Lissack-Fu Distance Based Pruning for Gaussian Mixtures

2nd International Workshop on Data - Algorithms - Decision Making 2006

Ondřej Straka and Miroslav Šimandl

Research Center Data, Algorithms, and Decision
Department of Cybernetics
University of West Bohemia in Pilsen
Czech Republic

December 9.-12., 2006

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
0000000

Numerical illustration
000

Conclusion

Outline

1 Introduction

2 Gaussian Mixture Filter

3 Lissack Fu distance based pruning
- main idea
- algorithm
- thrifty implementation

4 Numerical Illustration

5 Conclusion

Motivation

- Gaussian Mixture (GM) filter is a very popular tool in nonlinear state estimation.
- Represents a bank of Kalman filters (for linear nongaussian system) or a bank of extended Kalman filters (for nonlinear nongaussian system)
- Considering nongaussian noises, the number of Kalman filters usually grows *exponentialy* $\Rightarrow$ application of a reduction technique is *inevitable*.
- The reduction technique can significantly affect estimate quality.

Present approaches to reduction in GM

Pruning

cutting off some terms of the GM e.g. maximum posterior probability pruning, threshold-based pruning

Merging

joining two or more GM terms into one e.g. generalized pseudo Bayes approach

Optimization

finding parameters of a new GM using a parameter optimization technique, e.g. statistical parameter optimization (EM)

Combined

combining the previous three approaches parameter optimization and structural adaptation presented in papers by prof. Hanenbeck et al.

Present approaches to reduction in GM

Pruning

cutting off some terms of the GM e.g. maximum posterior probability
pruning, threshold-based pruning

Merging

joining two or more GM terms into one e.g. generalized pseudo
Bayes approach

Optimization

finding parameters of a new GM using a parameter optimization
technique, e.g. statistical parameter optimization (EM)

Combined

combining the previous three approaches parameter optimization and
structural adaptation presented in papers by prof. Hanenbeck et al.

Present approaches to reduction in GM

Pruning

cutting off some terms of the GM e.g. maximum posterior probability
pruning, threshold-based pruning

Merging

joining two or more GM terms into one e.g. generalized pseudo
Bayes approach

Optimization

finding parameters of a new GM using a parameter optimization
technique, e.g. statistical parameter optimization (EM)

Combined

combining the previous three approaches parameter optimization and
structural adaptation presented in papers by prof. Hanenbeck et al.

Ondřej Straka and Miroslav Šimandl   **University of West Bohemia in Pilsen Czech Republic**

Lissack-Fu Distance Based Pruning for Gaussian Mixtures

Present approaches to reduction in GM

Pruning

cutting off some terms of the GM e.g. maximum posterior probability pruning, threshold-based pruning

Merging

joining two or more GM terms into one e.g. generalized pseudo Bayes approach

Optimization

finding parameters of a new GM using a parameter optimization technique, e.g. statistical parameter optimization (EM)

Combined

combining the previous three approaches parameter optimization and structural adaptation presented in papers by prof. Hanenbeck et al.

Goal of the paper

- To propose a *pruning technique* that analyzes importance of each GM term in a clever way utilizing more information than just the weights of the GM terms (e.g. using a distance of two pdf's).
- To design a *thrifty implementation* of the proposed technique to keep computational requirements low as in other simple pruning techniques.

Goal of the paper

- To propose a *pruning technique* that analyzes importance of each GM term in a clever way utilizing more information than just the weights of the GM terms (e.g. using a distance of two pdf's).
- To design a *thrifty implementation* of the proposed technique to keep computational requirements low as in other simple pruning techniques.

