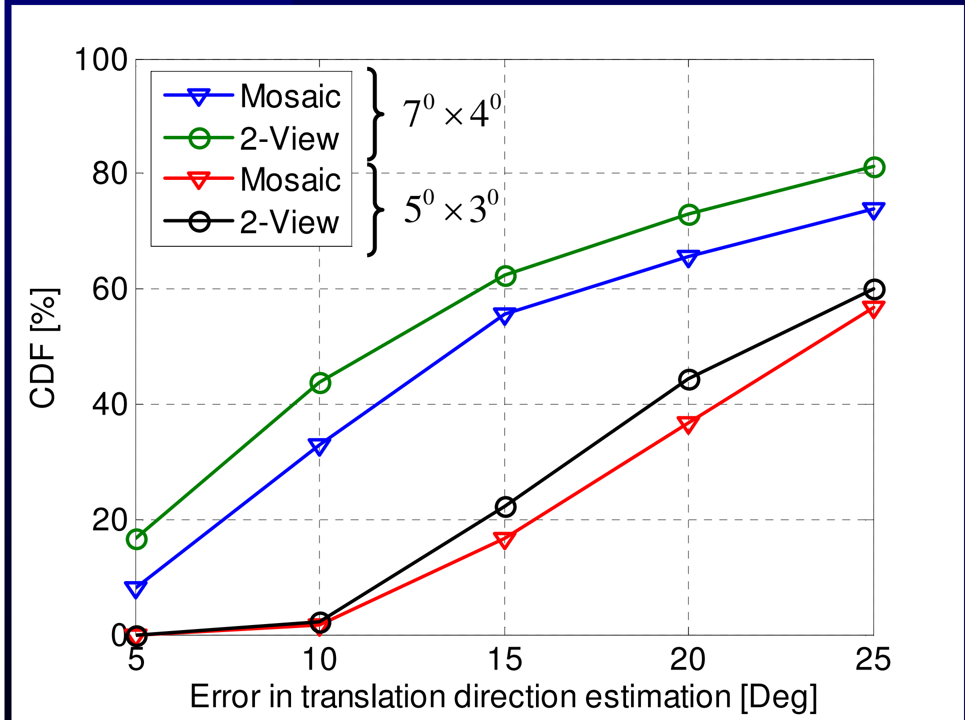
7^0x4^0 FOV

Images from Google Earth

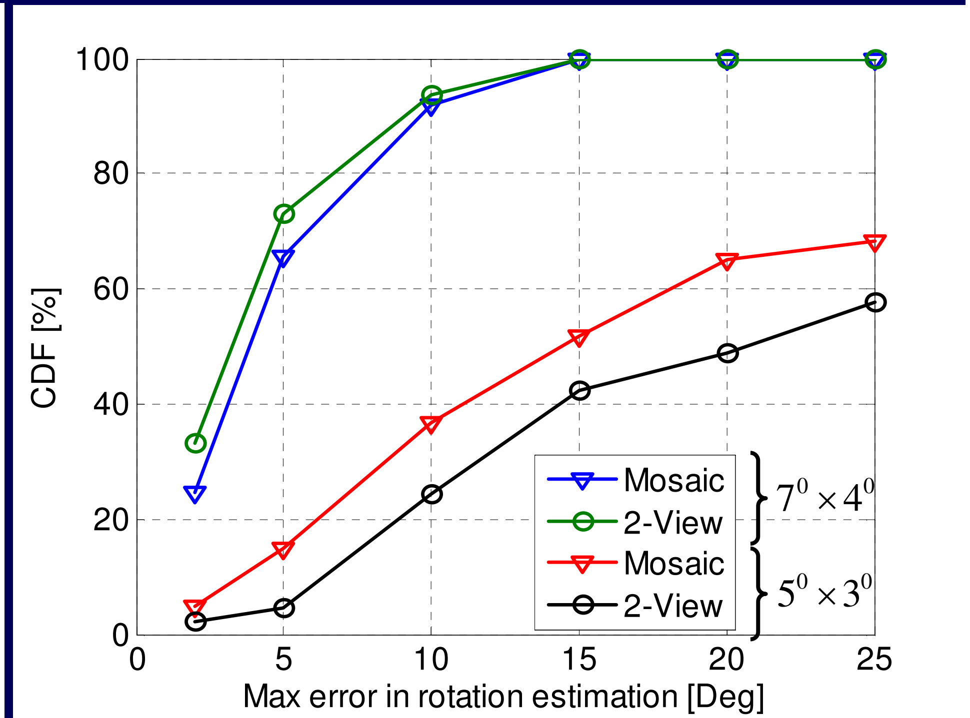
Mosaic-Based vs. Two-View Motion Estimation (Cont.)

