Computer Vision 2
WS 2018/19

Part 16 – Visual SLAM II
08.01.2019

Prof. Dr. Bastian Leibe

RWTH Aachen University, Computer Vision Group
http://www.vision.rwth-aachen.de

Course Outline

- Single-Object Tracking
- Bayesian Filtering
- Multi-Object Tracking
- Visual Odometry
  - Sparse interest-point based methods
  - Dense direct methods
- Visual SLAM & 3D Reconstruction
  - Online SLAM methods
  - Full SLAM methods
- Deep Learning for Video Analysis

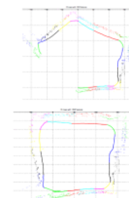


image source: [Clemente et al., RSS 2007]

Topics of This Lecture

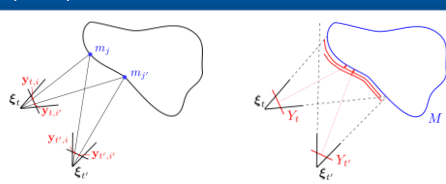
- Recap: Online SLAM methods
- EKF SLAM
  - Extended Kalman Filter formulation
  - 2D EKF SLAM example
  - Detailed analysis
- Loop Closure
- Case study: MonoSLAM
- Full SLAM methods
  - SLAM graph optimization
  - Pose graph optimization

Recap: Definition of Visual SLAM

- Visual SLAM
  - The process of simultaneously estimating the egomotion of an object and the environment map using only inputs from visual sensors on the object
- **Inputs:** images at discrete time steps t,
  - Monocular case: Set of images $I_{0:t} = \{I_0, \ldots, I_t\}$
  - Stereo case: Left/right images $I^l_{0:t} = \{I^l_0, \ldots, I^l_t\}$, $I^r_{0:t} = \{I^r_0, \ldots, I^r_t\}$
  - RGB-D case: Color/depth images $I_{0:t} = \{I_0, \ldots, I_t\}$, $Z_{0:t} = \{Z_0, \ldots, Z_t\}$
  - Robotics: **control inputs** $U_{1:t}$
- **Output**:
  - Camera pose estimates $\mathbf{T}_t \in \mathbf{SE}(3)$ in world reference frame. For convenience, we also write $\xi_t = \xi(\mathbf{T}_t)$
  - Environment map M

Slide credit: Jörg Stückler

Recap: Map Observations in Visual SLAM

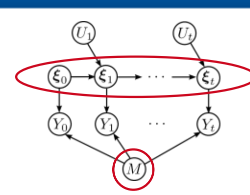


With Y_t we denote observations of the environment map in image I_t, e.g.,
- Indirect point-based method: $Y_t = \{\mathbf{y}_{t,1}, \ldots, \mathbf{y}_{t,N}\}$ (2D or 3D image points)
- Direct RGB-D method: $Y_t = \{I_t, Z_t\}$ (all image pixels)
- …

- Involves data association to map elements $M = \{m_1, \ldots, m_S\}$
  - We denote correspondences by $c_{t,i} = j,\ 1 \le i \le N,\ 1 \le j \le S$

Slide credit: Jörg Stückler

Recap: Probabilistic Formulation of Visual SLAM



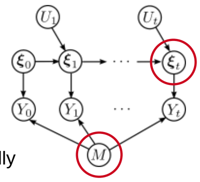
- SLAM posterior probability: $p(\xi_{0:t}, M \mid Y_{0:t}, U_{1:t})$
- Observation likelihood: $p(Y_t \mid \xi_t, M)$
- State-transition probability: $p(\xi_t \mid \xi_{t-1}, U_t)$

Slide credit: Jörg Stückler

Online SLAM Methods

- Marginalize out previous poses

$$p(\boldsymbol{\xi}_t, M \mid Y_{0:t}, U_{1:t}) =$$
$$\int \cdots \int p(\boldsymbol{\xi}_{0:t}, M \mid Y_{0:t}, U_{1:t})\, d\boldsymbol{\xi}_{t-1} \ldots d\boldsymbol{\xi}_0$$



- Poses can be marginalized individually in a recursive way
- Variants:
  - Tracking-and-Mapping: Alternating pose and map estimation
  - Probabilistic filters, e.g., EKF-SLAM

Topics of This Lecture

- Recap: Online SLAM methods
- EKF SLAM
  - Extended Kalman Filter formulation
  - 2D EKF SLAM example
  - Detailed analysis
- Loop Closure
- Case study: MonoSLAM
- Full SLAM methods
  - SLAM graph optimization
  - Pose graph optimization

