Robot Mapping

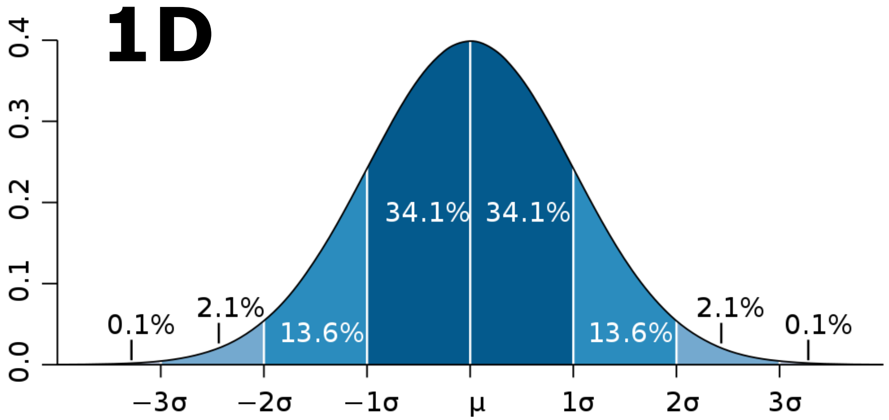
Extended Information Filter

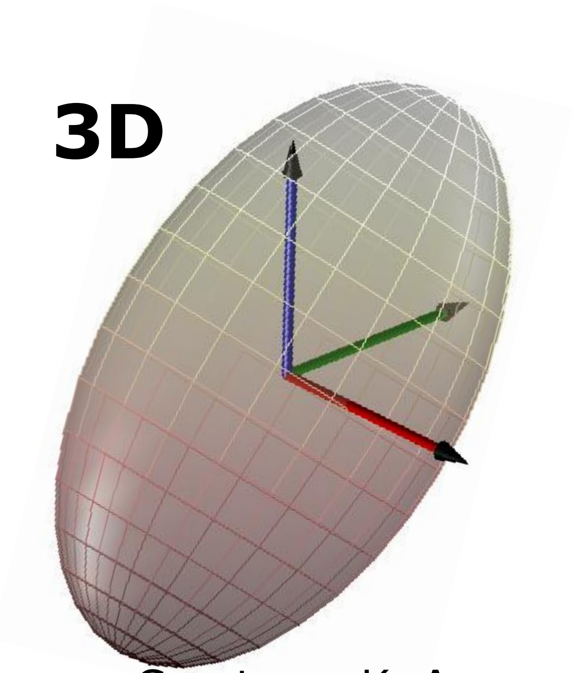
Gian Diego Tipaldi, Wolfram Burgard

Gaussians

- Gaussian described by **moments** μ, Σ

$$p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$





Courtesy: K. Arras

Canonical Parameterization

- Alternative representation for Gaussians
- Described by **information matrix** Ω and **information vector** ξ

Canonical Parameterization

- Alternative representation for Gaussians
- Described by **information matrix** Ω

$$\Omega = \Sigma^{-1}$$

- and **information vector** ξ

$$\xi = \Sigma^{-1}\mu$$

Complete Parameterizations

moments

$$\Sigma = \Omega^{-1}$$

$$\mu = \Omega^{-1}\xi$$

canonical

$$\Omega = \Sigma^{-1}$$

$$\xi = \Sigma^{-1}\mu$$

Towards the Information Form

$$
\begin{aligned}
& p(x) \\
= \quad & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)
\end{aligned}
$$

Towards the Information Form

$$\begin{aligned} p(x) & \\ &= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big) \\ &= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu - \frac{1}{2}\mu^T\Sigma^{-1}\mu\Big) \end{aligned}$$

Towards the Information Form

$$
\begin{aligned}
p(x) & \\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) \\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(\underline{-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu} - \underline{\frac{1}{2}\mu^T\Sigma^{-1}\mu}\right) \\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mu^T\Sigma^{-1}\mu\right) \\
&\qquad \exp\left(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu\right)
\end{aligned}
$$

Towards the Information Form

$$
\begin{aligned}
& p(x) \\
= \quad & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) \right) \\
= \quad & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu - \frac{1}{2}\mu^T\Sigma^{-1}\mu \right) \\
= \quad & \underline{\det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mu^T\Sigma^{-1}\mu \right)} \\
& \qquad \exp\left(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu \right) \\
= \quad & \eta \exp\left(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu \right)
\end{aligned}
$$

