

# NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

---

Finite State Models for Non-linear Estimation

Didier J. Le Gall

10 Rue Paul Vaillant Couturier, Levallois - Perret 92300 France

---

## RESUME

On étudie une nouvelle approche du problème d'estimation non linéaire. La solution de l'équation d'état est remplacée par un modèle Markovien à complexité finie. On résout le problème de l'estimation maximum à posteriori M.A.P. en utilisant un algorithme séquentiel comme l'algorithme de Viterbi.

On présente une application au problème classique de poursuite de phase et on compare les résultats à ceux d'un filtre de Kalman étendu; quand le niveau de bruit est fort l'amélioration est spectaculaire. Pour cette application le coût des calculs est faible et l'algorithme de Viterbi permet également des lissages avec délais. Cette méthode présente un grand intérêt dans le cas des canaux de transmission haute fréquence.

## SUMMARY

A new approach to non-linear estimation is studied where the solution of the dynamic equation is modeled as a finite state Markov Chain. The Maximum a posteriori M.A.P estimation problem can be solved using a sequential search procedure such as the Viterbi algorithm.

An application to the classical phase tracking problem is given and the results are compared to the Extended Kalman filter results; in that case the improvement is striking when the noise level is high. The Viterbi algorithm allows filtering as well as sequence estimation and fixed-lag smoothing with a small computational complexity. Also an application to demodulation with Radio Frequency channel is presented.



Finite State Models  
for Non-linear Estimation

I Introduction

Given a dynamic system with discrete time representation such as

$$\begin{aligned} X_{k+1} &= f(X_k) + g(X_k, W_k) & /1/ \\ Y_k &= h(X_k) + V_k & /2/ \end{aligned}$$

where the state process  $X_k$  is driven by the noise process  $W_k$  and the observation process  $Y_k$  is perturbed by the noise process  $V_k$ .

( $W_k$  and  $V_k$  are independant i.i.d. processes)

The task of non-linear filtering is to estimate the state  $X_k$  given the observation up to time,  $k$   $y_0^k = (y_1, y_2, \dots, y_k)$ . However estimates of the state  $X_k$  with nice properties such as the minimum variance estimate  $\hat{X}_{M.S} = E \left\{ X_k / Y_k \right\}$  or the M.A.P estimate Maximum a posteriori, require the computation of the conditional distribution of the state given the observation. This computation via the conditional density has led to very complex algorithms not always easily instrumentable even with the most powerful digital computer.

Most of the efforts to find instrumentable solution to the non-linear filtering problem have been concerned with the generalization of the Kalman Filtering technique and are thus known as "Extended Kalman Filters". They provide good results in many cases, however in high noise environment they do not perform satisfactorily. Phenomenon often called Kalman Filter divergence.

Finite State Models

We propose here a new approach to non-linear estimation that seems better suited

to the high noise case, without having the computational cost that the so-called "true non-linear filters" have. We realise a finite partition quantization of the state space  $\mathcal{S}$  with  $M$  elements and we construct a finite State Markov Chain on  $S_M$  that matches the original dynamic process described by /1/; this is possible provided the equation /1/ admits a stationary solution. Hence the approximation does not lie in the estimation but in the model itself. A Finite State Model for the equation /1/ is given by the state space  $S_M$  and the transition probabilities

$$P_{ij} = \Pr \left( S_i \rightarrow S_j \right) \quad i, j = 1, \dots, M.$$

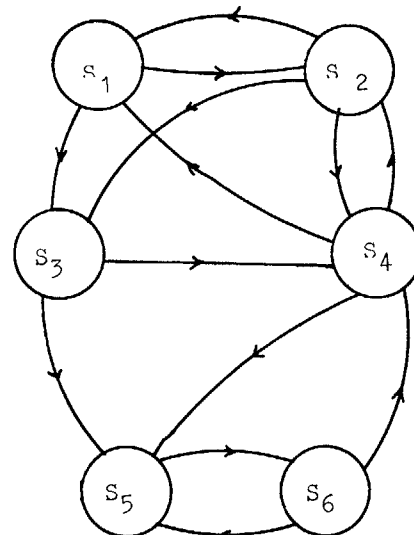


Fig.1. Finite State Model

Sequential estimation

The advantage of the finite state description is that we have only a finite number of state trajectories. The modified non-linear filtering problem can be posed in

