

A power spectral density estimation method and apparatus.

TECHNICAL FIELD

The present invention relates to a bias compensated spectral estimation method and apparatus based on a parametric autoregressive model.

BACKGROUND OF THE INVENTION

The present invention may be applied, for example, to noise suppression [1, 2] in telephony systems, conventional as well as cellular, where adaptive algorithms are used in order to model and enhance noisy speech based on a single microphone measurement.

Speech enhancement by spectral subtraction relies on, explicitly or implicitly, accurate power spectral density estimates calculated from the noisy speech. The classical method for obtaining such estimates is periodogram based on the Fast Fourier Transform (FFT). However, lately another approach has been suggested, namely parametric power spectral density estimation, which gives a less distorted speech output, a better reduction of the noise level and remaining noise without annoying artifacts ("musical noise"). For details on parametric power spectral density estimation in general, see [3, 4].

In general, due to model errors, there appears some bias in the spectral valleys of the parametric power spectral density estimate. In the output from a spectral subtraction based noise canceler this bias gives rise to an undesirable "level pumping" in the background noise.

SUMMARY OF THE INVENTION

An object of the present invention is a method and apparatus that eliminates or reduces this "level pumping" of the background

noise with relatively low complexity and without numerical stability problems.

This object is achieved by a method and apparatus in accordance with the enclosed claims.

- 5 The key idea of this invention is to use a data dependent (or adaptive) dynamic range expansion for the parametric spectrum model in order to improve the audible speech quality in a spectral subtraction based noise canceler.

BRIEF DESCRIPTION OF THE DRAWINGS

- 10 The invention, together with further objects and advantages thereof, may best be understood by making reference to the following description taken together with the accompanying drawings, in which:

- 15 FIGURE 1 is a block diagram illustrating an embodiment of an apparatus in accordance with the present invention;
- FIGURE 2 is a block diagram of another embodiment of an apparatus in accordance with the present invention;
- 20 FIGURE 3 is a diagram illustrating the true power spectral density, a parametric estimate of the true power spectral density and a bias compensated estimate of the true power spectral density;
- FIGURE 4 is another diagram illustrating the true power spectral density, a parametric estimate of the true power spectral density and a bias compensated estimate of the true power spectral density;
- 25 FIGURE 5 is a flow chart illustrating the method performed by the embodiment of Fig. 1; and

FIGURE 6 is a flow chart illustrating the method performed by the embodiment of Fig. 2.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Throughout the drawings the same reference designations will be used for corresponding or similar elements.

Furthermore, in order to simplify the description of the present invention, the mathematical background of the present invention has been transferred to the enclosed appendix. In the following description numerals within parentheses will refer to corresponding equations in this appendix.

Figure 1 shows a block diagram of an embodiment of the apparatus in accordance with the present invention. A frame of speech $\{x(k)\}$ is forwarded to a LPC analyzer (LPC analysis is described in, for example, [5]). LPC analyzer 10 determines a set of filter coefficients (LPC parameters) that are forwarded to a PSD estimator 12 and an inverse filter 14. PSD estimator 12 determines a parametric power spectral density estimate of the input frame $\{x(k)\}$ from the LPC parameters (see (1) in the appendix). In Fig. 1 the variance of the input signal is not used as an input to PSD estimator 12. Instead a unit signal "1" is forwarded to PSD estimator 12. The reason for this is simply that this variance would only scale the PSD estimate, and since this scaling factor has to be canceled in the final result (see (9) in the appendix), it is simpler to eliminate it from the PSD calculation. The estimate from PSD estimator 12 will contain the "level pumping" bias mentioned above.

In order to compensate for the "level pumping" bias the input frame $\{x(k)\}$ is also forwarded to inverse filter 14 for forming a residual signal (see (7) in the appendix), which is forwarded to another LPC analyzer 16. LPC analyzer 16 analyses the residual signal and forwards corresponding LPC parameters (variance and filter coefficients) to a residual PSD estimator 18, which forms

a parametric power spectral density estimate of the residual signal (see (8) in the appendix).

Finally the two parametric power spectral density estimates of the input signal and residual signal, respectively, are multiplied by each other in a multiplier 20 for obtaining a bias compensated parametric power spectral density estimate of input signal frame $\{x(k)\}$ (this corresponds to equation (9) in the appendix).

