

Fuzja sygnałów i filtry bayesowskie

Roboty Manipulacyjne i Mobilne

dr inż. Janusz Jakubiak

Katedra Cybernetyki i Robotyki
Wydział Elektroniki, Politechnika Wrocławska

Wrocław, 10.03.2015

Dlaczego potrzebna jest fuzja sygnałów?

Rodzaje informacji:

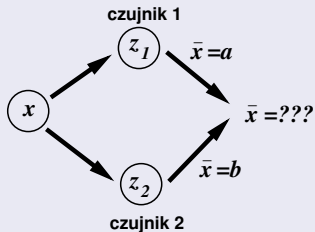
- rozłączne,
- komplementarne,
- substytucyjne.

Dlaczego potrzebna jest fuzja sygnałów?

Rodzaje informacji:

- rozłączne,
- komplementarne,
- substytucyjne.

Problem:



Dlaczego potrzebna jest fuzja sygnałów?

Fuzja sygnałów

Przetworzenie sygnałów pochodzących z różnych czujników w taki sposób, że wynikowa informacja jest lepsza od informacji możliwych do uzyskania z każdego z czujników oddzielnie.

Dlaczego potrzebna jest fuzja sygnałów?

Fuzja sygnałów

Przetworzenie sygnałów pochodzących z różnych czujników w taki sposób, że wynikowa informacja jest lepsza od informacji możliwych do uzyskania z każdego z czujników oddzielnie.

Rozróżnia się fuzje

- *bezpośrednią*, nazywaną także *fuzją sensorów*, w której przetwarzane są wyłącznie dane z czujników oraz
- *pośrednią (fuzja informacji)*, w której wykorzystywane są także informacje z innych źródeł (np. wiedza *a priori*, informacje od użytkownika).

Notacja

- x – stan rzeczywisty układu (R^n), \bar{x} – estymata stanu,
- z – obserwacja (R^m),

- x – stan rzeczywisty układu (R^n), \bar{x} – estymata stanu,
- z – obserwacja (R^m),
- $p(A|B)$ – prawdopodobieństwo zdarzenia A pod warunkiem B ,
- $\mathcal{N}(\mu, \Sigma)$ – wielowymiarowy rozkład normalny ze średnią μ i macierzą kowariancji Σ (parametrami: momenty)

$$\mathcal{N}(\mu, \Sigma) \sim p(X) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\}$$

- x – stan rzeczywisty układu (R^n), \bar{x} – estymata stanu,
- z – obserwacja (R^m),
- $p(A|B)$ – prawdopodobieństwo zdarzenia A pod warunkiem B ,
- $\mathcal{N}(\mu, \Sigma)$ – wielowymiarowy rozkład normalny ze średnią μ i macierzą kowariancji Σ (parametrami: momenty)

$$\mathcal{N}(\mu, \Sigma) \sim p(X) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp \left\{ -\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu) \right\}$$

- Twierdzenie Bayesa

$$p(X|Z) = \frac{p(Z|X)p(X)}{p(Z)} = \eta p(Z|X)p(X),$$

gdzie η jest współczynnikiem normalizującym.

Notacja (2)

Zmiany momentów dystrybucji gaussowskich przy przekształceniach

- liniowym ($Y = AX + B$)

$$p(X) = \mathcal{N}(\mu, \Sigma) \quad \rightarrow \quad p(Y) = \mathcal{N}(A\mu + B, A\Sigma A^T)$$

Notacja (2)

Zmiany momentów dystrybucji gaussowskich przy przekształceniach

- liniowym ($Y = AX + B$)

$$p(X) = \mathcal{N}(\mu, \Sigma) \quad \rightarrow \quad p(Y) = \mathcal{N}(A\mu + B, A\Sigma A^T)$$

- sumie ($Y = X_1 + X_2$)

$$p(X_1) = \mathcal{N}(\mu_1, \Sigma_1), p(X_2) = \mathcal{N}(\mu_2, \Sigma_2),$$
$$\rightarrow p(Y) = \mathcal{N}(\mu_1 + \mu_2, \Sigma_1 + \Sigma_2)$$

Notacja (2)

Zmiany momentów dystrybucji gaussowskich przy przekształceniach

- liniowym ($Y = AX + B$)

$$p(X) = \mathcal{N}(\mu, \Sigma) \quad \rightarrow \quad p(Y) = \mathcal{N}(A\mu + B, A\Sigma A^T)$$