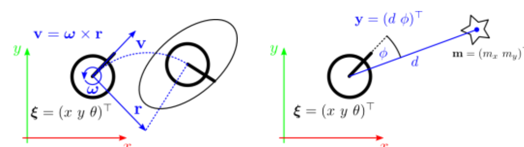
SLAM with Extended Kalman Filters

- Detected keypoint y_i in an image observes „landmark“ position m_j in the map $M = \{m_1, \ldots, m_S\}$.
- Idea: Include landmarks into state variable

$$\mathbf{x}_t = \begin{pmatrix} \boldsymbol{\xi}_t \\ \mathbf{m}_{t,1} \\ \vdots \\ \mathbf{m}_{t,S} \end{pmatrix} \qquad \Sigma_t = \begin{pmatrix} \Sigma_{t,\xi\xi} & \Sigma_{t,\xi\mathbf{m}_1} & \cdots & \Sigma_{t,\xi\mathbf{m}_S} \\ \Sigma_{t,\mathbf{m}_1\xi} & \Sigma_{t,\mathbf{m}_1\mathbf{m}_1} & \cdots & \Sigma_{t,\mathbf{m}_1\mathbf{m}_S} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{t,\mathbf{m}_S\xi} & \Sigma_{t,\mathbf{m}_S\mathbf{m}_1} & \cdots & \Sigma_{t,\mathbf{m}_S\mathbf{m}_S} \end{pmatrix}$$

$$= \begin{pmatrix} \Sigma_{t,\xi\xi} & \Sigma_{t,\xi\mathbf{m}} \\ \Sigma_{t,\mathbf{m}\xi} & \Sigma_{t,\mathbf{mm}} \end{pmatrix}$$

Example: EKF-SLAM in a 2D World

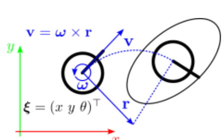


- For simplicity, let‘s assume
  - 3-DoF camera motion on a 2D plane
  - 2D range-and-bearing measurements of 2D landmarks
  - Only one measurement at a time
  - Known data association

2D EKF-SLAM State-Transition Model

- State/control variables

$$\boldsymbol{\xi}_t = (x_t\ y_t\ \theta_t)^\top \qquad \mathbf{m}_{t,j} = (m_{t,j,x}\ m_{t,j,y})^\top$$
$$\mathbf{u}_t = (v_t\ \omega_t)^\top = (\|\mathbf{v}\|_2\ \ \|\boldsymbol{\omega}\|_2)^\top$$



- State-transition model
  - Pose:

$$\boldsymbol{\xi}_t = g_\xi(\boldsymbol{\xi}_{t-1}, \mathbf{u}_t) + \boldsymbol{\epsilon}_{\xi,t} \qquad \boldsymbol{\epsilon}_{\xi,t} \sim \mathcal{N}(\mathbf{0}, \Sigma_{d_t,\xi})$$
$$g_\xi(\boldsymbol{\xi}_{t-1}, \mathbf{u}_t) = \begin{pmatrix} x_{t-1} \\ y_{t-1} \\ \theta_{t-1} \end{pmatrix} + \begin{pmatrix} -\frac{v_t}{\omega_t}\sin\theta_{t-1} + \frac{v_t}{\omega_t}\sin(\theta_t + \omega_t \Delta t) \\ \frac{v_t}{\omega_t}\cos\theta_{t-1} - \frac{v_t}{\omega_t}\cos(\theta_t + \omega_t \Delta t) \\ \omega_t \Delta t \end{pmatrix}$$

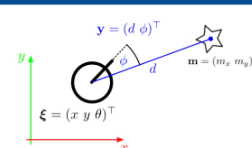
  - Landmarks: $\mathbf{m}_t = g_\mathbf{m}(\mathbf{m}_{t-1}) = \mathbf{m}_{t-1}$
  - Combined:

$$\mathbf{x}_t = g(\mathbf{x}_{t-1}, \mathbf{u}_t) + \boldsymbol{\epsilon}_t, \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_{d_t}) \qquad g(\mathbf{x}_{t-1}, \mathbf{u}_t) = \begin{pmatrix} g_\xi(\boldsymbol{\xi}_{t-1}, \mathbf{u}_t) \\ g_\mathbf{m}(\mathbf{m}_{t-1}) \end{pmatrix} \qquad \Sigma_{d_t} = \begin{pmatrix} \Sigma_{d_t,\xi} & 0 \\ 0 & 0 \end{pmatrix}$$

2D EKF-SLAM Observation Model

- State/measurement variables

$$\mathbf{y}_t = (d_t\ \phi_t)^\top \qquad \mathbf{m}_{t,j} = (m_{t,j,x}\ m_{t,j,y})^\top$$



