TABLE I

<table>
<thead>
<tr>
<th>N</th>
<th>SCER for Various Schemes and N</th>
</tr>
</thead>
<tbody>
<tr>
<td>(m)</td>
<td>8 (200)</td>
</tr>
<tr>
<td>SM</td>
<td>48.86</td>
</tr>
<tr>
<td>2C (a)</td>
<td>53.81</td>
</tr>
<tr>
<td>2C (b)</td>
<td>53.81</td>
</tr>
<tr>
<td>2C 6 1C</td>
<td>51.32</td>
</tr>
<tr>
<td>2CRM (a)</td>
<td>47.54</td>
</tr>
<tr>
<td>2CRM (b)</td>
<td>47.46</td>
</tr>
</tbody>
</table>

Note: $m$—number of FFT's averaged, $a$—coefficient resolution 12 bits, $b$—coefficient resolution 10 bits.

Fig. 3. 2C processing with SM multiplier in an FFT butterfly.

Fig. 4. Reduced multiplier array with rounding off.

If, instead of the 2C to SM conversion in Fig. 3, a 2C to 1C conversion is made, this gives a saving in hardware, though the SCER goes down, as shown in Table I for the scheme 2CRM (2C with reduced multiplier) with 8 columns (about 30 percent array elements) removed from a 12 X 12-bit SM multiplier. This loss of SCER reduces with increasing $N$. It is felt that this is due to increased interleaving of errors for larger $N$.

Thus, the techniques described here will help the designer of the FFT hardware in choosing various schemes considering hardware cost and computational accuracy.

REFERENCES


Improved Approximation of Bias in Squared Coherence Estimates for Weakly Smooth Spectra

PETER LÖHNBERG

Abstract—Bias in squared coherence estimates for normal processes is approximated by a function of coherence and second derivatives divided

Manuscript received February 17, 1977; revised September 29, 1977 and November 23, 1977.

The author is with the Department of Electrical Engineering, Twente University of Technology, Enschede, The Netherlands.

0096-3518/78/0400-0172$00.75 © 1978 IEEE
The squared coherence at a certain frequency is defined as
\[ C = \frac{G_{xy} G_{yx}}{G_{xx} G_{yy}} \]  
(1)
with the spectra: cross spectral density \( G_{xy} \) with complex conjugate \( G_{x}^{\ast} \) and power spectral densities \( G_{xx} \) and \( G_{yy} \) of stationary real random processes \( x \) and \( y \).

Spectra can be estimated under ergodicity assumption from a realization of \( x \) and \( y \) over time \( T \). An efficient procedure consists of dividing this interval into \( N \) equally long segments, forming products per segment of FFT's and their complex conjugates, multiplication of inverse FFT's by a lag window, and averaging over FFT's of all \( N \) results. Substitution of estimates \( \hat{G}_{xx}, \hat{S}_{xy}, \hat{G}_{yy} \) into (1) gives estimate \( \hat{C} \).

For optimal window selection and bias correction it is desirable that bias [\( \hat{C} \)] and var [\( \hat{C} \)] are known for arbitrary spectra. Bias [\( \hat{C} \)] was approximated \(^1\) in (1) with weak restrictions on spectral variations over the window length, defined as the width of the range of frequencies \( f \) for which the Fourier transform \( W(f) \) of the lag window has significant values. Stronger restrictions were made in an exact derivation \(^2\) and apparently white processes were used in experiments \(^3\), \(^4\). References \(^2\)-\(^4\) showed bias [\( \hat{C} \)] to be a decreasing function of \( C \) reaching 0 for \( C = 1 \). But according to \(^1\), for these stronger restrictions bias [\( \hat{C} \)] would be independent of \( C \). So use of \(^1\) might cause serious errors.

Therefore, an improved expression for bias [\( \hat{C} \)] is required which indicates effects of spectral variations over the window width. Extension of the exact derivation to weaker restrictions causes mathematical problems, whereas repetition of experiments for a sufficient number of combinations of spectra is hardly feasible. A more accurate approximation for even stronger spectral variations than \(^1\) is possible however; for \(^1\) only a first-order Taylor approximation of [\( \hat{C}(G_{xy}, \hat{G}_{yx}) \)] at point \( \hat{G}_{xy}, \hat{G}_{yx} \) was used. It is only slightly more difficult to approximate \( C \) as a function of all spectral estimates by a second-order Taylor series. This expresses bias [\( \hat{C} \)] into first and second moments of spectral estimates, for which expressions are known. These moments were already used maximally in \(^1\) to approximate \(^2\) var [\( \hat{C} \)]. Therefore, only bias [\( \hat{C} \)] will be approximated here such as to make maximal use of these moments.

**Approximate Bias of Squared Coherence Estimate**

For notational simplicity, define for estimates of arbitrary quantities \( X \) and \( Y \) the normalizations \( B[X] = \text{bias} [\hat{X}|\hat{X}, \hat{Y}] \) and \( B[X|Y] = \text{cov} [\hat{X}|\hat{X}, \hat{Y}] \).