- sumie ($Y = X_1 + X_2$)

$$p(X_1) = \mathcal{N}(\mu_1, \Sigma_1), p(X_2) = \mathcal{N}(\mu_2, \Sigma_2), \\ \rightarrow \quad p(Y) = \mathcal{N}(\mu_1 + \mu_2, \Sigma_1 + \Sigma_2)$$

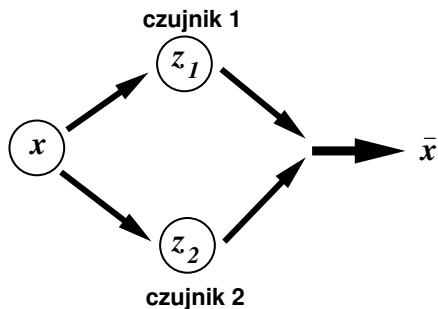
- iloczyn

$$p(X_1)p(X_2) = \mathcal{N}(\mu_1, \Sigma_1)\mathcal{N}(\mu_2, \Sigma_2) \quad \rightarrow \quad \mathcal{N}(\mu_Y, \Sigma_Y),$$

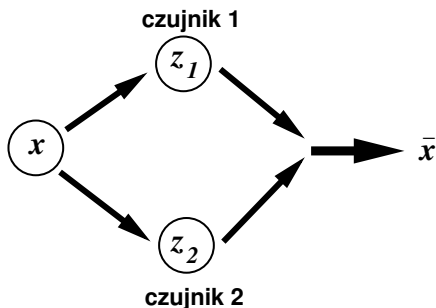
$$\Sigma_Y = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}$$

$$\mu_Y = \Sigma_Y \Sigma_1^{-1} \mu_1 + \Sigma_Y \Sigma_2^{-1} \mu_2$$

Niezależne pomiary



$$p(x|z_1 \wedge z_2) = p(x|z_1)p(x|z_2) = \eta p(z_1|x)p(z_2|x)$$



Niezależne pomiary – przykład

Założmy, że mamy dwa czujniki A i B, dające niezależne pomiary. Czujnik A ma dokładność ± 0.8 , czujnik B: ± 0.5 . Rozkłady prawdopodobieństwa są jednorodne w podanym zakresie dokładności.

Niezależne pomiary – przykład

Założmy, że mamy dwa czujniki A i B, dające niezależne pomiary.

Czujnik A ma dokładność ± 0.8 , czujnik B: ± 0.5 . Rozkłady prawdopodobieństwa są jednorodne w podanym zakresie dokładności.

Przy pomiarze czujniki zwróciły – A:2.0, a B:2.9.

- 1 Jak wyglądają rozkłady prawdopodobieństwa pomiarów z obu czujników?

Niezależne pomiary – przykład

Założmy, że mamy dwa czujniki A i B, dające niezależne pomiary.

Czujnik A ma dokładność ± 0.8 , czujnik B: ± 0.5 . Rozkłady prawdopodobieństwa są jednorodne w podanym zakresie dokładności.

Przy pomiarze czujniki zwróciły – A:2.0, a B:2.9.

- 1 Jak wyglądają rozkłady prawdopodobieństwa pomiarów z obu czujników?
- 2 Jak wygląda rozkład gęstości pomiaru łącznego, jaka jest jego dokładność?

Niezależne pomiary – przykład

Założmy, że mamy dwa czujniki A i B, dające niezależne pomiary.

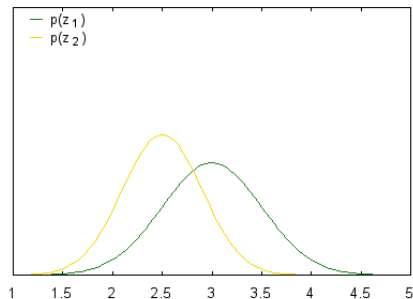
Czujnik A ma dokładność ± 0.8 , czujnik B: ± 0.5 . Rozkłady prawdopodobieństwa są jednorodne w podanym zakresie dokładności.

Przy pomiarze czujniki zwróciły – A:2.0, a B:2.9.

- 1 Jak wyglądają rozkłady prawdopodobieństwa pomiarów z obu czujników?
- 2 Jak wygląda rozkład gęstości pomiaru łącznego, jaka jest jego dokładność?
- 3 (*) Jak dokładność zależy od różnicy między odczytami z sensorów A i B?

