ALGORITHM STRUCTURES FOR CYCLIC PHASE ESTIMATION

José M. N. Leitão and José M. P. Moura

Complexo I, Instituto Superior Técnico,
Av. Rovisco Pais, 1000 LISBOA, PORTUGAL

RESUMÉ

On étudie des estimateurs de phase rapides en utilisant la méthode présentée en [1]. L'idée centrale est la représentation du trait disant facteur d'observation par un train approprié de fonctions de Gauss. Le démodulateur qui en résulte a une structure très simple vis-à-vis des techniques alternatives pour la mise en œuvre de filtres non linéaires. La performance, mesurée par l'erreur carrée moyenne est comparable à celles du filtre de masses ponctuelles [2] et du filtre de Fourier [3]. De l'autre côté, il n'a pas les problèmes associés au filtre de Fourier, à savoir: 1) la négativité introduite quand on tronque la série; 2) le besoin de calculer des fonctions de Bessel; et, 3) la sensibilité aux variations du rapport signal-bruit, ce que se traduit par la croissance du nombre de coefficients de Fourier quand ce rapport là devient plus fort [4].

SUMMARY

The paper studies fast cyclic phase demodulators. They are derived by using the method presented in [1]. The central idea involves the representation of the so called sensor factor by a suitable train of Gaussian functions. The resulting demodulator is of a very simple structure when compared with alternative techniques for the implementation of the optimal nonlinear filter. The mean square performance compares to the ones of the point mass filter of [2] and of the Fourier filter of [3]. However, it does not exhibit the drawbacks associated with the Fourier filter, namely, i) the considerable negativity introduced by the truncated series, ii) the requirement of complex arithmetic and Bessel function evaluation, and iii) the inconsistent behavior for high signal to noise ratio, where the Fourier filter requires a larger number of Fourier terms to be retained [4].

1 - This work was partially supported by INIC, by NATO Scientific Affairs Division under grant no. 1557 and by JNICT under research contract no. 118.79.80.
I - INTRODUCTION

A common practice in nonlinear estimation linearizes the model nonlinearities about the last estimate. By assuming a Gaussian a priori probability density function, the resulting structure is the Extended Kalman-Bucy Filter. The Phase Locked Loop (PLL) is an example of such a class of filters applied to the phase estimation problem. If the mean square estimation error is kept below the typical value of 0.25 rad$^2$, this device can be considered a near optimal estimator. Above that threshold there is a performance degradation. It is due to the fact that, for larger errors (i.e. for stronger noise conditions), the probability density function of the phase process conditioned on the observations can no longer be considered Gaussian. The density becomes multimodal, being impossible to take this information into account by linearizing techniques.

If optimal nonlinear filtering is carried out, by implementing the Bayes’s Law, better performance is expected. Previous studies have corroborated, by means of Monte Carlo simulation, the preceding assertion. Direct implementation of Bayes’s Law leads however to estimation structures that are more complex than the PLL. This complexity, requiring a larger computational effort, is the tradeoff paid for the performance improvement. As will be shown, the scheme proposed buys the performance gains without the computational penalties of more complex structures.

II - MODEL

A discrete scalar Brownian motion is taken as the phase process.

\[
\begin{align*}
x_{n+1} &= x_n + u_n \\
x_1 &= c
\end{align*}
\]

where \(\{u_n\}\) is a white Gaussian sequence of variance \(q_u\), which is related to the parameter \(q_c\) of the continuous corresponding process by \(q_c = \Delta q_u\). The sampling interval \(\Delta\) must be sufficiently small in order to guarantee that the continuous and the discrete models describe essentially the same phenomenon. The initial condition \(x_1\) is a random variable independent of \(\{u_n\}\) and has probability density function \(p(c)\).

The observation model is the vector stochastic difference equation

\[
\begin{bmatrix}
z_{1n} \\
z_{2n}
\end{bmatrix} = \begin{bmatrix}
\cos x_n + v_{1n} \\
\sin x_n + v_{2n}
\end{bmatrix},
\]

where \(\{v_{1n}\}\) and \(\{v_{2n}\}\) are two mutually independent white Gaussian sequences of equal variance \(\sigma\). They are also assumed to be independent of \(\{u_n\}\) and \(c\). The variance \(\sigma\) is related to the equivalent continuous variance \(\sigma_c\) by \(\sigma = \sigma_c/\Delta\).