- Observation model:

$$\mathbf{y}_t = h(\boldsymbol{\xi}_t, \mathbf{m}_{t,c_t}) + \boldsymbol{\delta}_t \qquad \boldsymbol{\delta}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_{m_t})$$
$$h(\boldsymbol{\xi}_t, \mathbf{m}_{t,c_t}) = \begin{pmatrix} \|\mathbf{m}^{\text{rel}}_{t,c_t}\|_2 \\ \operatorname{atan2}\left(\mathbf{m}^{\text{rel}}_{t,c_t,y}, \mathbf{m}^{\text{rel}}_{t,c_t,x}\right) \end{pmatrix}$$
$$\mathbf{m}^{\text{rel}}_{t,c_t} := \mathbf{R}(-\theta_t)\left(\mathbf{m}_{t,c_t} - (x_t\ y_t)^\top\right)$$

State Initialization

- First frame:
  - Anchor reference frame at initial pose $\mathbf{x}_0 = \mathbf{0}$
  - Set pose covariance to zero $\Sigma_{0,\xi\xi}^- = \mathbf{0}$
- New landmark:
  - Initial position unknown $\Sigma_{0,\xi\mathbf{m}}^- = \Sigma_{0,\mathbf{m}\xi}^- = \mathbf{0}$
  - Initialize mean at zero
  - Initialize covariance to infinity (large value) $\Sigma_{0,\mathbf{mm}}^- = \infty\mathbf{I}$

Topics of This Lecture

- Recap: Online SLAM methods
- EKF SLAM
  - Extended Kalman Filter formulation
  - 2D EKF SLAM example
  - Detailed analysis
- Loop Closure
- Case study: MonoSLAM
- Full SLAM methods
  - SLAM graph optimization
  - Pose graph optimization

Evolution of State Estimate on Prediction

- *How is the state estimate modified on a state-transition?*
- Recap: EKF Prediction

$$\mathbf{x}_t^- = g(\mathbf{x}_{t-1}^+, \mathbf{u}_t) \qquad \Sigma_t^- = \mathbf{G}_t \Sigma_{t-1}^+ \mathbf{G}_t^\top + \Sigma_{d_t}$$

$$\mathbf{G}_t = \nabla_{\mathbf{x}} g(\mathbf{x}, \mathbf{u}_t)|_{\mathbf{x}=\mathbf{x}_{t-1}^+} \qquad \mathbf{G}_{t,\xi} := \nabla_\xi g_\xi(\xi, \mathbf{u}_t)|_{\xi=\xi_{t-1}^+}$$

$$\mathbf{x}_t^- = \begin{pmatrix} g_\xi(\xi_{t-1}^+, \mathbf{u}_t) \\ \mathbf{m}_{t-1}^+ \end{pmatrix}$$

only the mean pose is updated!

$$\mathbf{x}_t = \begin{pmatrix} \xi_t \\ \mathbf{m}_{t,1} \\ \vdots \\ \mathbf{m}_{t,S} \end{pmatrix}$$

Evolution of State Estimate on Prediction

- How is the state estimate modified on a state-transition?
- Recap: EKF Prediction

$$\mathbf{x}_t^- = g(\mathbf{x}_{t-1}^+, \mathbf{u}_t) \qquad \Sigma_t^- = \mathbf{G}_t \Sigma_{t-1}^+ \mathbf{G}_t^\top + \Sigma_{d_t}$$

$$\mathbf{G}_t = \nabla_{\mathbf{x}} g(\mathbf{x}, \mathbf{u}_t)|_{\mathbf{x}=\mathbf{x}_{t-1}^+} \qquad \mathbf{G}_{t,\xi} := \nabla_\xi g_\xi(\xi, \mathbf{u}_t)|_{\xi=\xi_{t-1}^+}$$

$$\begin{pmatrix} \mathbf{G}_{t,\xi} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \Sigma_{t-1,\xi\xi}^+ & \Sigma_{t-1,\xi\mathbf{m}}^+ \\ \Sigma_{t-1,\mathbf{m}\xi}^+ & \Sigma_{t-1,\mathbf{mm}}^+ \end{pmatrix} \begin{pmatrix} \mathbf{G}_{t,\xi}^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} + \begin{pmatrix} \Sigma_{d_t,\xi} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} =$$

$$\begin{pmatrix} \mathbf{G}_{t,\xi}\Sigma_{t-1,\xi\xi}^+\mathbf{G}_{t,\xi}^\top + \Sigma_{d_t,\xi} & \mathbf{G}_{t,\xi}\Sigma_{t-1,\xi\mathbf{m}}^+ \\ \Sigma_{t-1,\mathbf{m}\xi}^+\mathbf{G}_{t,\xi}^\top & \Sigma_{t-1,\mathbf{mm}}^+ \end{pmatrix}$$

covariances are transformed to the new pose!