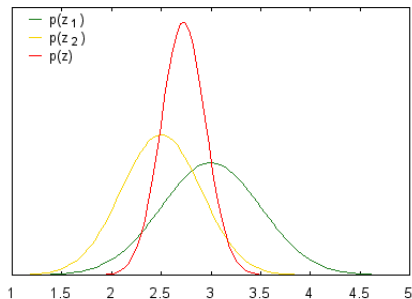
Niezależne pomiary – c.d.

$$p(x|z_1 \wedge z_2) = p(x|z_1)p(x|z_2) = \eta p(z_1|x)p(z_2|x)$$



Niezależne pomiary – c.d.

$$p(x|z_1 \wedge z_2) = p(x|z_1)p(x|z_2) = \eta p(z_1|x)p(z_2|x)$$



Proces ma własność Markowa, jeśli warunkowe rozkłady przyszłych stanów procesu zależą tylko od stanu bieżącego, a nie od stanów przeszłych

Proces ma własność Markowa, jeśli warunkowe rozkłady przyszłych stanów procesu zależą tylko od stanu bieżącego, a nie od stanów przeszłych

Własność ta w robotyce mobilnej często nie jest spełniona np. z powodu następujących czynników:

- model nie uwzględnia dynamiki otoczenia (np. poruszający się ludzie i ich wpływ na pomiary czujników),
- niedokładności modeli probabilistycznych $p(z_t|x_t)$, $p(x_t|x_{t-1}, u_t)$,
- błędy przybliżonych reprezentacji (np. rozkładu prawdopodobieństwa).

Niektóre z nich mogą być uwzględnione w modelu, jednak kosztem jego znacznej komplikacji.

Wszystkie wykorzystują ideę wielowymiarowej dystrybucji Gaussa, z parametryzacją za pomocą momentów, lub kanoniczną.

Zakładają zatem *unimodalność*, z czego wynika:

- + dobrze sprawdzają się w zadaniach ze zmienną losową w pobliżu stanu rzeczywistego
- działają gorzej lub wcale w zadaniach z wieloma równoprawnymi hipotezami

Filtr Kalmana (KF) (Swerling 1958, Kalman 1960)

- Stosowany dla układów liniowych, ciągłych.
- W chwili t szacujemy rozkład prawdopodobieństwa X przez $\mathcal{N}(\mu_t, \Sigma_t)$.
- Zakładamy, że spełnione są założenia Markowa oraz:

- 1) prawdopodobieństwo dla funkcji przejścia jest f.liniową

$$x_t = A_t x_{t-1} + B_t u_t + \epsilon_t, \quad \epsilon_t \rightarrow \mathcal{N}(0, R_t),$$

- 2) prawdopodobieństwo pomiarów – f. liniowa

$$z_t = C_t x_t + \delta_t, \quad \delta_t \rightarrow \mathcal{N}(0, Q_t),$$

- 3) prawdopodobieństwo początkowe

$$p(x_0) = \mathcal{N}(\mu_0, \Sigma_0).$$

Wtedy także prawdopodobieństwo *a posteriori* jest gaussowskie.

Filtr Kalmana – algorytm

$$\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$$

$$\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$$

$$K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t)$$

$$\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$$

Obliczenia przewidywanego rozkładu nowego stanu po zastosowaniu u_t .

Filtr Kalmana – algorytm

$$\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$$

$$\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$$

$$K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t)$$

$$\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$$

Wyznaczenie tzw. wzmocnienia Kalmana.

Filtr Kalmana – algorytm

$$\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$$

$$\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$$

$$K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t)$$

$$\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$$

Wyznaczenie wypadkowego rozkładu nowego stanu.

(jeśli wzmacnienie Kalmana wynosi 0 – nowy rozkład wyliczany jest wyłącznie na podstawie wiedzy *a priori*, jeśli zaś $K_t = C_t^{-1}$ – wyłącznie na podstawie pomiarów)

Filtr Kalmana – algorytm

$$\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$$

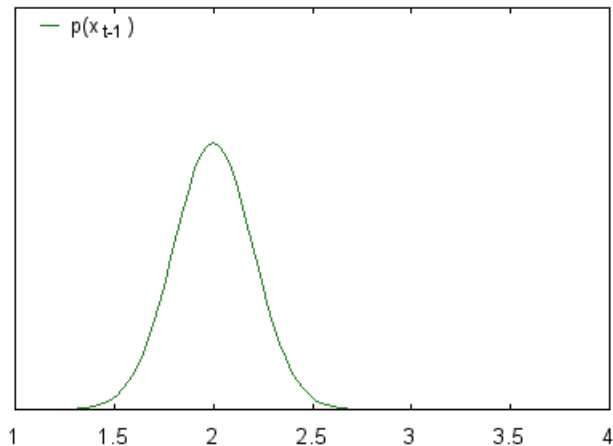
$$\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$$