III - FILTER ALGORITHM

The problem consists on constructing an estimate of \(x_n\) by processing the set of observations \(z_n = [z_{kn}, 1 \leq k \leq n]\). The estimate is defined by choosing a suitable cost function \(L(x; \lambda)\). The solution requires, at each time step, knowledge of the conditional density function \(p(x_n/z_n)\). We shall denote \(p(x_n/z_n)\) by \(P(n)\) and call it the filter density.

The main task lies in the propagation of \(P(n)\). This is recursively accomplished by implementing the equations

\[
\begin{align*}
\text{Prediction:} & \quad P(n+1) = S(n) \ast P(n) \\
\text{Filtering:} & \quad F(n) = C(n) H(n) P(n)
\end{align*}
\]

The symbol \(\ast\) in (3) denotes convolution.

The convolution kernel \(S(n)\), depending exclusively on the phase dynamics, apart a normalizing constant, is

\[
S(n) = \exp[-\frac{1}{2q}(x_{n+1}^2 - x_n^2)].
\]

The result of convolution, \(P(n+1)\), is the predictor density. The symbol \(\ast\) in (4) represents a pointwise multiplication of \(H(n)\) by \(P(n)\). The function \(H(n)\) is called the sensor factor. It takes into account the observation structure, being given by the expression

\[
H(n) = \exp[(z_{1n} \cos x + z_{2n} \sin x)/\xi].
\]

The term \(C(n)\) in (6) is a normalization constant. The a priori knowledge about the phase process is given by \(P(1) = p(c)\).

In order to implement (3) and (4) in a digital computer a suitable representation
of the densities has to be used. In [1], a Gaussian sum description of $H(n)$ has been proposed. The procedure is as follows:

1. **Sensor Factor Representation**

   Represent $H(n)$ as an infinite set of Gaussian terms given by
   
   $$ R(n) = \sum_{i=-\infty}^{+\infty} \exp \left[ -\frac{1}{2\sigma_n^2} (x-\eta_n^{H})^2 \right]. $$

   (7)

   The means $\eta_{ln}^H$ are computed according to
   
   $$ \eta_{ln}^H = \eta_{ln}^H + 2\pi i \quad i=\ldots,-1,0,1,\ldots $$

   (8)

   where $\eta_{ln}^H$, the abcissa of the maximum of $H(n)$ in the interval $[-\pi, +\pi)$, is given by
   
   $$ \eta_{ln}^H = \tan^{-1} \frac{z_{ln}}{z_{ln}} $$

   (9)

   The variance $\sigma_n^H$ is obtained by fitting one Gaussian function to each period of $H(n)$ as shown in fig. 1. Its value is
   
   $$ \sigma_n^H = 2(\pi/8) \sqrt{z_{ln}^2 + z_{ln}^2} $$

   (10)

   FIG. 1 - Fitting one Gaussian function (---)

   to each period of the sensor factor (----)

   As a consequence of this representation of $H(n)$ and from the fact that $S(n)$ is Gaussian, all the densities will be Gaussian sums, provided $P(1)$, the initial condition, has the same form.

2. **Filtering**

2.1 - **Multiplication**

   Assume $P(n)$ is given by
   
   $$ P(n) = \sum_{i=1}^{N_p} P_n^{i} \exp \left[ -\frac{1}{2\sigma_n^2} (x-\eta_n^{P})^2 \right], $$

   (11)

   where $P_n^{i}$ are the predictor weighting factors, $\sigma_n^P$ is the predictor common variance and $\eta_n^{P}$ are the predictor means.