Towards the Information Form

$$
\begin{aligned}
p(x) & \\
= \; & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big) \\
= \; & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu - \frac{1}{2}\mu^T\Sigma^{-1}\mu\Big) \\
= \; & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}\mu^T\Sigma^{-1}\mu\Big) \\
& \qquad \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu\Big) \\
= \; & \eta \exp\Big(-\frac{1}{2}x^T\underline{\Sigma^{-1}}x + x^T\underline{\Sigma^{-1}\mu}\Big) \\
= \; & \eta \exp\Big(-\frac{1}{2}x^T\Omega x + x^T\xi\Big)
\end{aligned}
$$

Dual Representation

$$p(x) = \frac{\exp(-\frac{1}{2}\mu^T\xi)}{\det(2\pi\Omega^{-1})^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}x^T\Omega x + x^T\xi\Big)$$

canonical parameterization

$$p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big)$$

moments parameterization

Marginalization and Conditioning

$$p(\boldsymbol{\alpha},\boldsymbol{\beta}) = \mathcal{N}\left(\begin{bmatrix}\boldsymbol{\mu}_\alpha \\ \boldsymbol{\mu}_\beta\end{bmatrix}, \begin{bmatrix}\Sigma_{\alpha\alpha} & \Sigma_{\alpha\beta} \\ \Sigma_{\beta\alpha} & \Sigma_{\beta\beta}\end{bmatrix}\right) = \mathcal{N}^{-1}\left(\begin{bmatrix}\boldsymbol{\eta}_\alpha \\ \boldsymbol{\eta}_\beta\end{bmatrix}, \begin{bmatrix}\Lambda_{\alpha\alpha} & \Lambda_{\alpha\beta} \\ \Lambda_{\beta\alpha} & \Lambda_{\beta\beta}\end{bmatrix}\right)$$

	MARGINALIZATION	CONDITIONING
	$p(\boldsymbol{\alpha}) = \int p(\boldsymbol{\alpha},\boldsymbol{\beta})\,d\boldsymbol{\beta}$	$p(\boldsymbol{\alpha} \mid \boldsymbol{\beta}) = p(\boldsymbol{\alpha},\boldsymbol{\beta})/p(\boldsymbol{\beta})$
COV. FORM	$\boldsymbol{\mu} = \boldsymbol{\mu}_\alpha$ $\Sigma = \Sigma_{\alpha\alpha}$	$\boldsymbol{\mu}' = \boldsymbol{\mu}_\alpha + \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)$ $\Sigma' = \Sigma_{\alpha\alpha} - \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\alpha}$
INFO. FORM	$\boldsymbol{\eta} = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\boldsymbol{\eta}_\beta$ $\Lambda = \Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\Lambda_{\beta\alpha}$	$\boldsymbol{\eta}' = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\boldsymbol{\beta}$ $\Lambda' = \Lambda_{\alpha\alpha}$

trivial

expensive

Courtesy: R. Eustice

From the Kalman Filter to the Information Filter

- Two parameterization for Gaussian
- Same expressiveness
- Marginalization and conditioning have different complexities
- We learned about Gaussian filtering with the Kalman filter in Chapter 4
- Kalman filtering in information from is called information filtering

Kalman Filter Algorithm

$$\textbf{Kalman\_filter}(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)\textbf{:}$$

$$\bar{\mu}_t = A_t\, \mu_{t-1} + B_t\, u_t$$
$$\bar{\Sigma}_t = A_t\, \Sigma_{t-1}\, A_t^T + R_t$$

$$K_t = \bar{\Sigma}_t\, C_t^T (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1}$$
$$\mu_t = \bar{\mu}_t + K_t(z_t - C_t\, \bar{\mu}_t)$$
$$\Sigma_t = (I - K_t\, C_t)\, \bar{\Sigma}_t$$
$$return\ \mu_t, \Sigma_t$$

Prediction Step (1)

- Transform $\bar{\Sigma}_t = A_t \, \Sigma_{t-1} \, A_t^T + R_t$
- Using $\Sigma_{t-1} = \Omega_{t-1}^{-1}$
- Leads to

$$\bar{\Omega}_t \quad = \quad (A_t \, \Omega_{t-1}^{-1} \, A_t^T + R_t)^{-1}$$

Prediction Step (2)