$$K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t)$$

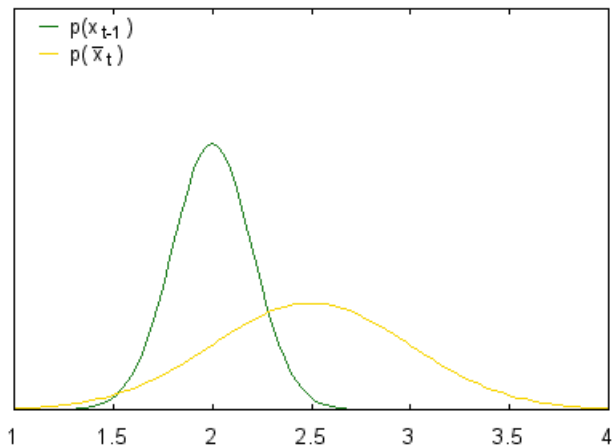
$$\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$$

KF – przykład



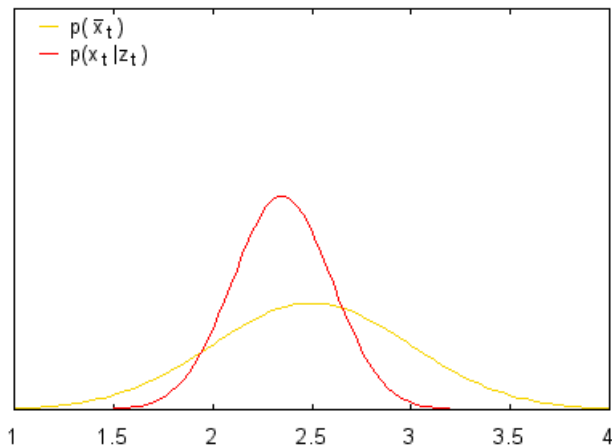
Rozpoczynamy od rozkładu w poprzedniej iteracji x_{t-1} .

KF – przykład



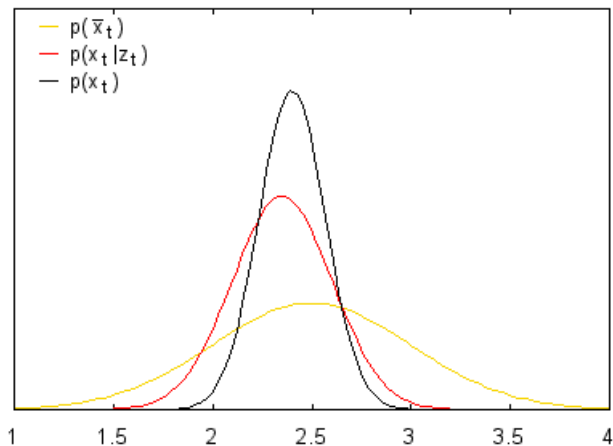
Wyznaczamy przewidywany rozkład \bar{x}_t po przemieszczeniu.

KF – przykład



Jednocześnie dokonujemy pomiaru z_t .

KF – przykład



Wyliczamy rozkład bieżący z_t .

Rozszerzony filtr Kalmana (EKF)

EKF

Stosowany dla układów nieliniowych postaci

$$x_t = g(u_t, x_{t-1}) + \epsilon_t$$

$$z_t = h(x_t) + \delta_t$$

EKF – algorytm

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

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

$$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - h_t(\bar{\mu}_t))$$

$$\Sigma_t = (I - K_t H_t) \bar{\Sigma}_t$$

Obliczane jacobiany:

$$G_t = g'(u_t, \mu_{t-1}), \quad g'(u_t, x_{t-1}) = \frac{\partial g(u_t, x_{t-1})}{\partial x_{t-1}}$$

EKF – algorytm

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

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

$$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - h_t(\bar{\mu}_t))$$

$$\Sigma_t = (I - K_t H_t) \bar{\Sigma}_t$$

Obliczane jacobiany:

$$H_t = h'(\mu_t) \quad h'(x_t) = \frac{\partial h(x_t)}{\partial x_t}$$

EKF – algorytm

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

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

$$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - h_t(\bar{\mu}_t))$$

$$\Sigma_t = (I - K_t H_t) \bar{\Sigma}_t$$

Obliczane jacobiany:

$$G_t = g'(u_t, \mu_{t-1}), \quad g'(u_t, x_{t-1}) = \frac{\partial g(u_t, x_{t-1})}{\partial x_{t-1}}$$