$$\Sigma_t = \begin{pmatrix} \Sigma_{t,\xi\xi} & \Sigma_{t,\xi\mathbf{m}_1} & \cdots & \Sigma_{t,\xi\mathbf{m}_S} \\ \Sigma_{t,\mathbf{m}_1\xi} & \Sigma_{t,\mathbf{m}_1\mathbf{m}_1} & \cdots & \Sigma_{t,\mathbf{m}_1\mathbf{m}_S} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{t,\mathbf{m}_S\xi} & \Sigma_{t,\mathbf{m}_S\mathbf{m}_1} & \cdots & \Sigma_{t,\mathbf{m}_S\mathbf{m}_S} \end{pmatrix}$$

Evolution of State Estimate on Correction

- *How is the state estimate modified on a landmark measurement?*
- Recap: EKF Correction

$$\mathbf{K}_t = \Sigma_t^- \mathbf{H}_t^\top \left(\mathbf{H}_t \Sigma_t^- \mathbf{H}_t^\top + \Sigma_{m_t}\right)^{-1} \qquad \mathbf{H}_t = \nabla_{\mathbf{x}} h(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_t^-}$$

$$\mathbf{x}_t^+ = \mathbf{x}_t^- + \mathbf{K}_t \left(\mathbf{y}_t - h\left(\mathbf{x}_t^-\right)\right)$$

$$\Sigma_t^+ = \left(\mathbf{I} - \mathbf{K}_t\mathbf{H}_t\right)\Sigma_t^-$$

- How do correlations propagate onto mean and covariance through the Kalman gain?

Evolution of State Estimate on Correction

- Let's have a closer look at the Kalman gain

$$\mathbf{K}_t = \Sigma_t^- \mathbf{H}_t^\top \left(\mathbf{H}_t \Sigma_t^- \mathbf{H}_t^\top + \Sigma_{m_t}\right)^{-1} \qquad \mathbf{H}_t = \nabla_{\mathbf{x}} h(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_t^-}$$

- The Jacobian of the observation function is only non-zero for the pose and the measured landmark:

$$\mathbf{H}_t = \begin{pmatrix} \mathbf{H}_{t,\xi} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{H}_{t,\mathbf{m}_{c_t}} & \mathbf{0} & \ldots & \mathbf{0} \end{pmatrix}$$

$$\mathbf{H}_{t,\xi} = \nabla_\xi h\left(\xi, \mathbf{m}_{t,c_t}\right)|_{\xi=\xi_t^-} \qquad \mathbf{H}_{t,\mathbf{m}_{c_t}} = \nabla_{\mathbf{m}_{c_t}} h\left(\xi_t, \mathbf{m}_{c_t}\right)|_{\mathbf{m}_{c_t}=\mathbf{m}_{t,c_t}^-}$$

Evolution of State Estimate on Correction

- Let's have a closer look at the Kalman gain

$$\mathbf{K}_t = \Sigma_t^- \mathbf{H}_t^\top \left(\boxed{\mathbf{H}_t \Sigma_t^- \mathbf{H}_t^\top} + \Sigma_{m_t} \right)^{-1} \qquad \mathbf{H}_t = \nabla_{\mathbf{x}} h(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_t^-}$$

- The matrix $\mathbf{H}_t \Sigma_t^- \mathbf{H}_t^\top$ only involves covariances between pose and the measured landmark:

$$\mathbf{H}_t \Sigma_t^- \mathbf{H}_t^\top = \mathbf{H}_{t,\xi} \Sigma_{t,\xi\xi}^- \mathbf{H}_{t,\xi}^\top + \mathbf{H}_{t,\mathbf{m}_{c_t}} \Sigma_{t,\mathbf{m}_{c_t}\xi}^- \mathbf{H}_{t,\xi}^\top + \mathbf{H}_{t,\xi} \Sigma_{t,\xi\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top + \mathbf{H}_{t,\mathbf{m}_{c_t}} \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top$$

$$\Sigma_t = \begin{pmatrix} \Sigma_{t,\xi\xi}^- & \Sigma_{t,\xi\mathbf{m}_1}^- & \cdots & \Sigma_{t,\xi\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\xi\mathbf{m}_S}^- \\ \Sigma_{t,\mathbf{m}_1\xi}^- & \Sigma_{t,\mathbf{m}_1\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_1\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_1\mathbf{m}_S}^- \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \Sigma_{t,\mathbf{m}_{c_t}\xi}^- & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_S}^- \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \Sigma_{t,\mathbf{m}_S\xi}^- & \Sigma_{t,\mathbf{m}_S\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_S\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_S\mathbf{m}_S}^- \end{pmatrix}$$