- High-texture scene – Varying camera FOV

CDF of <u>translation</u> motion estimation error



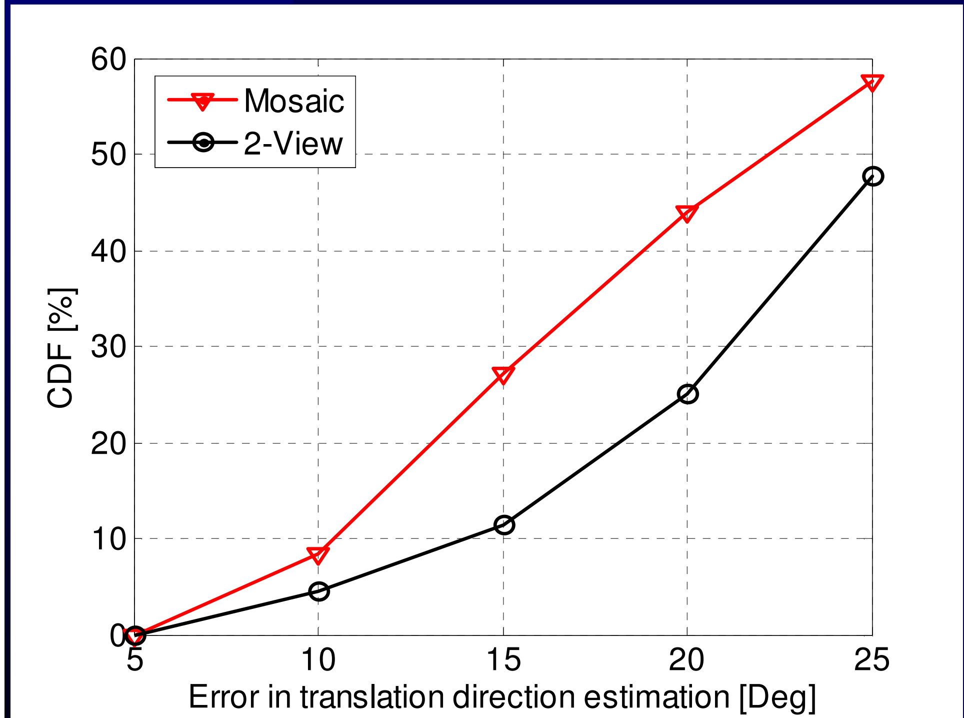
CDF of <u>rotation</u> motion estimation error



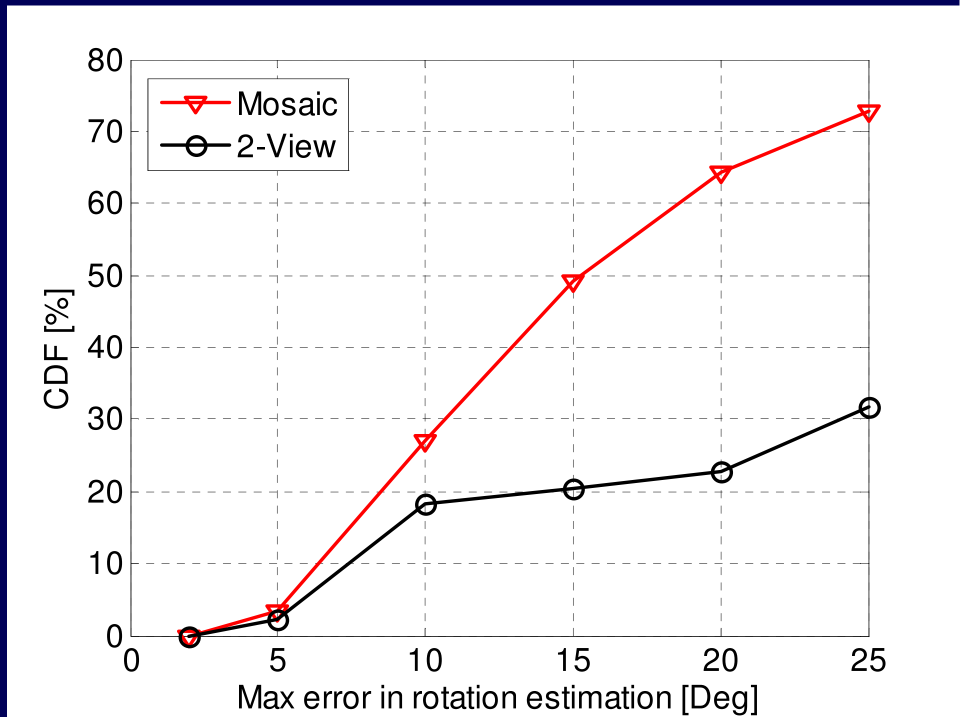
Mosaic-Based vs. Two-View Motion Estimation (Cont.)