Example

The following scenario is considered: The frame length $N=1024$ and the AR (AR=AutoRegressive) model order $p=10$. The underlying true system is modeled by the ARMA (ARMA=AutoRegressive-Moving Average) process

$$x(k) = \frac{1 - z^{-1} + 0.9z^{-2}}{1 - 3.0z^{-1} + 4.64z^{-2} - 4.44z^{-3} + 2.62z^{-4} - 0.77z^{-5}} e(k)$$

where $e(k)$ is white noise.

Figure 3 shows the true power spectral density of the above process (solid line), the biased power spectral density estimate from PSD estimator 12 (dash-dotted line) and the bias compensated power spectral density estimate in accordance with the present invention (dashed line). From Fig. 3 it is clear that the bias compensated power spectral density estimate in general is closer to the underlying true power spectral density. Especially in the deep valleys (for example for $\omega/(2\pi) \approx 0.17$) the bias compensated estimate is much closer (by 5 dB) to the true power spectral density.

In a preferred embodiment of the present invention a design parameter γ may be used to multiply the bias compensated estimate. In Fig. 3 parameter γ was assumed to be equal to 1. Generally γ is a positive number near 1. In the preferred embodiment γ has the value indicated in the algorithm section of the appendix. Thus, in this case γ differs from frame to frame. Fig. 4 is a diagram similar to the diagram in Fig. 3, in which

the bias compensated estimate has been scaled by this value of γ .

The above described embodiment of Fig. 1 may be characterized as a frequency domain compensation, since the actual compensation is performed in the frequency domain by multiplying two power spectral density estimates with each other. However, such an operation corresponds to convolution in the time domain. Thus, there is an equivalent time domain implementation of the invention. Such an embodiment is shown in Fig. 2.

In Fig. 2 the input signal frame is forwarded to LPC analyzer 10 as in Fig. 1. However, no power spectral density estimation is performed with the obtained LPC parameters. Instead the filter parameters from LPC analysis of the input signal and residual signal are forwarded to a convolution circuit 22, which forwards the convoluted parameters to a PSD estimator 12', which forms the bias compensated estimate, which may be multiplied by γ . The convolution step may be viewed as a polynomial multiplication, in which a polynomial defined by the filter parameters of the input signal is multiplied by the polynomial defined by the filter parameters of the residual signal. The coefficients of the resulting polynomial represent the bias compensated LPC-parameters. The polynomial multiplication will result in a polynomial of higher order, that is, in more coefficients. However, this is no problem, since it is customary to "zero pad" the input to a PSD estimator to obtain a sufficient number of samples of the PSD estimate. The result of the higher degree of the polynomial obtained by the convolution will only be fewer zeroes.

Flow charts corresponding to the embodiments of Figs. 1 and 2 are given in Figs. 5 and 6, respectively. Furthermore, the corresponding frequency and time domain algorithms are given in the appendix.

A rough estimation of the numerical complexity may be obtained as follows. The residual filtering (7) requires $\sim Np$ operations (sum + add). The LPC analysis of $e(k)$ requires $\sim Np$ operations to form

the covariance elements and $\sim p^2$ operations to solve the corresponding set of equations (3). Of the algorithms (frequency and time domain) the time domain algorithm is the most efficient, since it requires $\sim p^2$ operation for performing the convolution. To summarize, the bias compensation can be performed in $\sim 2p(N+p)$ operations/frame. For example, with $n=256$ and $p=10$ and 50% frame overlap, the bias compensation algorithm requires approximately $0,5 \times 10^6$ instructions/s.

In this specification the invention has been described with reference to speech signals. However, the same idea is also applicable in other applications that rely on parametric spectral estimation of measured signals. Such applications can be found, for example, in the areas of radar and sonar, economics, optical interferometry, biomedicine, vibration analysis, image processing, radio astronomy, oceanography, etc.

It will be understood by those skilled in the art that various modifications and changes may be made to the present invention without departure from the spirit and scope thereof, which is defined by the appended claims.