Evolution of State Estimate on Correction

$$\mathbf{K}_t = \boxed{\Sigma_t^- \mathbf{H}_t^\top} \left(\mathbf{H}_t \Sigma_t^- \mathbf{H}_t^\top + \Sigma_{m_t} \right)^{-1} \qquad \mathbf{H}_t = \nabla_{\mathbf{x}} h(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_t^-}$$

- The matrix $\Sigma_t^- \mathbf{H}_t^\top$ stacks the covariances between the pose/the measured landmark and all state variables (pose+landmarks)

$$\Sigma_t^- \mathbf{H}_t^\top = \begin{pmatrix} \Sigma_{t,\xi\xi}^- \mathbf{H}_{t,\xi}^\top + \Sigma_{t,\xi\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top \\ \Sigma_{t,\mathbf{m}_1\xi}^- \mathbf{H}_{t,\xi}^\top + \Sigma_{t,\mathbf{m}_1\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top \\ \vdots \\ \Sigma_{t,\mathbf{m}_S\xi}^- \mathbf{H}_{t,\xi}^\top + \Sigma_{t,\mathbf{m}_S\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top \end{pmatrix}$$

$$\Sigma_t = \begin{pmatrix} \Sigma_{t,\xi\xi}^- & \Sigma_{t,\xi\mathbf{m}_1}^- & \cdots & \Sigma_{t,\xi\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\xi\mathbf{m}_S}^- \\ \Sigma_{t,\mathbf{m}_1\xi}^- & \Sigma_{t,\mathbf{m}_1\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_1\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_1\mathbf{m}_S}^- \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \Sigma_{t,\mathbf{m}_{c_t}\xi}^- & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_S}^- \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \Sigma_{t,\mathbf{m}_S\xi}^- & \Sigma_{t,\mathbf{m}_S\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_S\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_S\mathbf{m}_S}^- \end{pmatrix}$$

Evolution of State Estimate on Correction

- Hence, the Kalman gain distributes information onto all state dimensions that are correlated with the pose or the measured landmark

$$\mathbf{K}_t = \Sigma_t^- \mathbf{H}_t^\top \left(\mathbf{H}_t \Sigma_t^- \mathbf{H}_t^\top + \Sigma_{m_t} \right)^{-1}$$

$$\Sigma_t^- \mathbf{H}_t^\top = \begin{pmatrix} \Sigma_{t,\xi\xi}^- \mathbf{H}_{t,\xi}^\top + \Sigma_{t,\xi\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top \\ \Sigma_{t,\mathbf{m}_1\xi}^- \mathbf{H}_{t,\xi}^\top + \Sigma_{t,\mathbf{m}_1\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top \\ \vdots \\ \Sigma_{t,\mathbf{m}_S\xi}^- \mathbf{H}_{t,\xi}^\top + \Sigma_{t,\mathbf{m}_S\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top \end{pmatrix}$$

- The correction step updates all state dimensions in the mean that are correlated with the pose or measured landmark

$$\mathbf{x}_t^+ = \mathbf{x}_t^- + \mathbf{K}_t \left(\mathbf{y}_t - h\left(\mathbf{x}_t^-\right) \right)$$

Evolution of State Estimate on Correction

- How is the state covariance updated in the correction step?

$$\Sigma_t^+ = (\mathbf{I} - \boxed{\mathbf{K}_t \mathbf{H}_t)\, \Sigma_t^-}$$

$$\begin{pmatrix} \mathbf{K}_{t,\xi}\mathbf{H}_{t,\xi} & 0 & \ldots & 0 & \mathbf{K}_{t,\xi}\mathbf{H}_{t,\mathbf{m}_{c_t}} & 0 & \ldots & 0 \\ \mathbf{K}_{t,\mathbf{m}_1}\mathbf{H}_{t,\xi} & 0 & \ldots & 0 & \mathbf{K}_{t,\mathbf{m}_1}\mathbf{H}_{t,\mathbf{m}_{c_t}} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{K}_{t,\mathbf{m}_S}\mathbf{H}_{t,\xi} & 0 & \ldots & 0 & \mathbf{K}_{t,\mathbf{m}_S}\mathbf{H}_{t,\mathbf{m}_{c_t}} & 0 & \ldots & 0 \end{pmatrix} \begin{pmatrix} \Sigma_{t,\xi\xi}^- & \Sigma_{t,\xi\mathbf{m}_1}^- & \cdots & \Sigma_{t,\xi\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\xi\mathbf{m}_S}^- \\ \Sigma_{t,\mathbf{m}_1\xi}^- & \Sigma_{t,\mathbf{m}_1\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_1\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_1\mathbf{m}_S}^- \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \Sigma_{t,\mathbf{m}_{c_t}\xi}^- & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_{c_t}\mathbf{m}_S}^- \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \Sigma_{t,\mathbf{m}_S\xi}^- & \Sigma_{t,\mathbf{m}_S\mathbf{m}_1}^- & \cdots & \Sigma_{t,\mathbf{m}_S\mathbf{m}_{c_t}}^- & \cdots & \Sigma_{t,\mathbf{m}_S\mathbf{m}_S}^- \end{pmatrix}$$