State estimation

Consider a linear nongaussian system:

$$\mathbf{x}_{k+1} = \mathbf{F}_k \cdot \mathbf{x}_k + \mathbf{w}_k, \quad k = 0, 1, 2, \ldots$$
$$\mathbf{z}_k = \mathbf{H}_k \cdot \mathbf{x}_k + \mathbf{v}_k, \quad k = 0, 1, 2, \ldots$$

- $\mathbf{x}_k$ - state, $\mathbf{z}_k$ - measurement
- $\mathbf{w}_k$ and $\mathbf{v}_k$ are mutually independent white noises with
  $p(\mathbf{w}_k) = \sum_{n=1}^{q_k} \beta_{kn} \mathcal{N} \{\mathbf{w}_k : \hat{\mathbf{w}}_{kn}, \mathbf{Q}_{kn}\}$
  $p(\mathbf{v}_k) = \sum_{m=1}^{r_k} \gamma_{km} \mathcal{N} \{\mathbf{v}_k : \hat{\mathbf{v}}_{km}, \mathbf{R}_{km}\}$
  and independent of $\mathbf{x}_0$.

The aim of state estimation here is to find

the filtering pdf $p(\mathbf{x}_k | \mathbf{z}^k)$, $\mathbf{z}^k = [\mathbf{z}_0^{\mathrm{T}}, \ldots \mathbf{z}_k^{\mathrm{T}}]^{\mathrm{T}}$

Gaussian Mixture (GM) filter

- GM filter solves the Bayesian Recursive Relations (BRR) using a bank of Kalman filters.
- At each time instant k the filtering pdf is given as $p(\mathbf{x}_k|\mathbf{z}^k) = \sum_{j=1}^{\xi_k} \alpha_{kj} \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{kj}, \mathbf{P}_{kj})$.
- The means $\hat{\mathbf{x}}_{kj}$ and the covariance matrices $\mathbf{P}_{kj}$ are calculated using individual Kalman filters.
- The number of terms in the mixture grows exponentially as

$$\xi_k = q_k r_k \xi_{k-1}$$

.

- Thus, a reduction technique has to be implemented to keep computational demands of the algorithm within reasonable bounds.

Distance-based pruning - an idea

1 to start with the approximate filtering pdf containing one term of the GM of the filtering pdf

2 to *append* other terms of the filtering pdf to the approximate filtering pdf successively

3 to *assess* significancy of the appended terms with respect to approximation quality

4 if the term is insignificant, it is discarded

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
●0000000

Numerical illustration
000

Conclusion

Distance-based pruning - an idea

1. to start with the approximate filtering pdf containing one term of the GM of the filtering pdf

2. to *append* other terms of the filtering pdf to the approximate filtering pdf successively

3. to *assess* significancy of the appended terms with respect to approximation quality

4. if the term is insignificant, it is discarded

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
●0000000

Numerical illustration
000

Conclusion

Distance-based pruning - an idea

1. to start with the approximate filtering pdf containing one term of the GM of the filtering pdf

2. to *append* other terms of the filtering pdf to the approximate filtering pdf successively

3. to *assess* significancy of the appended terms with respect to approximation quality

4. if the term is insignificant, it is discarded

Distance-based pruning - an idea

1. to start with the approximate filtering pdf containing one term of the GM of the filtering pdf

2. to *append* other terms of the filtering pdf to the approximate filtering pdf successively

3. to *assess* significancy of the appended terms with respect to approximation quality

4. if the term is insignificant, it is discarded

Introduction	Gaussian Mixture filter	**Distance-based pruning**	Numerical illustration	Conclusion
ooo	oo	o●oooooo	ooo	

Distance-based pruning - an idea

- Significancy of an appended term is measured using the Lissack Fu (LF) distance between the approximate pdf with the considered term and the original unpruned pdf

The LF distance between $p_1(\mathbf{x}_k|\mathbf{z}^k)$ and $p_2(\mathbf{x}_k|\mathbf{z}^k)$ is given as

$$L_\gamma(p_1, p_2) \triangleq \int |p_1(\mathbf{x}_k|\mathbf{z}^k) - p_2(\mathbf{x}_k|\mathbf{z}^k)|^\gamma \mathrm{d}\mathbf{x}_k,$$

where γ is chosen as $\gamma = 2$ in the paper.