- Transform $\bar{\mu}_t = A_t \, \mu_{t-1} + B_t \, u_t$
- Using $\bar{\mu}_{t-1} = \Omega_{t-1}^{-1} \xi_{t-1}$
- Leads to

$$\begin{aligned} \bar{\xi}_t &= \bar{\Omega}_t (A_t \, \mu_{t-1} + B_t \, u_t) \\ &= \bar{\Omega}_t (A_t \, \Omega_{t-1}^{-1} \xi_{t-1} + B_t \, u_t) \end{aligned}$$

Information Filter Algorithm

$$\textbf{Information\_filter}(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)\textbf{:}$$

$$\bar{\Omega}_t = (A_t\, \Omega_{t-1}^{-1}\, A_t^T + R_t)^{-1}$$

$$\bar{\xi}_t = \bar{\Omega}_t (A_t\, \Omega_{t-1}^{-1}\, \xi_{t-1} + B_t\, u_t)$$

Correction Step

- Use the Bayes filter measurement update and replace the components

$$bel(x_t) = \eta \, p(z_t \mid x_t) \, \overline{bel}(x_t)$$

$$= \quad \eta' \, \exp\left(-\frac{1}{2} \, (z_t - C_t x_t)^T \, Q_t^{-1} \, (z_t - C_t x_t)\right) \, \exp\left(-\frac{1}{2} \, (x_t - \bar{\mu}_t)^T \, \bar{\Sigma}_t^{-1} \, (x_t - \bar{\mu}_t)\right)$$

Correction Step

- Use the Bayes filter measurement update and replace the components

$$
\begin{aligned}
bel(x_t) &= \eta \, p(z_t \mid x_t) \, \overline{bel}(x_t) \\
&= \eta' \exp\left(-\frac{1}{2} (z_t - C_t x_t)^T Q_t^{-1} (z_t - C_t x_t)\right) \exp\left(-\frac{1}{2} (x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1} (x_t - \bar{\mu}_t)\right) \\
&= \eta' \exp\left(-\frac{1}{2} (z_t - C_t x_t)^T Q_t^{-1} (z_t - C_t x_t) - \frac{1}{2} (x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1} (x_t - \bar{\mu}_t)\right)
\end{aligned}
$$

Correction Step

- Use the Bayes filter measurement update and replace the components

$$
\begin{aligned}
bel(x_t) &= \eta\, p(z_t \mid x_t)\, \overline{bel}(x_t) \\
&= \eta' \exp\left(-\frac{1}{2}\,(z_t - C_t x_t)^T\, Q_t^{-1}\,(z_t - C_t x_t)\right) \exp\left(-\frac{1}{2}\,(x_t - \bar{\mu}_t)^T\, \bar{\Sigma}_t^{-1}\,(x_t - \bar{\mu}_t)\right) \\
&= \eta' \exp\left(-\frac{1}{2}\,(z_t - C_t x_t)^T\, Q_t^{-1}\,(z_t - C_t x_t) - \frac{1}{2}\,(x_t - \bar{\mu}_t)^T\, \bar{\Sigma}_t^{-1}\,(x_t - \bar{\mu}_t)\right) \\
&= \eta'' \exp\left(-\frac{1}{2}\, x_t^T\, C_t^T\, Q_t^{-1}\, C_t\, x_t + x_t^T\, C_t^T\, Q_t^{-1}\, z_t - \frac{1}{2}\, x_t^T\, \bar{\Omega}_t x_t + x_t^T \bar{\xi}_t\right)
\end{aligned}
$$

Correction Step

- Use the Bayes filter measurement update and replace the components

$$
\begin{aligned}
bel(x_t) &= \eta \, p(z_t \mid x_t) \, \overline{bel}(x_t) \\
&= \eta' \; \exp\left(-\frac{1}{2} \, (z_t - C_t x_t)^T \; Q_t^{-1} \; (z_t - C_t x_t)\right) \; \exp\left(-\frac{1}{2} \, (x_t - \bar{\mu}_t)^T \; \bar{\Sigma}_t^{-1} \; (x_t - \bar{\mu}_t)\right) \\
&= \eta' \; \exp\left(-\frac{1}{2} \, (z_t - C_t x_t)^T \; Q_t^{-1} \; (z_t - C_t x_t) - \frac{1}{2} \, (x_t - \bar{\mu}_t)^T \; \bar{\Sigma}_t^{-1} \; (x_t - \bar{\mu}_t)\right) \\
&= \eta'' \; \exp\left(-\frac{1}{2} \, x_t^T \; C_t^T \; Q_t^{-1} \; C_t \; x_t + x_t^T \; C_t^T \; Q_t^{-1} \; z_t - \frac{1}{2} \, x_t^T \; \bar{\Omega}_t x_t + x_t^T \bar{\xi}_t\right) \\
&= \eta'' \; \exp\left(-\frac{1}{2} \, x_t^T \; \underbrace{\left[C_t^T \; Q_t^{-1} \; C_t + \bar{\Omega}_t\right]}_{\Omega_t} \; x_t + x_t^T \; \underbrace{\left[C_t^T \; Q_t^{-1} \; z_t + \bar{\xi}_t\right]}_{\xi_t}\right)
\end{aligned}
$$