- Covariance change for a landmark that is not the measured landmark:

$$\Sigma_{t,\mathbf{m}_1\mathbf{m}_{c_t}}^+ =$$

$$\mathbf{K}_{t,\mathbf{m}_1} = \left(\boxed{\Sigma_{t,\mathbf{m}_1\xi}^-} \mathbf{H}_{t,\xi}^\top + \Sigma_{t,\mathbf{m}_1\mathbf{m}_{c_t}}^- \mathbf{H}_{t,\mathbf{m}_{c_t}}^\top \right) \left(\mathbf{H}_t \Sigma_t^- \mathbf{H}_t^\top + \Sigma_{m_t} \right)^{-1}$$

non-zero!

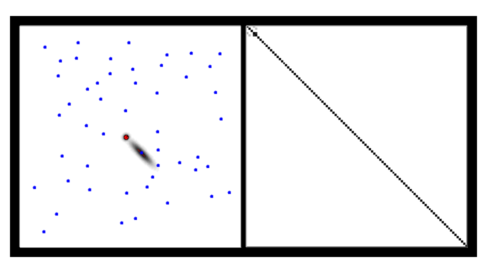
Evolution of State Estimate on Correction

- The correction step updates all state dimensions in the state covariance that correlate with the pose or measured landmark

$$\Sigma_t^+ = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)\, \Sigma_t^-$$

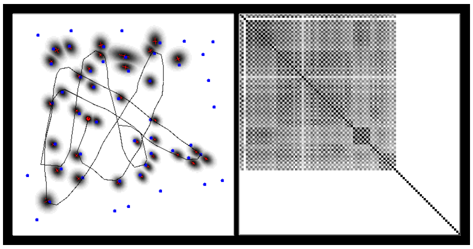
- Since all landmarks are correlated with pose, all landmark correlations with the measured landmark get updated
- Hence, all state variables become correlated: The state covariance is dense!
- Measurement information propagates on all landmarks along the trajectory

Example Evolution of the Covariance



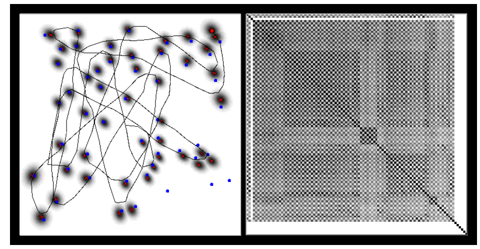
Pose and map Correlation matrix

Example Evolution of the Covariance



Pose and map | Correlation matrix

Example Evolution of the Covariance

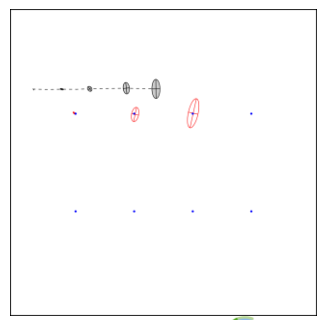


Pose and map | Correlation matrix

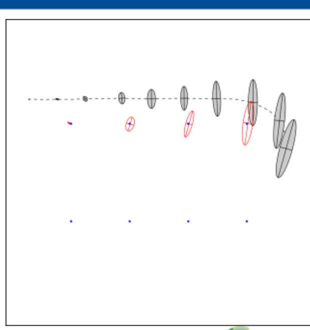
Topics of This Lecture

- Recap: Online SLAM methods
- EKF SLAM
  - Extended Kalman Filter formulation
  - 2D EKF SLAM example
  - Detailed analysis
- **Loop Closure**
- Case study: MonoSLAM
- Full SLAM methods
  - SLAM graph optimization
  - Pose graph optimization