-

Considering $p_1(\mathbf{x}_k|\mathbf{z}^k) = \sum_{i=1}^{\xi 1} \alpha_{ki}^1 \mathcal{N}\{\mathbf{x}_k; \hat{\mathbf{x}}_{ki}^1, \mathbf{P}_{ki}^1\}$ and
$p_2(\mathbf{x}_k|\mathbf{z}^k) = \sum_{j=1}^{\xi 2} \alpha_{kj}^2 \mathcal{N}\{\mathbf{x}_k; \hat{\mathbf{x}}_{kj}^2, \mathbf{P}_{kj}^2\}$,
$L_2(p_1, p_2) = A_{11} - 2A_{12} + A_{22},$

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
00●00000

Numerical illustration
000

Conclusion

Distance-based pruning - an idea

where

$$A_{11} = \sum_{i=1}^{\xi^1} \sum_{j=1}^{\xi^1} \alpha_{ki}^1 \alpha_{kj}^1 \mathcal{N}\{\hat{\mathbf{x}}_{ki}^1 : \hat{\mathbf{x}}_{kj}^1, \mathbf{P}_{ki}^1 + \mathbf{P}_{kj}^1\},$$

$$A_{12} = \sum_{i=1}^{\xi^1} \sum_{j=1}^{\xi^2} \alpha_{ki}^1 \alpha_{kj}^2 \mathcal{N}\{\hat{\mathbf{x}}_{ki}^1 : \hat{\mathbf{x}}_{kj}^2, \mathbf{P}_{ki}^1 + \mathbf{P}_{kj}^2\},$$

$$A_{22} = \sum_{i=1}^{\xi^2} \sum_{j=1}^{\xi^2} \alpha_{ki}^2 \alpha_{kj}^2 \mathcal{N}\{\hat{\mathbf{x}}_{ki}^2 : \hat{\mathbf{x}}_{kj}^2, \mathbf{P}_{ki}^2 + \mathbf{P}_{kj}^2\}.$$

Distance-based pruning - algorithm

0. Initialization Sort the terms of the GM. The initial $\hat{p}(\mathbf{x}_k|\mathbf{z}^k)$ is given by the highest-weighted term. Specify r and $L_{2,k}^{\max} = L_2\left(p(\mathbf{x}_k|\mathbf{z}^k), \hat{p}(\mathbf{x}_k|\mathbf{z}^k)\right)$.

1. Term appending Append the next term to the approximate filtering pdf $\hat{p}(\mathbf{x}_k|\mathbf{z}^k)$.

2. Distance computing Compute $L_2\left(p(\mathbf{x}_k|\mathbf{z}^k), \hat{p}(\mathbf{x}_k|\mathbf{z}^k)\right)$.

3. Assessment If $L_2\left(p(\mathbf{x}_k|\mathbf{z}^k), \hat{p}(\mathbf{x}_k|\mathbf{z}^k)\right) \leq L_{2,k}^{\max} \cdot \exp(-r\xi_k^{\mathrm{red}})$, the term is accepted, otherwise the term is rejected and discarded.

Note: For $r = 0$ no term is pruned, with increasing r more terms are pruned.

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
0000●000

Numerical illustration
000

Conclusion

Distance based pruning - thrifty implementation

Idea

computing as many terms as possible in advance and multiple utilization of the values for evaluation of the distance

- Construct a $\xi_k \times \xi_k$ matrix $\mathbf{L}_k$ with the elements $\mathbf{L}_k(i,j) = \alpha_{ki}\alpha_{kj}\mathcal{N}\{\hat{\mathbf{x}}_{ki} : \hat{\mathbf{x}}_{kj}, \mathbf{P}_{ki} + \mathbf{P}kj\}$. The matrix is symmetric.
- A_{11} is given by the normalized sum of all elements $\mathbf{L}_k(l_i, l_j)$ $i = 1, 2, \ldots, \xi_k^{\mathrm{red}} j = 1, 2, \ldots, \xi_k^{\mathrm{red}}$, where l_i is the index of the ith term of $\hat{p}(\mathbf{x}_k|\mathbf{z}^k)$ in $p(\mathbf{x}_k|\mathbf{z}^k)$.
- A_{12} is given by the normalized sum of all elements $\mathbf{L}_k(l_i, j)$ $i = 1, 2, \ldots, \xi_k^{\mathrm{red}}, j = 1, 2, \ldots, \xi_k$.
- A_{22} is computed as the sum of all elements of $\mathbf{L}_k$.