REFERENCES

- [1] S.F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral subtraction", IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-27, April 1979, pp 113-120.
- [2] J.S. Lim and A.V. Oppenheim, "Enhancement and Bandwidth Compression of Noisy Speech", Proceedings of the IEEE, Vol. 67, No. 12, December 1979, pp. 1586-1604.
- [3] S.M. Kay, Modern Spectral estimation: Theory and Application, Prentice Hall, Englewood Cliffs, NJ, 1988, pp 237-240.
- [4] J.G. Proakis et al, Advanced Digital Signal Processing, Macmillan Publishing Company, 1992, pp. 498-510.
- [5] J.G. Proakis, Digital Communications, MacGraw Hill, 1989, pp. 101-110.
- [6] P. Händel et al, "Asymptotic variance of the AR spectral estimator for noisy sinusoidal data", Signal Processing, Vol. 35, No. 2, January 1994, pp. 131-139.

8
APPENDIX

Consider the real-valued zero mean signal $\{x(k)\}$, $k = 1, \dots, N$ where N denotes the frame length ($N = 160$, for example). The autoregressive spectral estimator (ARSPE) is given by, see [3, 4]

$$\hat{\Phi}_x(\omega) = \frac{\hat{\sigma}_x^2}{|\hat{A}(e^{j\omega})|^2} \quad (1)$$

where ω is the angular frequency $\omega \in (0, 2\pi)$. In (1), $\hat{A}(z)$ is given by

$$\hat{A}(z) = 1 + \hat{a}_1 z + \dots + \hat{a}_p z^p \quad (2)$$

where $\hat{\theta}_x = (\hat{a}_1 \dots \hat{a}_p)^T$ are the estimated AR coefficients (found by LPC analysis, see [5]) and $\hat{\sigma}_x^2$ is the residual error variance. The estimated parameter vector $\hat{\theta}_x$ and $\hat{\sigma}_x^2$ are calculated from $\{x(k)\}$ as follows:

$$\begin{aligned} \hat{\theta}_x &= -\hat{\mathbf{R}}^{-1} \hat{\mathbf{r}} \\ \hat{\sigma}_x^2 &= \hat{r}_0 + \hat{\mathbf{r}}^T \hat{\theta}_x \end{aligned} \quad (3)$$

where

$$\hat{\mathbf{R}} = \begin{pmatrix} \hat{r}_0 & \dots & \hat{r}_{p-1} \\ \vdots & \ddots & \vdots \\ \hat{r}_{p-1} & \dots & \hat{r}_0 \end{pmatrix} \quad \mathbf{r} = \begin{pmatrix} \hat{r}_1 \\ \vdots \\ \hat{r}_p \end{pmatrix} \quad (4)$$

and, where

$$\hat{r}_k = \frac{1}{N} \sum_{t=1}^{N-k} x(t+k)x(t) \quad \hat{r}_{-k} = \hat{r}_k \quad k = 0, \dots, p \quad (5)$$

The set of linear equations (3) can be solved using the Levinson-Durbin algorithm, see [3]. The spectral estimate (1) is known to be smooth and its statistical properties have been analyzed in [6] for broad-band and noisy narrow-band signals, respectively.

In general, due to model errors there appears some bias in the spectral valleys. Roughly, this bias can be described as

$$\hat{\Phi}_x(\omega) - \Phi_x(\omega) \begin{cases} \approx 0 & \text{for } \omega \text{ such that } \Phi_x(\omega) \approx \max_{\omega} \Phi_x(\omega) \\ \gg 0 & \text{for } \omega \text{ such that } \Phi_x(\omega) \ll \max_{\omega} \Phi_x(\omega) \end{cases} \quad (6)$$

where $\hat{\Phi}_x(\omega)$ is the estimate (1) and $\Phi_x(\omega)$ is the true (and unknown) power spectral density of $x(k)$.

In order to reduce the bias appearing in the spectral valleys, the residual is calculated according to

$$\varepsilon(k) = \hat{A}(z^{-1})x(k) \quad k = 1, \dots, N \quad (7)$$

Performing another LPC analysis on $\{\varepsilon(k)\}$, the residual power spectral density can be calculated from. cf. (1)

$$\hat{\Phi}_\varepsilon(\omega) = \frac{\hat{\sigma}_\varepsilon^2}{|\hat{B}(e^{j\omega})|^2} \quad (8)$$

where, similarly to (2), $\hat{\theta}_\varepsilon = (\hat{b}_1 \dots \hat{b}_q)^T$ denotes the estimated AR coefficients and $\hat{\sigma}_\varepsilon^2$ the error variance. In general, the model order $q \neq p$, but here it seems reasonable to let $p = q$. Preferably $p \approx \sqrt{N}$, for example N may be chosen around 10.