Finite State Models for Non-linear Estimation

the following terms: "Among all possible trajectories, find the one that maximizes a given performance index and thus best accounts for the observation". If we denote by  $\mathcal{Y}_0^N$  the set of all state trajectories of length N the modified problem is stated as follows:

$$\begin{aligned}
 & X_0^N \in \mathcal{Y}_0^N && /3/ \\
 & Y_k = g(X_k) + V_k \quad k = 0, 1, 2, \dots, N && /4/
 \end{aligned}$$

the maximum a posteriori estimate  $\hat{X}_{MAP}^N$  maximizes the conditional density

$$\hat{X}_{MAP}^N = \max_{X_0^N \in \mathcal{Y}_0^N}^{-1} P(X_0^N / Y_0^N) \quad /5/$$

we can write also

$$\begin{aligned}
 \hat{X}_{MAP}^N = \max_{X_0^N \in \mathcal{Y}_0^N}^{-1} & \sum_{k=1}^N \left\{ \log P(Y_k / X_k) + \right. \\
 & \left. + \log P(X_k / X_{k-1}) \right\}
 \end{aligned}$$

and if we define

$$\begin{aligned}
 m(X_k, X_{k-1}; Y_k) &= \log P(Y_k / X_k) + \\
 &+ \log P(X_k / X_{k-1}), \\
 \hat{X}_{MAP}^N &= \max_{X_0^N \in \mathcal{Y}_0^N}^{-1} \sum_1^N m(X_k, X_{k-1}; Y_k) \quad /6/
 \end{aligned}$$

and this modified non-linear filtering problem amounts to minimizing a metric in a tree, hence it can be solved using a sequential algorithm.

The Viterbi algorithm

Among all sequential search procedures we shall rely particularly on the Viterbi algorithm. The idea of using the Viterbi algorithm together with a finite description of the dynamic system, has been first suggested by J.K.Omura [1].

The Viterbi algorithm was introduced in relation to the decoding of convolutional codes. The trees generated by such codes

have a strong regularity since the encoder has a finite memory and this allows us to describe them by a trellis diagram.

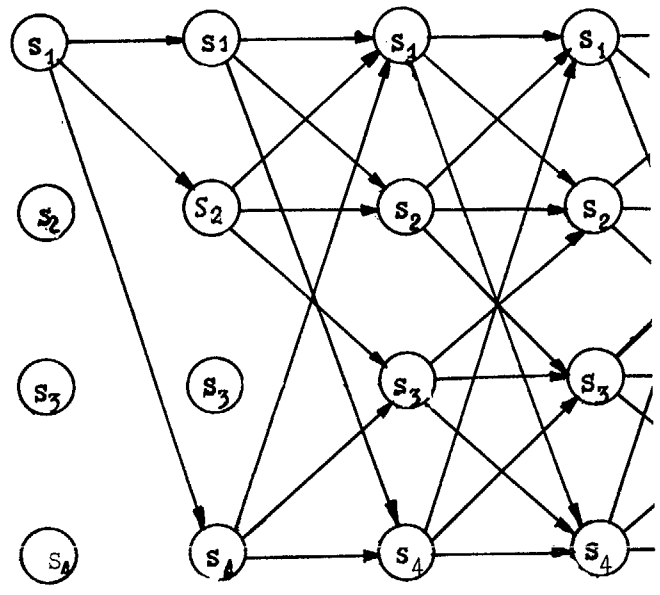


Fig.2. Trellis diagram and Viterbi algorithm

The original state sequence is represented by a path on the trellis; the estimated state sequence is given by the path on the trellis that most closely matched the observations in term of the metric. In comparison with other tree search algorithms the Viterbi algorithm does an exhaustive search with a minimum complexity. Omura in [2] has shown that the Viterbi algorithm is the same as the forward dynamic programming algorithm. Specialized processors that perform the Viterbi algorithm efficiently are quite common in the industry.

The success of the Finite State Model approach depends much on the existence of low complexity Finite State Models (F.S.M.) A general algorithm has been established to construct a F.S.M. for a stationary Markov process [3]. The rate-distorsion theory



and its related concept of  $\bar{d}$  - distance between processes provide us with some insight in model design as well as a relation between the fidelity and the complexity of the F.S.M. [3], pp 30,45. In what follows we give a practical example of implementation of a filtering algorithm via F.S.M. The simple example of phase tracking is a classic of non-linear filtering and a somewhat similar filter making use of the Viterbi algorithm has also been suggested in [4] with a different point of view.