Distance based pruning - thrifty implementation

Example for computing A_{11}, A_{12}, A_{22}

Original unpruned pdf $p(\mathbf{x}|\mathbf{z}^k)$ contains $\xi_k = 4$ sorted terms.
The pruned pdf $\hat{p}(\mathbf{x}|\mathbf{z}^k)$ considers the terms number 1 and 3.

$$A_{11}: \quad \mathbf{L}_k = \begin{matrix} \mathbf{L}_k(1,1) & \mathbf{L}_k(1,2) & \mathbf{L}_k(1,3) & \mathbf{L}_k(1,4) \\ \mathbf{L}_k(2,1) & \mathbf{L}_k(2,2) & \mathbf{L}_k(2,3) & \mathbf{L}_k(2,4) \\ \mathbf{L}_k(3,1) & \mathbf{L}_k(3,2) & \mathbf{L}_k(3,3) & \mathbf{L}_k(3,4) \\ \mathbf{L}_k(4,1) & \mathbf{L}_k(4,2) & \mathbf{L}_k(4,3) & \mathbf{L}_k(4,4) \end{matrix}$$

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
00000000

Numerical illustration
000

Conclusion

Distance based pruning - thrifty implementation

Example for computing A_{11}, A_{12}, A_{22}

Original unpruned pdf $p(\mathbf{x}_|\mathbf{z}^k)$ contains $\xi_k = 4$ sorted terms.
The pruned pdf $\hat{p}(\mathbf{x}_|\mathbf{z}^k)$ considers the terms number 1 and 3.

$$A_{12}: \quad \mathbf{L}_k = \begin{matrix} \mathbf{L}_k(1,1) & \mathbf{L}_k(1,2) & \mathbf{L}_k(1,3) & \mathbf{L}_k(1,4) \\ \mathbf{L}_k(2,1) & \mathbf{L}_k(2,2) & \mathbf{L}_k(2,3) & \mathbf{L}_k(2,4) \\ \mathbf{L}_k(3,1) & \mathbf{L}_k(3,2) & \mathbf{L}_k(3,3) & \mathbf{L}_k(3,4) \\ \mathbf{L}_k(4,1) & \mathbf{L}_k(4,2) & \mathbf{L}_k(4,3) & \mathbf{L}_k(4,4) \end{matrix}$$

Introduction
ooo

Gaussian Mixture filter
oo

Distance-based pruning
oooooooo●

Numerical illustration
ooo

Conclusion

Distance based pruning - thrifty implementation

Example for computing A_{11}, A_{12}, A_{22}

Original unpruned pdf $p(\mathbf{x}|\mathbf{z}^k)$ contains $\xi_k = 4$ sorted terms.
The pruned pdf $\hat{p}(\mathbf{x}|\mathbf{z}^k)$ considers the terms number 1 and 3.

$$A_{12}: \quad \mathbf{L}_k = \begin{matrix}
\mathbf{L}_k(1,1) & \mathbf{L}_k(1,2) & \mathbf{L}_k(1,3) & \mathbf{L}_k(1,4) \\
\mathbf{L}_k(2,1) & \mathbf{L}_k(2,2) & \mathbf{L}_k(2,3) & \mathbf{L}_k(2,4) \\
\mathbf{L}_k(3,1) & \mathbf{L}_k(3,2) & \mathbf{L}_k(3,3) & \mathbf{L}_k(3,4) \\
\mathbf{L}_k(4,1) & \mathbf{L}_k(4,2) & \mathbf{L}_k(4,3) & \mathbf{L}_k(4,4)
\end{matrix}$$