Correction Step

- This results in a simple update rule

$$bel(x_t) \quad = \quad \eta \exp\left(-\frac{1}{2}\, x_t^T \, \underbrace{[C_t^T \, Q_t^{-1} \, C_t + \bar{\Omega}_t]}_{\Omega_t} \, x_t + x_t^T \, \underbrace{[C_t^T \, Q_t^{-1} \, z_t + \bar{\xi}_t]}_{\xi_t} \right)$$

$$\Downarrow$$

$$\begin{aligned} \Omega_t &= C_t^T \, Q_t^{-1} \, C_t + \bar{\Omega}_t \\ \xi_t &= C_t^T \, Q_t^{-1} \, z_t + \bar{\xi}_t \end{aligned}$$

Information Filter Algorithm

1: $\textbf{Information\_filter}(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)\textbf{:}$

2: $\bar{\Omega}_t = (A_t\ \Omega_{t-1}^{-1}\ A_t^T + R_t)^{-1}$

3: $\bar{\xi}_t = \bar{\Omega}_t (A_t\ \Omega_{t-1}^{-1}\ \xi_{t-1} + B_t\ u_t)$

4: $\Omega_t = C_t^T\ Q_t^{-1}\ C_t + \bar{\Omega}_t$

5: $\xi_t = C_t^T\ Q_t^{-1}\ z_t + \bar{\xi}_t$

6: $return\ \xi_t, \Omega_t$

Prediction and Correction

- Prediction

$$
\begin{aligned}
\bar{\Omega}_t &= (A_t\, \Omega_{t-1}^{-1}\, A_t^T + R_t)^{-1} \\
\bar{\xi}_t &= \bar{\Omega}_t(A_t\, \Omega_{t-1}^{-1}\, \xi_{t-1} + B_t\, u_t)
\end{aligned}
$$

- Correction

$$
\begin{aligned}
\Omega_t &= C_t^T\, Q_t^{-1}\, C_t + \bar{\Omega}_t \\
\xi_t &= C_t^T\, Q_t^{-1}\, z_t + \bar{\xi}_t
\end{aligned}
$$

Discuss differences to the KF!

Complexity

- Kalman filter
  - Efficient prediction step: $\mathcal{O}(n^2)^*$
  - Costly correction step: $\mathcal{O}(n^2 + k^{2.4})$
- Information filter
  - Costly prediction step: $\mathcal{O}(n^{2.4})$
  - Efficient correction step: $\mathcal{O}(n^2)^*$
- Transformation between both parameterizations is costly: $\mathcal{O}(n^{2.4})$

*Potentially faster, especially for SLAM; depending on type of controls and observations

Extended Information Filter

- As the Kalman filter, the information filter suffers from the linear models
- The extended information filter (EIF) uses a similar trick as the EKF
- Linearization of the motion and observation function

Linearization of the EIF

- Taylor approximation analog to the EKF (see Chapter 4)

$$
\begin{aligned}
g(u_t, x_{t-1}) &\approx g(u_t, \mu_{t-1}) + G_t \, (x_{t-1} - \mu_{t-1}) \\
h(x_t) &\approx h(\bar{\mu}_t) + H_t \, (x_t - \bar{\mu}_t)
\end{aligned}
$$

- with the Jacobians G_t and H_t

Prediction: From EKF of EIF

- Substitution of the moments brings us from the EKF

$$\begin{aligned} \bar{\Sigma}_t &= G_t \, \Sigma_{t-1} \, G_t^T + R_t \\ \bar{\mu}_t &= g(u_t, \mu_{t-1}) \end{aligned}$$