Then, by second-order Taylor expansion of \( B[\hat{C}(G_{xy}, \hat{G}_{yx}, \hat{G}_{xx}, \hat{G}_{yy})] \) at point \( \hat{G}_{xy}, \hat{G}_{yx}, \hat{G}_{xx}, \hat{G}_{yy} \) follows
\[
B[\hat{C}] \approx B[\hat{C}_{xy}] + B[\hat{G}_{xy}] - B[\hat{G}_{yx}] - B[\hat{G}_{yy}] \\
+ B[\hat{G}_{xy}] B[\hat{G}_{yx}] + V[\hat{G}_{xy}, \hat{G}_{yx}], \\
- B[\hat{G}_{xy}] B[\hat{G}_{xx}] - V[\hat{G}_{xy}, \hat{G}_{xx}], \\
- B[\hat{G}_{xy}] B[\hat{G}_{yy}] - V[\hat{G}_{xy}, \hat{G}_{yy}] \\
(2)
\]

Bias of spectral estimates may be derived similarly to \(^5\). This requires small variation of the spectra over frequency width \( N/T \) and of second derivatives to frequency \( G_{xy}, \hat{G}_{xx}, \) and \( \hat{G}_{yy} \) over the window width. Then
\[
B[\hat{G}_{xy}] \approx SG_{xy} G_{yx} \\
(3)
\]
and analogous expressions for \( \hat{G}_{xx} \) and \( \hat{G}_{yy} \) with systematic error constant
\[
S = \frac{1}{2} \int_{-\infty}^{\infty} f^2 W(f) \, df. \\
(4)
\]

Covariance of spectral estimates also has been derived \(^1\).

For that it is required that \( G_{ac} \) and \( G_{bd} \) of arbitrary normal processes \( a, b, c, \) and \( d \) vary little over the window width. Then
\[
V[\hat{G}_{ab}, \hat{G}_{cd}] \approx RG_{ac} G_{db}/R_{ab} G_{cd} \\
(4)
\]
with random error constant
\[
R = \frac{1}{T} \int_{-\infty}^{\infty} W^2(j) \, df < 1. \\
(4)
\]

\( R \) is obtainable for usual windows by substitution of the values of \( "D" \) from \(^5\) into \( S = "D"/(2m)^2 \).

Substitution of (3), (4) into (2) yields
\[
B[\hat{C}] \approx SG_{xy} G_{yx} + SG_{yx} G_{xy} - SG_{xx} G_{xx} - SG_{yy} G_{yy} \\
+ S^2 G_{xy} G_{yx}/G_{xx} G_{yy} + R/C \\
- S^2 G_{xy} G_{xx}/G_{xy} G_{xx} - R - S^2 G_{yy} G_{yy}/G_{xy} G_{yy} - R \\
- S^2 G_{yx} G_{xx}/G_{yx} G_{xx} - R - S^2 G_{yx} G_{yy}/G_{yx} G_{yy} - R \\
+ S^2 G_{xx} G_{xy}/G_{xx} G_{xy} + R + S^2 G_{yy} G_{xx}/G_{yy} G_{xx} + R \\
+ S^2 G_{xx} G_{yy}/G_{xx} G_{yy} + R/C. \\
(5)
\]

Substitution into (5) of the result from (1) \( G_{xy} = \sqrt{G_{xx} G_{yy}} \) shows that relative spectral second and not first derivatives influence bias. As expected, bias is decreased by decreasing \( S \) or \( R \). However, for fixed \( T \) and window type decrease of either increases the other. So for each \( T \), an optimal window for minimal coherence bias might be derived from (5) with appropriate spectral information. By also taking into account var[\( \hat{C} \)], the mean-square error in \( \hat{C} \) might be minimized instead.

**Comparison with Literature for Very Smooth Spectra**

Several simplifications are needed for comparison of (5) and \(^1\) to \(^2\)-\(^4\). For \(^1\) \( G_{xy} \) was required to be smooth, that is \( S|G_{xy}/G_{yy}| \ll 1 \). Then the fifth term in (5) can be
\[^4\]Formula (A.9.1.28) in \(^1\).
omitted. Furthermore, by the rough approximation, only the first, second, and sixth terms in (5) were considered for [1].

References [2]--[4] concern very smooth spectra, that is $S(G_{yy}/G_{xy}) < R$ and $G_{yy}$ and $G_{xy}$ similar. Then all systematic error terms containing $S$ and second derivatives may be omitted, and only the random error terms in (5) remain. So according to (5)

$$\text{bias} [\hat{C}]_{\text{very smooth}} = CB[\hat{C}]_{\text{very smooth}} \approx R(1 - C)^2 = F(C).$$

(6)

The version of (1) for very smooth spectra is

$$\text{bias} [\hat{C}]_{\text{very smooth}} \approx R.$$  

(7)

Approximation $F(C)$ in (6) is $R$ for $C = 0$ and $0$ for $C = 1$. This agrees with theoretical bias $[\hat{C}]_I$, [2], with simulation results for $R = \frac{1}{4}$, [4], and even with degenerate procedure effect $\hat{C} = 1$ for $R = 1$. But for $0 < C < 1$, $F(C)$ is lower than bias according to all of these. These lower values must be due to approximation (2) and possibly (4). Nevertheless, $F(C)$ differs less from bias $[\hat{C}]$ according to [2], [4] than approximation $R$ in (7) does for any $R$ and $C > 0$.