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
00000000

Numerical illustration
●00

Conclusion

Consider a simple system

$$x_{k+1} = 0.9 \cdot x_k + w_{k+1}, \, z_k = x_k + v_k$$
$$p(w_k) = 0.2\mathcal{N}\{w_k : -1, 0.1\} + 0.6\mathcal{N}\{w_k : 0, 0.1\} + 0.2\mathcal{N}\{w_k : 1, 0.1\},$$
$$p(v_k) = 0.8\mathcal{N}\{v_k : 0, 0.01\} + 0.2\mathcal{N}\{v_k : 0, 10\}, \, p(x_0) = \mathcal{N}\{x_0 : 0, 0.1\}$$

Criteria

$$J(x_k) = \mathsf{E}_{p(z^k)} \left(1 - \int \min\left[\hat{p}(x_k|z^k), p(x_k|z^k)\right] \mathrm{d}x_k \right)$$
$$\text{MSE: } \bar{\Pi} = \frac{1}{100}\sum_{k=1}^{100}(x_k - \hat{x}_k)^2$$

	DBP-0.4	MPPP-10	MPPP-200	EM-4	GPB1
$\bar{J}$	0.02	0.12	0.09	0.03	0.41
$\bar{\Pi}$	0.26	0.38	0.30	0.25	0.26
$\bar{T}$ [s]	0.26	0.03	0.57	11.49	0.01
$\bar{\xi}$	13.2	10	200	4	1

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
00000000

Numerical illustration
0●0

Conclusion

$J(\mathbf{x}_k)$ - various methods

$J(\mathbf{x}_k)$ DBP with variable r
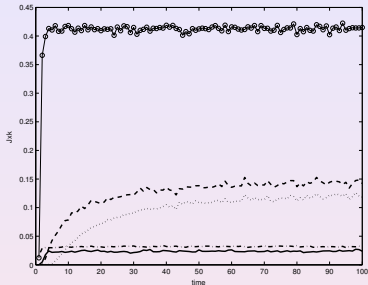
DBP-0.4 - solid
MPPP-10 - dashed
MPPP-10 - dotted
EM-4 - dash-dotted
GPB-1 - circles

DBP-0.6
DBP-0.5
DBP-0.4
DBP-0.3

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
00000000

Numerical illustration
00●

Conclusion
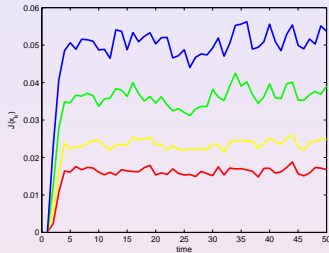
Example of filtering pdf approximation
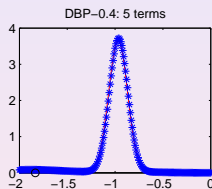
$$p(x_{35}|z^{35}), \hat{p}(x_{35}|z^{35})$$

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
00000000

Numerical illustration
00●

Conclusion
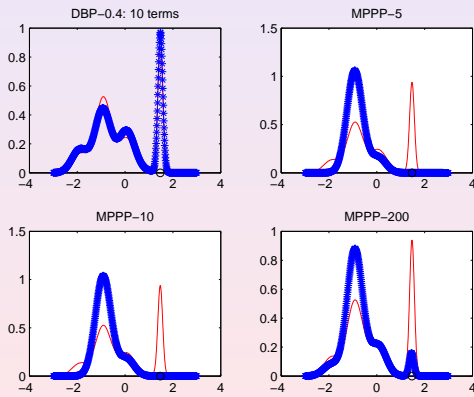
Example of filtering pdf approximation