$$H_t = h'(\mu_t) \quad h'(x_t) = \frac{\partial h(x_t)}{\partial x_t}$$

EKF – algorytm

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

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

$$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - h_t(\bar{\mu}_t))$$

$$\Sigma_t = (I - K_t H_t) \bar{\Sigma}_t$$

Obliczane jacobiany:

$$G_t = g'(u_t, \mu_{t-1}), \quad g'(u_t, x_{t-1}) = \frac{\partial g(u_t, x_{t-1})}{\partial x_{t-1}}$$

$$H_t = h'(\mu_t) \quad h'(x_t) = \frac{\partial h(x_t)}{\partial x_t}$$

- Złożoność obliczeniowa jednej iteracji algorytmu $O(k^{2.4} + n^2)$.
- Założenie o unimodalności (jednak w wielu praktycznych przypadkach, mimo niespełnienia tego założenia, okazuje się skuteczny).
- Jakość estymacji zależy od nieliniowości funkcji w pobliżu punktu przybliżenia oraz od niepewności początkowej.

- MHEKF (*multi-hypothesis*),
- ADF (*assumed density*) – dopasowanie momentów *a posteriori*,
- UKF (*unscented*) – analiza stochastyczna z użyciem tzw. punktów σ ,
- IF/EIF (*(extended) information filter*) – dualny do f. Kalmana, wykorzystujący reprezentację kanoniczną rozkładu, czyli macierz informacji $\Omega = \Sigma^{-1}$ oraz wektor informacji $\xi = \Sigma^{-1}\mu$.

(E)IF

- + reprezentacja globalnej niepewności
- + lepsza stabilność numeryczna
- + w środowiskach wielorobotowych

(E)KF

- + mniejsza złożoność obliczeniowa

- MHEKF (*multi-hypothesis*),
- ADF (*assumed density*) – dopasowanie momentów *a posteriori*,
- UKF (*unscented*) – analiza stochastyczna z użyciem tzw. punktów σ ,
- IF/EIF (*(extended) information filter*) – dualny do f. Kalmana, wykorzystujący reprezentację kanoniczną rozkładu, czyli macierz informacji $\Omega = \Sigma^{-1}$ oraz wektor informacji $\xi = \Sigma^{-1}\mu$.

(E)IF

- + reprezentacja globalnej niepewności
- + lepsza stabilność numeryczna
- + w środowiskach wielorobotowych

(E)KF

- + mniejsza złożoność obliczeniowa

- MHEKF (*multi-hypothesis*),
- ADF (*assumed density*) – dopasowanie momentów *a posteriori*,
- UKF (*unscented*) – analiza stochastyczna z użyciem tzw. punktów σ ,
- IF/EIF (*(extended) information filter*) – dualny do f. Kalmana, wykorzystujący reprezentację kanoniczną rozkładu, czyli macierz informacji $\Omega = \Sigma^{-1}$ oraz wektor informacji $\xi = \Sigma^{-1}\mu$.

(E)IF

- + reprezentacja globalnej niepewności
- + lepsza stabilność numeryczna
- + w środowiskach wielorobotowych

(E)KF

- + mniejsza złożoność obliczeniowa

- MHEKF (*multi-hypothesis*),
- ADF (*assumed density*) – dopasowanie momentów *a posteriori*,
- UKF (*unscented*) – analiza stochastyczna z użyciem tzw. punktów σ ,
- IF/EIF (*(extended) information filter*) – dualny do f. Kalmana, wykorzystujący reprezentację kanoniczną rozkładu, czyli macierz informacji $\Omega = \Sigma^{-1}$ oraz wektor informacji $\xi = \Sigma^{-1}\mu$.

(E)IF

- + reprezentacja globalnej niepewności
- + lepsza stabilność numeryczna
- + w środowiskach wielorobotowych

(E)KF

- + mniejsza złożoność obliczeniowa

- MHEKF (*multi-hypothesis*),
- ADF (*assumed density*) – dopasowanie momentów *a posteriori*,
- UKF (*unscented*) – analiza stochastyczna z użyciem tzw. punktów σ ,
- IF/EIF (*(extended) information filter*) – dualny do f. Kalmana, wykorzystujący reprezentację kanoniczną rozkładu, czyli macierz informacji $\Omega = \Sigma^{-1}$ oraz wektor informacji $\xi = \Sigma^{-1}\mu$.

(E)IF

- + reprezentacja globalnej niepewności
- + lepsza stabilność numeryczna
- + w środowiskach wielorobotowych

(E)KF

- + mniejsza złożoność obliczeniowa