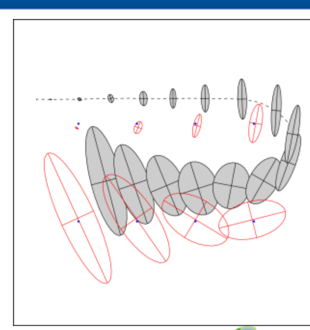
Closing a Loop



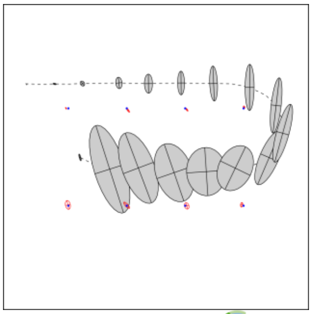
Closing a Loop



Closing a Loop



Closing a Loop



Slide credit: Jörg Stückler

Image: Michael Montemerlo

Closing a Loop

- Effect of loop closure
  - On loop closure, old landmarks in the map get reobserved
  - Strong correlations are added between older parts of the map that were not observed for some time and the current pose / recently observed landmarks
  - Pose and landmarks are corrected to make the estimate more consistent with the reobservation
- Loop closure reduces uncertainty in pose and landmark estimates
  - High certainty in the old part of the map propagates to current pose and recent landmark estimates
  - But: wrong correspondences can lead to divergence towards a wrong estimate!

Slide credit: Jörg Stückler

Topics of This Lecture

- Recap: Online SLAM methods
- EKF SLAM
  - Extended Kalman Filter formulation
  - 2D EKF SLAM example
  - Detailed analysis
- Loop Closure
- Case study: MonoSLAM
- Full SLAM methods
  - SLAM graph optimization
  - Pose graph optimization

MonoSLAM: Monocular EKF-SLAM

Slide credit: Jörg Stückler

Video source: Youtube / A. J. Davison

MonoSLAM: State Parametrization

- Camera motion

$$\xi_t = \begin{pmatrix} \mathbf{p}_t \\ \mathbf{q}_t \\ \mathbf{v}_t \\ \boldsymbol{\omega}_t \end{pmatrix}$$

3D position in world frame
Quaternion for rotation from camera to world frame
Linear velocity in world frame
Angular velocity of camera in world frame

- Landmarks

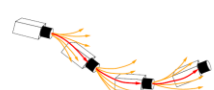
$$\mathbf{m}_{t,j} = \begin{pmatrix} m_{t,j,x} \\ m_{t,j,y} \\ m_{t,j,z} \end{pmatrix}$$

3D position in world frame

Slide credit: Jörg Stückler

MonoSLAM: State Transition Model

- 6-DoF camera dynamics model (constant-velocity)



$$\xi_t = g_{\xi}(\xi_{t-1}) = \begin{pmatrix} \mathbf{p}_{t-1} + (\mathbf{v}_{t-1} + \epsilon_{\mathbf{v},t})\,\Delta t \\ \mathbf{q}_{t-1}\mathbf{q}\left((\boldsymbol{\omega}_{t-1} + \epsilon_{\omega,t})\,\Delta t\right) \\ \mathbf{v}_{t-1} + \epsilon_{\mathbf{v},t} \\ \boldsymbol{\omega}_{t-1} + \epsilon_{\omega,t} \end{pmatrix}$$

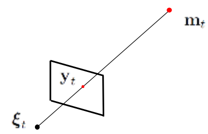
Gaussian noise

- Map remains static, $\mathbf{m}_t = g_{\mathbf{m}}(\mathbf{m}_{t-1}) = \mathbf{m}_{t-1}$

Slide credit: Jörg Stückler

MonoSLAM: Observation Model

- Bearing-only observation model
  - Depth is not measured in a monocular image

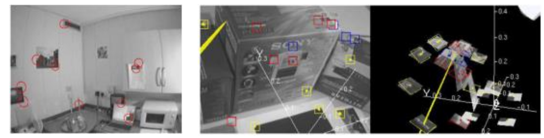


  - Landmark observation model

$$\bar{y}_t = h(\xi_t, m_{t,c_t}) + \delta_t = C\pi\left(R(q_t)^\top (m_{t,c_t} - p_t)\right) + \delta_t \qquad \delta_t \sim \mathcal{N}(0, \Sigma_{m_t})$$

- MonoSLAM additionally considers the radial distortion in a wide-angle camera image using an analytically invertible model

MonoSLAM: Data Association



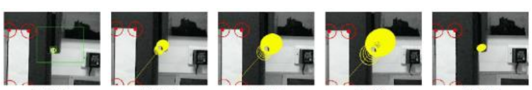
- Active search:
  - Likely region of measurement from innovation covariance

$$H_t \Sigma_t^- H_t^\top + \Sigma_{m_t}$$

- Correspondence measure
  - Matching of small image patches (e.g., 9×9 to 15×15)
  - Projective warping using a patch normal estimate
  - Sum of squared intensity differences