II A Phase Tracking Problem

Phase tracking is the classic problem of coherent communication. We take the phase process to be described by a Random walk on the circle.

$$X_{k+1} = X_k \oplus W_k \quad /7/$$

(where  $\oplus$  denote the sum modulo  $2\pi$  and  $W_k$  is a i.i.d sequence). In most of what follows we assume  $W_k$  is gaussian with variance  $\sigma_w^2$ . however there is no such restriction on  $W_k$ . The observed data are given by

$$Y_{1k} = A \cos X_k + V_{1k} \quad /8/$$

$$Y_{2k} = A \sin X_k + V_{2k}$$

where  $V_{1k}$  and  $V_{2k}$  are independant white Gaussian processes with unit power spectrum.

As suggested we shall use a F.S.M. for the phase. The phase space  $(0, 2\pi)$  will be quantized into  $S_M = (s_1, s_2, \dots, s_M)$

where  $s_k$  is the interval

$$\left\{ \frac{2\pi}{M}k - \frac{\pi}{M}, \frac{2\pi}{M}k + \frac{\pi}{M} \right\}$$

centered on  $\hat{s}_k = \frac{2\pi}{M}k$ , the transition probability matrix is then

$$P_{ij} = \text{Prob} \left\{ X_k \in s_j \rightarrow X_{k+1} \in s_j \right\}$$

Reduced Complexity

It is interesting to note that in most cases the phase variance  $\sigma_w^2$  will be sufficiently small so that only transitions to the nearest neighbor have a significant probability of occurring hence we can reduce further the complexity of the model. The transition probability matrix  $p_{ij}$  is then given by

$$P_{ij} = \begin{cases} p_0 & i=j \\ \frac{1-p_0}{2} & i-j = 1 \\ 0 & i-j > 1 \end{cases}$$

where  $p_0 = \text{Prob} \left\{ W_k \in \left( -\frac{\pi}{M}, \frac{\pi}{M} \right) \right\}$

The reduced complexity Phase Model can be represented by the state diagram of Fig.3.

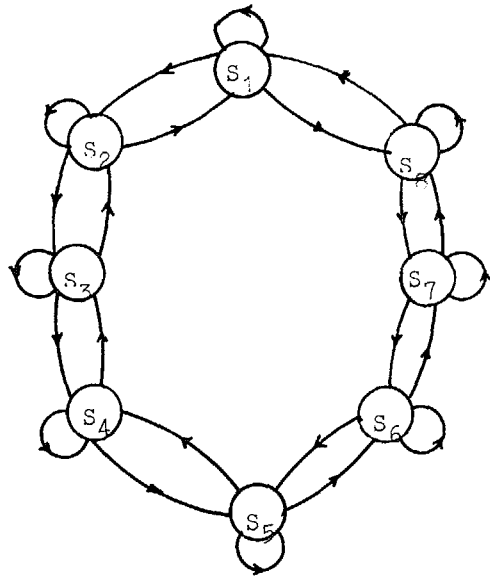


Fig.3. A Reduced Complexity Phase Model

The metric for the phase tracking problem can be computed easily. In the Gaussian case plugging the conditional density

$$P(Y_{1k}, Y_{2k}/X_k) = \frac{1}{2\pi} \exp - \frac{1}{2} \left\{ (Y_{1k} - A \cos X_k)^2 + (Y_{2k} - A \sin X_k)^2 \right\}$$

in the formula for the metric

$$m(X_k, X_{k-1}; Y_{1k}, Y_{2k}) = \text{Ln } P(Y_{1k}, Y_{2k}/X_k) + \text{Ln } P(X_k/X_{k-1})$$

yields for an admissible transition

$$m(X_k, X_{k-1}; Y_{1k}, Y_{2k}) = A \left( Y_{1k} \cos X_k + Y_{2k} \sin X_k \right) + \ln P(X_k/X_{k-1})$$

where  $P(X_k/X_{k-1})$  takes the two values  $P_0$  if  $X_k = X_{k-1}$  and  $(1-p_0)/2$  if  $X_k \neq X_{k-1}$ .

The form of this metric is very well suited to the realization of a fast tracking algorithm.