- **Low**-texture scene
- **Narrow** camera FOV – 5^0x3^0

CDF of translation motion estimation error



CDF of rotation motion estimation error



Summary

Introduction

Camera Scanning

Mosaic Construction

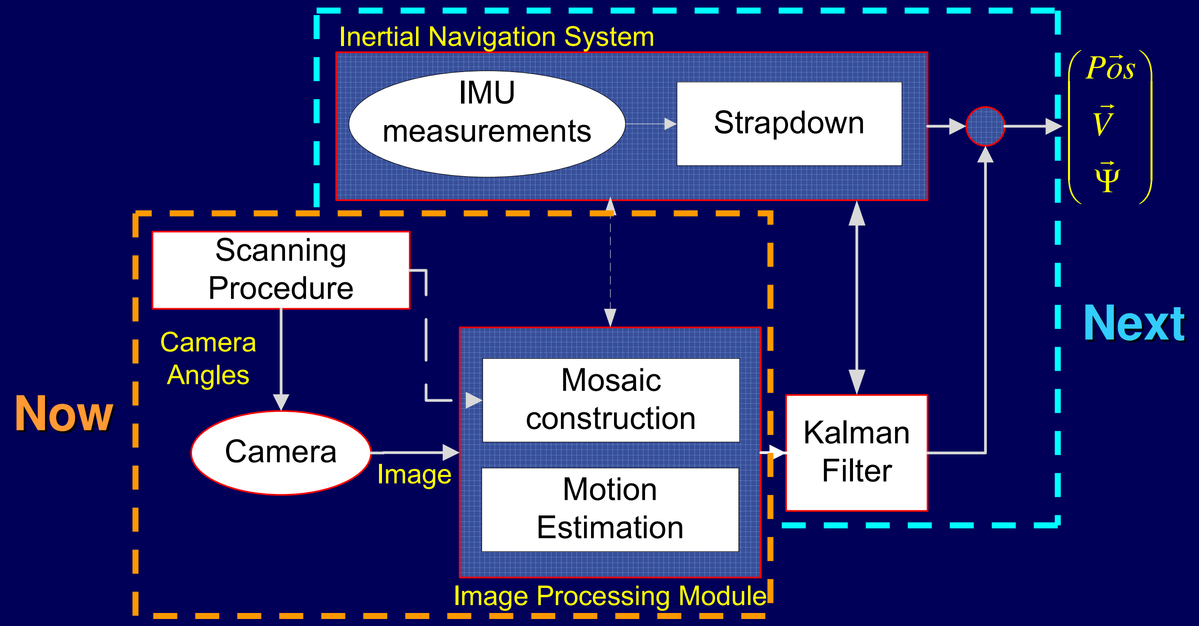
Motion Estimation

Performance Evaluation

Summary

- On-line mosaicking and camera scanning procedures were presented
  - Enlarged overlapping region
  - Improved motion estimation in challenging scenarios
- Performance evaluation based on images from Google Earth
  - Wide camera FOV, non-planar scene
  - Narrow camera FOV, high-texture and low-texture scenes
    - Mosaic-based motion estimation outperforms Two-view motion estimation in case of low-texture scene and narrow FOV camera

Next



Inertial Navigation System
IMU measurements
Strapdown
$\vec{Pos}$
$\vec{V}$
$\vec{\Psi}$
Next
Now
Scanning Procedure
Camera Angles
Camera
Image
Mosaic construction
Motion Estimation
Kalman Filter
Image Processing Module

Thank you …