In the proposed frequency domain algorithm below, the estimate (1) is compensated according to

$$\tilde{\Phi}_x(\omega) = \frac{\gamma}{\hat{\sigma}_x^2} \cdot \hat{\Phi}_\varepsilon(\omega) \cdot \hat{\Phi}_x(\omega) \quad (9)$$

where $\gamma (\approx 1)$ is a design variable. The frequency domain algorithm is summarized in the algorithms section below and in the block diagrams in Fig. 1 and 5.

A corresponding time domain algorithm is also summarized in the algorithms section and in Fig. 2 and 6. In this case the compensation is performed in a convolution step, in which the LPC filter coefficients $\hat{\theta}_x$ are compensated. This embodiment is more efficient, since one PSD estimation is replaced by a less complex convolution. In this embodiment the scaling factor γ may simply be set to a constant near or equal to 1. However, it is also possible to calculate γ for each frame, as in the frequency domain algorithm by calculating the root of the characteristic polynomial defined by $\hat{\theta}_\varepsilon$ that lies closest to the unit circle. If the angle of this root is denoted $\bar{\omega}$, then

$$\max_k \hat{\Phi}_\varepsilon(k) = \frac{\hat{\sigma}_\varepsilon^2}{|\hat{B}(e^{j\bar{\omega}})|^2}$$

ALGORITHMS

INPUTS

$\mathbf{x}$ input data $\mathbf{x} = (x(1) \dots x(N))^T$

p LPC model order

OUTPUTS

$\hat{\theta}_x$ signal LPC parameters $\hat{\theta}_x = (\hat{a}_1 \dots \hat{a}_p)^T$

$\hat{\sigma}_x^2$ signal LPC residual variance

$\hat{\Phi}_x$ signal LPC spectrum $\hat{\Phi}_x = (\hat{\Phi}_x(1) \dots \hat{\Phi}_x(N/2))^T$

$\tilde{\Phi}_x$ compensated LPC spectrum $\tilde{\Phi}_x = (\tilde{\Phi}_x(1) \dots \tilde{\Phi}_x(N/2))^T$

ϵ residual $\epsilon = (\epsilon(1) \dots \epsilon(N))^T$

$\hat{\theta}_\epsilon$ residual LPC parameters $\hat{\theta}_\epsilon = (\hat{b}_1 \dots \hat{b}_p)^T$

$\hat{\sigma}_\epsilon^2$ residual LPC error variance

γ design variable $(=1/(\max_k \hat{\Phi}_\epsilon(k)))$ in preferred embodiment)

FREQUENCY DOMAIN ALGORITHM

FOR EACH FRAME DO THE FOLLOWING STEPS:

(power spectral density estimation)

 $[\hat{\theta}_x, \hat{\sigma}_x^2] := \text{LPCanalyze}(x, p)$ signal LPC analysis $\hat{\Phi}_x := \text{SPEC}(\hat{\theta}_x, 1, N)$ signal spectral estimation, $\hat{\sigma}_x^2$ set to 1

(bias compensation)

 $\varepsilon := \text{FILTER}(\hat{\theta}_x, x)$ residual filtering $[\hat{\theta}_\varepsilon, \hat{\sigma}_\varepsilon^2] := \text{LPCanalyze}(\varepsilon, p)$ residual LPC analysis $\hat{\Phi}_\varepsilon := \text{SPEC}(\hat{\theta}_\varepsilon, \hat{\sigma}_\varepsilon^2, N)$ residual spectral estimation

FOR k=1 TO N/2 DO spectral compensation

 $\bar{\Phi}_x(k) := \gamma \cdot \hat{\Phi}_x(k) \cdot \hat{\Phi}_\varepsilon(k) \quad 1/(\max_k \hat{\Phi}_\varepsilon(k)) \leq \gamma \leq 1$