   In order to build a finite representation of $F(n)$, each term of $P(n)$ multiplies only the $J$ nearest terms of $H(n)$. The integer $J$, herein referred to as the multiplication parameter, is adjusted experimentally. The result of multiplication is
   
   $$ F(n) = \sum_{i=1}^{N_F} F_n^{i} \exp \left[ -\frac{1}{2\sigma_n^F} (x-\eta_n^{F})^2 \right], $$

   (12)

   with $\sigma_n^F = \sigma_n^P \sigma_n^H / (\sigma_n^P + \sigma_n^H)$

   (13)

   $$ F_n^{i} = F_n^{i} \exp \left[ -\frac{1}{2(\sigma_n^P + \sigma_n^H)} (\eta_n^{P} - \eta_n^{H})^2 \right]. $$

   (14)

   $\eta_n^{iJ}$ is explained in fig. 2.

   2.2 - **Agglutination**

   After multiplication, some of the modes of $F(n)$ fall very near each other. When $|\eta_n^{P} - \eta_n^{F}| < \beta_1$, the modes $i$ and $j$ are agglutinated. The resulting mean is the weighted average, the weighting factor being the sum of the weights. The variance remains unchanged. The threshold $\beta_1$, referred to as the agglutination parameter, is adjusted experimentally.

   $2\pi(J/2)$

   $2\pi(J/2)$

   FIG. 2 - Explaining the meaning of $\eta_n^{iJ}$.
2.3 - Elimination

Modes of $F(n)$ such that $K_{i+1}^F < \beta_2$ are eliminated. The constant $\beta_2$ is also set by simulation and is designated by the elimination parameter.

The elimination and agglutination procedures lead to a final number of terms of $F(n)$

$$N_n^F < N_n^F$$

(16)

The number $N_n^F$ is called the filter dimension.

3 - Prediction

Convolution of $F(n)$ by $S(n)$ leads to the predictor density

$$P(n+1) = \sum_{i=1}^{N_n^P} K_{i+1}^P \exp \left[-\frac{1}{2\sigma_{n+1}^2} (x - \eta_{i+1}^P) \right]$$

(17)

where

$$n_{n+1}^P = n_n^P$$

(17.a)

$$K_{i+1}^P = K_i^P$$

(17.b)

$$\eta_{i+1}^P = \eta_i^P$$

(17.c)

$$\sigma_{n+1}^P = \sigma_n^P + q$$

(17.d)

Formulae (16) and (17.a) indicate the possible change of the filter dimension at each iteration - an increase or a decrease.

4 - Estimation Criterion

Our interest lies on the cyclic phase demodulation. The loss function is periodic. The function $L(x_n - \tilde{x}_n) = 2[1 - \cos(x_n - \tilde{x}_n)]$ is a suitable choice [2].

By minimizing the conditional expectation

$$E[L(x_n - \tilde{x}_n) | Z_n] = \int_{-\infty}^{+\infty} L(x_n - \tilde{x}_n) F(n) dx_n$$

(18)

with $F(n)$ given by (14), one is conducted to the optimal estimate

$$\tilde{x}_n = \frac{\sum_{i=1}^{N_n^F} K_{i+1}^F \sin \eta_{i+1}^F}{\sum_{i=1}^{N_n^F} K_{i+1}^F \cos \eta_{i+1}^F}$$

(19)

IV - EXPERIMENTAL EVALUATION

The filter behavior is to be observed for different noise conditions.

The filter performance is quantified in terms of the mean square estimation error modulo $2\pi$. The noise condition is measured in terms of the theoretical performance that would be achieved if the observations were linear. This performance is the error variance in steady state, $\sigma^2$, which is supplied by the Riccati equation corresponding to the linearized continuous model. It is computed by

$$\sigma^2 = \frac{1}{\Delta n} \sigma^2$$

The linearized filter time constant $\tau$ is

$$\tau = \sqrt{\frac{\Delta}{\sigma^2}}$$

The sampling interval $\Delta$, on the basis of which the discrete model is built, is such that $\Delta = 0.1\tau$, see [2] for details.

For a given value of $R$, the mean square errors modulo $2\pi$, $\tilde{\sigma}^2$, are computed by running 500 sample functions, each with 130 points. The first 30 points of each run are discarded in order to measure only the steady state errors.

The value of $\tilde{\sigma}^2_M$ is obtained through

$$\tilde{\sigma}^2_n = \frac{1}{500} \sum_{i=1}^{500} ([x_{in} - \tilde{x}_{in}] \text{modulo } 2\pi)^2$$

(20)

and

$$\tilde{\sigma}^2 = \frac{1}{100} \sum_{n=31}^{130} \tilde{\sigma}^2_n$$

(21)

The PLL and the NLF are evaluated simultaneously and start with the same initial condition.