- to the EIF

$$\begin{aligned} \bar{\Omega}_t &= (G_t \, \Omega_{t-1}^{-1} \, G_t^T + R_t)^{-1} \\ \bar{\xi}_t &= \bar{\Omega}_t \, g(u_t, \Omega_{t-1}^{-1} \, \xi_{t-1}) \end{aligned}$$

Prediction: From EKF of EIF

1: **Extended_Kalman_filter**$(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)$**:**

2: $\bar{\mu}_t = g(u_t, \mu_{t-1})$

3: $\bar{\Sigma}_t = G_t \, \Sigma_{t-1} \, G_t^T + R_t$

1: **Extended_information_filter**$(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$**:**

2: $\mu_{t-1} = \Omega_{t-1}^{-1} \, \xi_{t-1}$

3: $\bar{\Omega}_t = (G_t \, \Omega_{t-1}^{-1} \, G_t^T + R_t)^{-1}$

4: $\bar{\mu}_t = g(u_t, \mu_{t-1})$

5: $\bar{\xi}_t = \bar{\Omega}_t \, \bar{\mu}_t$

Correction Step of the EIF

- As from the KF to IF transition, use substitute the moments in the measurement update

$$
\begin{aligned}
bel(x_t) \quad &= \quad \eta \, \exp\left(-\frac{1}{2} \left(z_t - h(\bar{\mu}_t) - H_t \, (x_t - \bar{\mu}_t) \right)^T Q_t^{-1} \right. \\
&\qquad \left. \left(z_t - h(\bar{\mu}_t) - H_t \, (x_t - \bar{\mu}_t) \right) - \frac{1}{2} (x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1} (x_t - \bar{\mu}_t) \right)
\end{aligned}
$$

- This leads to

$$
\begin{aligned}
\Omega_t \quad &= \quad \bar{\Omega}_t + H_t^T \, Q_t^{-1} \, H_t \\
\xi_t \quad &= \quad \bar{\xi}_t + H_t^T \, Q_t^{-1} \, (z_t - h(\bar{\mu}_t) + H_t \, \bar{\mu}_t)
\end{aligned}
$$

Extended Information Filter

Extended_information_filter$(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$**:**

$$\mu_{t-1} = \Omega_{t-1}^{-1}\ \xi_{t-1}$$

$$\bar{\Omega}_t = (G_t\ \Omega_{t-1}^{-1}\ G_t^T + R_t)^{-1}$$

$$\bar{\mu}_t = g(u_t, \mu_{t-1})$$

$$\bar{\xi}_t = \bar{\Omega}_t\ \bar{\mu}_t$$

$$\Omega_t = \bar{\Omega}_t + H_t^T\ Q_t^{-1}\ H_t$$

$$\xi_t = \bar{\xi}_t + H_t^T\ Q_t^{-1}\ (z_t - h(\bar{\mu}_t) + H_t\ \bar{\mu}_t)$$

return ξ_t, Ω_t

EIF vs. EKF

- The EIF is the EKF in information form
- Complexities of the prediction and correction steps can differ
- Same expressiveness than the EKF
- Unscented transform can also be used
- Reported to be numerically more stable than the EKF
- In practice, the EKF is more popular than the EIF

Summary

- Gaussians can also be represented using the canonical parameterization
- Allow for filtering in information form
- Information filter vs. Kalman filter
- KF: efficient prediction, slow correction
- IF: slow prediction, efficient correction
- The application determines which filter is the better choice!

Literature

Extended Information Filter

- Thrun et al.: “Probabilistic Robotics”, Chapter 3.5

Slide Information

- These slides have been created by Cyrill Stachniss as part of the robot mapping course taught in 2012/13 and 2013/14. I created this set of slides partially extending existing material of Edwin Olson, Pratik Agarwal, and myself.
- I tried to acknowledge all people that contributed image or video material. In case I missed something, please let me know. If you adapt this course material, please make sure you keep the acknowledgements.
- Feel free to use and change the slides. If you use them, I would appreciate an acknowledgement as well. To satisfy my own curiosity, I appreciate a short email notice in case you use the material in your course.
- My video recordings are available through YouTube:
  http://www.youtube.com/playlist?list=PLgnQpQtFTOGQrZ4O5QzbIHgl3b1JHimN_&feature=g-list

Cyrill Stachniss, 2014
cyrill.stachniss@igg.uni-bonn.de