END FOR

TIME DOMAIN ALGORITHM

FOR EACH FRAME DO THE FOLLOWING STEPS:

 $[\hat{\theta}_x, \hat{\sigma}_x^2] := \text{LPCanalyze}(x, p)$ signal LPC analysis $\varepsilon := \text{FILTER}(\hat{\theta}_x, x)$ residual filtering $[\hat{\theta}_\varepsilon, \hat{\sigma}_\varepsilon^2] := \text{LPCanalyze}(\varepsilon, p)$ residual LPC analysis $\hat{\theta} := \text{CONV}(\hat{\theta}_x, \hat{\theta}_\varepsilon)$ LPC compensation $\hat{\Phi} := \text{SPEC}(\hat{\theta}, \hat{\sigma}_\varepsilon^2, N)$ spectral estimation

FOR k=1 TO N/2 DO

 $\bar{\Phi}_x(k) := \gamma \cdot \hat{\Phi}(k)$ scaling

END FOR



CLAIMS

1. A power spectral density estimation method, comprising the steps of:

5 performing a LPC analysis on an input signal vector for determining a first set of LPC filter parameters;

determining a first power spectral density estimate of said input signal vector based on said first set of LPC filter parameters;

10 filtering said input signal vector through an inverse LPC filter determined by said first set of LPC filter parameters for obtaining a residual signal vector;

performing a LPC analysis on said residual signal vector for determining a second set of LPC filter parameters;

15 determining a second power spectral density estimate of said residual signal vector based on said second set of LPC filter parameters; and

forming a bias compensated power spectral estimate of said input signal vector that is proportional to the product of said first and second power spectral estimates.

20 2. The method of claim 1, wherein said product is multiplied by a positive scaling factor that is less than or equal to 1.

3. The method of claim 2, wherein said scaling factor is the inverted value of the maximum value of said second power spectral density estimate.

25 4. The method of claim 1, 2 or 3, wherein said input signal vector comprises speech samples.

5. A power spectral density estimation method, comprising the steps of:

30 performing a LPC analysis on an input signal vector for determining a first set of LPC filter parameters;

filtering said input signal vector through an inverse LPC filter determined by said first set of LPC filter parameters for



obtaining a residual signal vector;

performing a LPC analysis on said residual signal vector for determining a second set of LPC filter parameters;

convolving said first set of LPC filter parameters with said
5 second set of LPC filter parameters for forming a compensated set of LPC filter parameters;

determining a bias compensated power spectral density estimate of said input signal vector based on said compensated set of LPC filter parameters.

10 6. The method of claim 5, wherein said bias compensated power spectral density estimate is multiplied by a positive scaling factor that is less than or equal to 1.

7. The method of claim 6, wherein said scaling factor is the
15 inverted value of the maximum value of a power spectral density estimate of said residual signal vector.

8. The method of claim 5, 6 or 7, wherein said input signal vector comprises speech samples.

9. A power spectral density estimation apparatus, comprising:

20 means (10) for performing a LPC analysis on an input signal vector for determining a first set of LPC parameters;

means (12) for determining a first power spectral density estimate of said input signal vector based on said first set of LPC parameters;

25 means (14) for filtering said input signal vector through an inverse LPC filter determined by said first set of LPC parameters for obtaining a residual signal vector;

means (16) for performing a LPC analysis on said residual signal vector for determining a second set of LPC parameters;

30 means (18) for determining a second power spectral density estimate of said residual signal vector based on said second set of LPC parameters; and

means (20) for forming a bias compensated power spectral estimate of said input signal vector that is proportional to the



product of said first and second power spectral estimates.

10. A power spectral density estimation apparatus, comprising:
means (10) for performing a LPC analysis on an input signal
vector for determining a first set of LPC filter parameters;

5 means (14) for filtering said input signal vector through an
inverse LPC filter determined by said first set of LPC filter
parameters for obtaining a residual signal vector;

10 means (16) for performing a LPC analysis on said residual
signal vector for determining a second set of LPC filter
parameters;

means (22) for convolving said first set of LPC filter
parameters with said second set of LPC filter parameters for
forming a compensated set of LPC filter parameters;

15 means (12') for determining a bias compensated power spectral
density estimate of said input signal vector based on said
compensated set of LPC filter parameters.