1. A Simulation Result

Fig. 3 summarizes the results obtained for $R = 0$(dB) and the following filter parameters:

$$J = 2, \beta_1 = 0.25, \beta_2 = 0.001$$

The evolutions of $\tilde{\sigma}^2_n$(PLL) and $\tilde{\sigma}^2_n$(NLF) are plotted together with the average value of the filter dimension for each value of $n$,
ALGORITHM STRUCTURES FOR CYCLIC PHASE ESTIMATION

\[
\frac{N^F}{N_n} = \frac{1}{500} \sum_{i=1}^{500} N^F_n
\]  
(22)

\[
\begin{align*}
\delta^2_M &= 10 \log_{10} \frac{\delta^2_M(PLL)}{\delta^2_M(NLF)} - 10 \log_{10} \frac{\delta^2_F}{\delta^2_{n}} \\
\end{align*}
\]

one computes \( \delta^2_M = 0.65 \text{ dB} \), which is of the same order of magnitude as the ones reported in [2].

Since the error process is a diffusion and the NLF algorithm allows the expansion of \( F(n) \) on the real line, the filter dimension \( N^F_n \) increases with time as can be seen by the plot of \( N^F_n \) in fig. 3.

This is unsuitable for practical reasons: the error performance is good but the filter exhibits an increasing complexity. A simplification is required to keep the filter dimension constant and small.

2. Algorithm Simplification.

The Matching Technique.

It was observed by Youssef in [5], by using a Fourier series expansion as a representation of \( H(n) \), that \( F(n) \) could be replaced, at each iteration, by only one Gaussian function, without significant loss of performance.

Let us denote this new representation of \( F(n) \) by

\[
F(n) = \frac{1}{\sqrt{2\pi} \sigma^F_n} \exp \left[ -\frac{1}{2\sigma^F_n}(x-\eta^F_n)^2 \right].
\]  
(23)

According to [4], the parameters of \( F(n) \) will be such that \( F(n) \) and \( F(n) \) have the same first and second optimum moments relative to the same error criterion (the cyclic loss function). This condition corresponds to the matching equation (24).

\[
\int_{-\infty}^{+\infty} e^{jmx} [F(n) - F(n)]dx = 0 .
\]  
(24)

The new algorithm is now described:

A - The sensor factor representation is the same as in the preceding algorithm.

B - The predictor density \( P(n+1) \), obtained by convolution of two Gaussian functions \( S(n) \) and \( F(n) \), is now Gaussian, with parameters

\[
\sigma^P_{n+1} = \sigma^F_n + q
\]

\[
\eta^P_{n+1} = \eta^F_n
\]

(25-a)

(25-b)

C - When multiplying \( P(n) \) by \( H(n) \) only the \( J \) nearest terms of \( H(n) \) relative to the unique term of \( P(n) \) are considered. This leads to a sum of \( J \) Gaussian terms for \( F(n) \).

D - Matching - By application of the matching equation (24) one is conducted to

\[
\eta^F_n = tg^{-1} \frac{\delta^2_M}{\delta^2_F}
\]

\[
\delta^F_M = \delta^F_n - \ln (\sigma^2 + \sigma^2_c)
\]

(26)

(27)

where

\[
\begin{align*}
\delta^2_F &= \sum_{i=1}^{J} K^F_{i} \sin \eta^F_n \sin \eta^F_{i} \\
\delta^2_M &= \sum_{i=1}^{J} K^F_{i} \cos \eta^F_n \cos \eta^F_{i}
\end{align*}
\]  
(28-a)

(28-b)

FIG. 3 - Simulation results for the PLL and the NLF, corresponding to \( R = 0 \text{ dB} \).
The cyclic estimate is now

$$\tilde{\theta}_n = \eta_n^p$$  \hspace{1cm} (29)

Since $\eta_n^p$ belongs to $[-\pi, \pi)$ and $\eta_{n+1}^p = \eta_n^p$, the modes of $\tilde{\theta}(n)$ considered in the multiplication are the one in $[-\pi, \pi)$ and the adjacent ones. Next experiments correspond to $J=3$. Fig. 4 is a block diagram of the simplified (SNLF) algorithm just described.