Comparison of Performance

The Viterbi algorithm can be used on data sequences for a string of about 100 data or as a causal filter or as a fixed lag K estimator. When one compares this phase tracking system to a causal system such as the Extended Kalman Filter or a Phase Lock Loop phase tracker the comparison with our estimator is unfair, since we make use of more data and thus allow for a slight delay in the phase estimation. To make the comparison more justified, it is more proper to compare our results to those of a Fixed-lag extended Kalman smoother. One notes however than to go from the Extended Kalman Filter to the Fixed-lag K extended Kalman smoother one has to increase the complexity K-fold, wherever in the Viterbi phase tracker only slight modifications in the storage requirements are needed to go from filtering to smoothing.

In the simulations of Fig.4 we compare our results to both Kalman filtering and Kalman smoothing. We observe that the Viterbi algorithm is performing well in low noise environment where EKF is nearly optimal and strikingly better when the noise level increases.

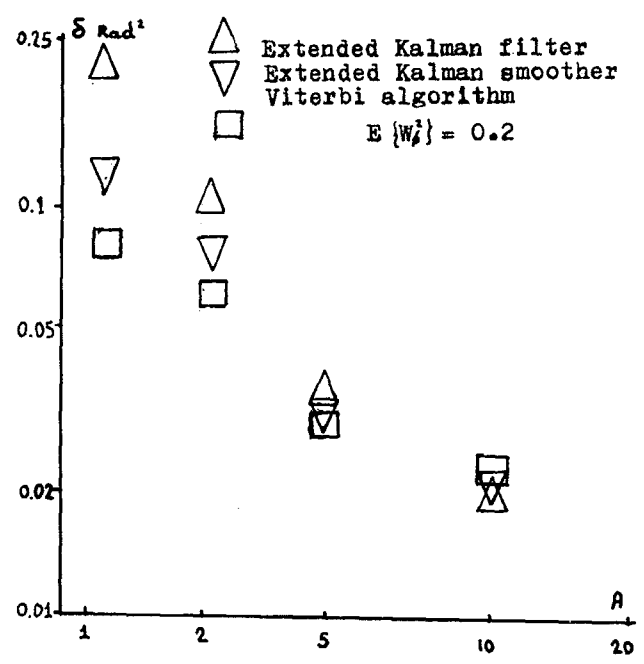


Fig.4. Comparison of Performance for Different Phase Trackers

The Viterbi algorithm is very robust wherever Extended Kalman smoothing presents some computational dangers (oversmoothing) that allow it to be used only with caution.

Finally one would like to show how Finite State Models estimation can be successful in simultaneously tracking the phase and recovering the data in a communication system.

In a B.P.S.K. modulation scheme transmitting one of two phases 0 or  $\pi$  over a Radio frequency communication channel, let  $u_k$  be the message taking values in  $\{0, \pi\}$  the received data, after taking the in-phase and quadrature components, are

$$Y_{1k} = A_k \cos(X_k + u_k) + V_{1k}$$

$$Y_{2k} = A_k \sin(X_k + u_k) + V_{2k}$$

where  $X_k$  is modeled as above,  $A_k$  is slowly varying and can be assumed constant for a string of data. If one extends the state



space  $S_M = (s_1, \dots, s_M)$  by taking into consideration pairs  $(X_k, u_k)$  where

$$(X_k, u_k) \in S_M \times \{0, \pi\}$$

the Viterbi algorithm still applies with only slight modifications; and provides us with simultaneous phase tracking and data demodulation.

### III Conclusion

This alternate approach seems to be very promising especially in high noise environment or whenever the EKF methods fail. Finite state models represent a serious reduction of complexity over other "true non-linear methods" such as those involving the conditional density with a discretization of the state space. The original complexity can be reduced by not allowing all state transitions to be represented and this only at a slight loss in performance.

### References

- [1] J.K.Omura and K.Kozlowski "Non-linear Smoothing Using Finite State Models" UCLA, School of Engineering and Applied Science 1980
- [2] J.K.Omura "On the Viterbi Decoding Algorithm", IEEE Transaction on Information Theory Correspondence, January 1969, pp177 - 179
- [3] D.J.Le Gall, Ph.D. Dissertation, University of California, Los Angeles 1981.
- [4] L.Scharf D.Cox C.Masreliez, "Modulo  $2\pi$  Phase Sequence Estimation" IEEE Transaction on Information Theory IT-26, September 1980.