MonoSLAM: Map Maintenance

- Heuristics to keep number of visible landmarks from any camera view point small (~12 landmarks)



- Special depth initialization for new landmark with a particle filter
- Map initialized with landmarks on a known 3D pattern
  - Sets metric scale
  - Good initial state for tracking
  - Stable pose for adding new landmarks

Summary: Online SLAM

- Online SLAM methods marginalize out past trajectory
- Tracking-and-Mapping approaches
  - Alternate optimization on map and camera pose estimate
  - Condition optimization of one estimate on the other
- Extended Kalman Filters can be used for online SLAM
  - Maintains correlations between camera pose and all landmarks
  - Quadratic update run-time complexity limits map size
- MonoSLAM:
  - Implements Visual EKF-SLAM for monocular cameras
  - Data association via active search and patch correlation

Topics of This Lecture

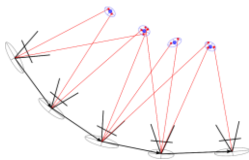
- Recap: Online SLAM methods
- EKF SLAM
  - Extended Kalman Filter formulation
  - 2D EKF SLAM example
  - Detailed analysis
- Loop Closure
- Case study: MonoSLAM
- Full SLAM methods
  - SLAM graph optimization
  - Pose graph optimization

Online SLAM vs. Full SLAM

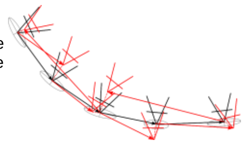
- Online SLAM
  - Only optimizes current camera pose (+ landmarks) with current measurements
  - Old pose estimates are not improved using newer measurements
  - At each time step, a linearization is performed at the current pose and landmark estimates to update the correlations of state variables
  - Linearization points are fixed while state estimates change later, correlations are not updated
    - No compensation for corrected estimates of landmark positions
    - No improvement of old pose estimates and reconsideration for linearization
- Full SLAM
  - Optimize for whole trajectory and all landmarks in the map at once
  - Uses „future" measurements as well to update „past" poses
  - Allows for relinearization of all state-transitions and measurements at each optimization step

Full SLAM Approaches

- SLAM graph optimization:
  - Joint optimization for poses and map elements from image observations of map elements and control inputs

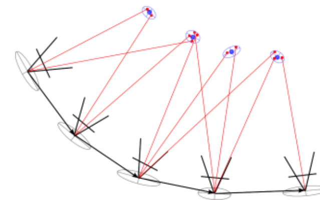


- Pose graph optimization:
  - Optimization of poses from relative pose constraints deduced from the image observations
  - Map recovered from the optimized poses



SLAM Graph Optimization

- Joint optimization for poses and map elements from image observations of map elements
  - Common map element observations induce constraints between the poses
  - Map elements correlate with each other through the common poses that observe them



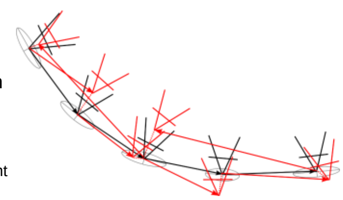
  - Without control inputs: Bundle Adjustment

Bundle Adjustment Example

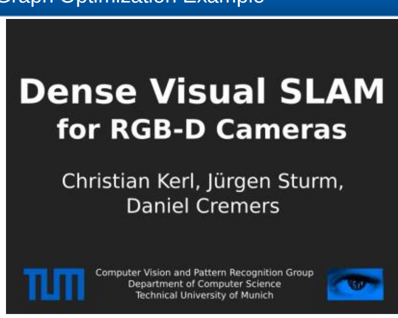
Agarwal et al., Building Rome in a Day. ICCV 2009, „Dubrovnik" image set

Pose Graph Optimization

- Optimization of poses
  - From relative pose constraints deduced from the image observations
  - Map recovered from the optimized poses
- Deduce relative constraints between poses from image observations, e.g.,
  - 8-point algorithm
  - Direct image alignment



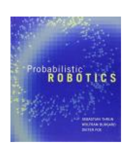
Pose Graph Optimization Example



Kerl et al., Dense Visual SLAM for RGB-D Cameras, IROS 2013

References and Further Reading

- Probabilistic Robotics textbook



Probabilistic Robotics, S. Thrun, W. Burgard, D. Fox, MIT Press, 2005

- Research papers
  - A.J. Davison et al., MonoSLAM: Real-Time Single Camera SLAM. IEEE Transaction on Patterm Analysis and Machine Intelligence, 2007
  - G. Klein and D. Murray, Parallel Tracking and Mapping for Small AR Workspaces. Int. Symposium on Mixed and Augmented Reality 2007