FIG. 4 - SNLF Block Diagram

V - SNLF PERFORMANCE EVALUATION

In order to see how this new filter (SNLF) compares with the preceding one (NLF), the same noise testing sequences ($R=0$ dB) were used. A good agreement between $\tilde{\theta}_n^2$ (NLF) and $\tilde{\theta}_n^2$ (SNLF) was observed. The dot points in fig. 3 represent only the values of $\tilde{\theta}_n^2$ (SNLF) for $n=5m$, $m=7, \ldots, 26$. The final mean square error was $\tilde{\theta}_M^2$ (SNLF) = 1.372 which is approximately the value of $\tilde{\theta}_M^2$ (NLF). Thus we conclude that the same performance is achieved by setting $J=3$ in the SNLF.

The continuous lines of fig. 5 represent the theoretical modulo $2\pi$ error variances versus $R$, for the PLL and the ideal linear filter. The upper dot points are the simulation results for the PLL; the lower ones correspond to the SNLF. Their respective values can be seen in table 1. One can see that: 1) the values of $\Delta \tilde{\theta}_M^2$, for $R=1,0,1$, are of the same order of magnitude of the ones reported in [2]; 2) $\Delta \tilde{\theta}_M^2$ is consistently positive for all values of $R$, tending to zero as $R$ approaches the region of linear behavior.

The dimension keeping constant. Point 2) emphasizes the difference relative to the Fourier filter which requires an increasing dimensionality as the SNR increases.

FIG. 5 - Simulation results (PLL, SNLF).

Table 1 - Mean square errors (modulo $2\pi$), (dB)

<table>
<thead>
<tr>
<th>$R$(dB)</th>
<th>$\tilde{\theta}_M^2$(P.L.L.)</th>
<th>$\tilde{\theta}_M^2$(N.L.F.)</th>
<th>Improvement $\Delta \tilde{\theta}_M^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6</td>
<td>5.371</td>
<td>5.402</td>
<td>0.031</td>
</tr>
<tr>
<td>-5</td>
<td>3.920</td>
<td>4.240</td>
<td>0.320</td>
</tr>
<tr>
<td>-4</td>
<td>2.711</td>
<td>3.026</td>
<td>0.315</td>
</tr>
<tr>
<td>-3</td>
<td>1.398</td>
<td>1.825</td>
<td>0.427</td>
</tr>
<tr>
<td>-2</td>
<td>0.105</td>
<td>0.644</td>
<td>0.539</td>
</tr>
<tr>
<td>-1</td>
<td>1.071</td>
<td>0.417</td>
<td>0.654</td>
</tr>
<tr>
<td>0</td>
<td>2.018</td>
<td>1.372</td>
<td>0.646</td>
</tr>
<tr>
<td>1</td>
<td>2.823</td>
<td>2.194</td>
<td>0.629</td>
</tr>
<tr>
<td>2</td>
<td>3.296</td>
<td>2.885</td>
<td>0.401</td>
</tr>
</tbody>
</table>

The computation effort associated with the SNLF is, loosely speaking, equivalent to: 1) three Kalman-Bucy filter steps operating in parallel, the associated weighting factors involving three exponentiations; 2) one matching operation - equations (26),(27) and (28-a,b) - and 3) a Kalman-Bucy predictor step.
VI - CONCLUSION

The paper studies a cyclic phase demodulator designed by application of the optimal nonlinear filtering techniques. The crucial step lies on a representation of the sensor factor, whereby each of its periods is adjusted by a single Gaussian. The resulting filter's performance compares to the one of the point mass filter [2]. Since this filter propagates at each step a line density, the number of Gaussian parameters increases with time. To obtain a further simplified algorithm, an adaptation of the filtered density by a single Gaussian function is carried out at each step. The simplified nonlinear filter exhibits no noticeable loss of performance, nor any of the difficulties reported for the Fourier filter [4].

BIBLIOGRAPHY