Reference [3] resulted in empirical estimates of bias $[\hat{C}]$ with fit $R(1 - \hat{C})$. For very low $C$ these results are slightly lower, and for high $C < 1$, higher than bias $[\hat{C}]$ according to [2], [4] and (6). These deviations increase with $R$. They must at least partly be due to the estimation in [3] of bias $[\hat{C}]$ by the inverse $Z^{-1}$ of the bias in $Z(C) = \frac{1}{4} \ln (1 + \sqrt{C}) / (1 - \sqrt{C})$. This procedure implies use of the approximation $F(C) = Z^{-1} \{E[Z(C)]\}$. The resulting approximation error in bias $[\hat{C}]$ increases with the second derivative of $Z(C)$ to $\hat{C}$ and with $\text{var} [\hat{C}]$, which is proportional to $R(1)^2$ These expected increases agree with the indicated ones of the actual deviations. Therefore, [3] will not be used for comparison.

CONCLUSION

Equation (5) approximates bias $[\hat{C}]$ of normal processes for values and second derivatives of spectra varying slightly over the window width. The derivation uses a more extensive and higher order Taylor approximation than [1]. Therefore, result (5) is expected to be more accurate than [1]. This might be supported by very smooth spectral version (6) of (5) agreeing closer with [2], [4] than corresponding version (7) of (1). A reason for the deviation of [3] from [2], [4], and (6) has been indicated.

ACKNOWLEDGMENT

The author wishes to thank Prof. E. W. Grønvevd, G. H. L. M. Heideman, and J. de Haas for their sincere interest in this subject, and Miss G. G. M. Steijlen for typing the manuscript.

REFERENCES


Notes on Linear Image Restoration by Maximizing the A Posteriori Probability

H. J. Trussell

Abstract—A distinction is made between the minimum mean-square error estimate derived for the linear case of maximum a posteriori restoration and the common Wiener estimate. The linear MAP filter is shown to produce good results with less a priori knowledge than is required for the Wiener filter.

Minimum mean-square error filters can be derived from both a priori and a posteriori knowledge. The minimum mean-square error filter derived from a priori knowledge is usually referred to as a Wiener filter [1]. This filter is a standard against which most other image restoration/enhancement filters are compared. The minimum mean-square error filter derived from a posteriori knowledge is the same as the filter which maximizes the a posteriori probability for a linear imaging system. Because of its relation to the maximum a posteriori (MAP) estimation method, we will refer to this filter as the linear MAP filter. This correspondence shows that restorations of quality comparable to Wiener restorations can be obtained by using the linear MAP filter. In actual implementation, the linear MAP restoration uses less a priori knowledge than the Wiener restoration; furthermore, it is less sensitive to perturbations of that a priori knowledge.

Hunt [2] derived the maximum a posteriori (MAP) equation for nonlinear image restoration. For the linear case (1), using lexicographic notation [3], the imaging system is described by

$$g = Hf + n$$

(1)

where

$\hat{f} = $ is the original image drawn from a Gaussian ensemble with mean $f$ and covariance $R_f$;

$H$ is the point-spread function matrix;

$n$ is zero-mean Gaussian noise with variance/covariance $R_n$;

$g$ is the recorded image.

The linear MAP estimate is then given by

$$\hat{f}_{MAP} = (R_f^{-1} + H^T R_n^{-1} H)^{-1}(H^T R_n^{-1} g + R_f^{-1} \hat{f}).$$

(2)

The Wiener estimate with ensemble mean $\hat{f}$ included can be shown [see [1] and model (1)] to be

$$\hat{f}_w = (R_f H^T + \hat{f} \hat{f}^T H^T)(H R_f H^T + \hat{f} \hat{f}^T H^T + R_n)^{-1} g.$$  

(3)

If $\hat{f} = 0$, as is usually assumed in Wiener estimation, then $\hat{f}_{MAP} = \hat{f}_w$.

The difference between the linear MAP filter and the usual form of the Wiener filter is the nonzero mean $\hat{f}$. The model used for the derivation of the Wiener filter assumes a zero-mean (or a constant-mean which can be subtracted). This implies that all spatial variation of the image ensemble is accounted for by the covariance matrix $R_f$. Since the mean $f$ has been assumed to have an extremely simple structure, the covariance must necessarily have a great deal of structure.

The model used for deriving the linear MAP filter assumes a nonstationary mean $\hat{f}$. For computation it is assumed that $f_D = f - \hat{f}$ is a stationary process. Since some of the ensemble

Manuscript received February 14, 1977; revised May 17, 1977 and November 1, 1977.

The author is with Los Alamos Scientific Laboratory, University of California, Los Alamos, NM 87545.