$$p(x_{85}|z^{85}), \hat{p}(x_{85}|z^{85})$$

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
0000000

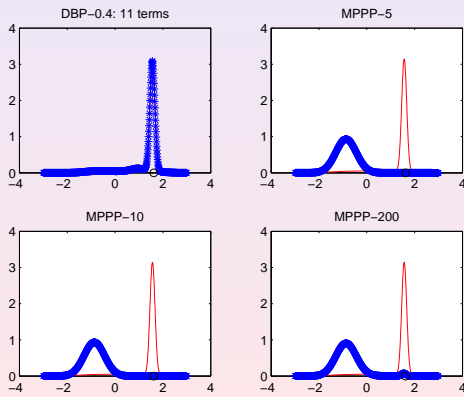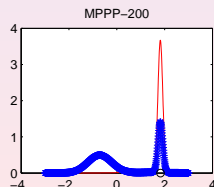**Numerical illustration**
00●
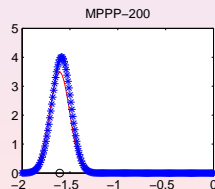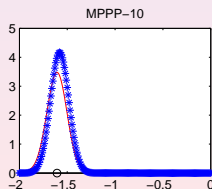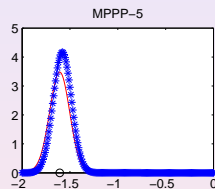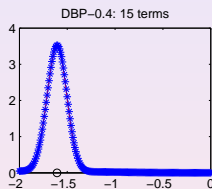
Conclusion

Example of filtering pdf approximation

$$p(x_{89}|z^{89}), \hat{p}(x_{89}|z^{89})$$

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
0000000

Numerical illustration
00●
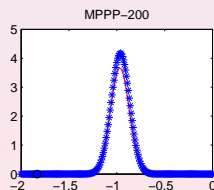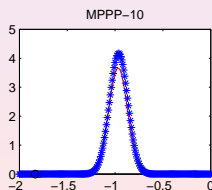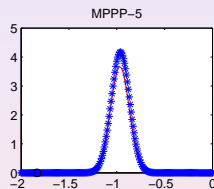
Conclusion

Example of filtering pdf approximation

$$p(x_{96}|z^{96}), \hat{p}(x_{96}|z^{96})$$

Introduction
000

Gaussian Mixture filter
00

Distance-based pruning
0000000

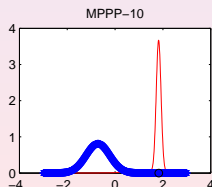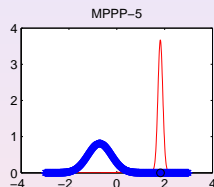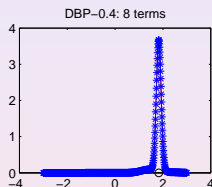**Numerical illustration**
00●
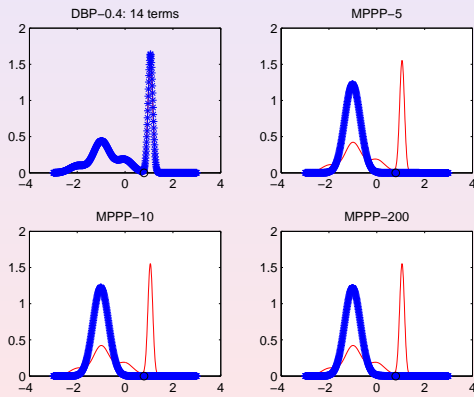
Conclusion

Example of filtering pdf approximation

$$p(x_{97}|z^{97}), \hat{p}(x_{97}|z^{97})$$

Example of filtering pdf approximation

$$p(x_{98}|z^{98}), \hat{p}(x_{98}|z^{98})$$

Introduction
○○○

Gaussian Mixture filter
○○

Distance-based pruning
○○○○○○○

Numerical illustration
○○○

Conclusion

Conclusion

- A new distance-based pruning technique has been proposed.
- It prunes a term according to its significance given by the LF distance between the unpruned pdf and the pruned pdf.
- The technique provides high quality approximations with respect to the other techniques (MPPP, GPB, EM).
- The technique is very efficient due to the proposed thrifty implementation which accomplishes as many computations as possible in advance. The results of the computations are used several times for calculation of the distance.