A Bayesian-Adaptive Decision Method for the V/UV/S Classification of Segments of a Speech Signal

G. BRUNO, M. D. DI BENEDETTO, M. G. DI BENEDETTO, A. GILIO, AND P. MANDARINI

Abstract—**In this correspondence, a method for voiced (V), unvoiced (UV), or silence (S) classification of speech segments, based on the maximum *a posteriori* probability criterion, is presented. The *a posteriori* probabilities of the three classes are determined using a vector $x = (f_1, \cdots, f_L)$ of measurements on the segment under consideration. It is assumed that the vector x has an L-dimensional Gaussian distribution with an expected random value also characterized by an L-dimensional Gaussian distribution. In addition, it is assumed that the sequence of the classes constitutes a first-order stationary Markov chain.**

The initial parameters are estimated in a training phase. During the application phase, the decision method is adapted by using the previous classifications in order to update the probability density function (pdf) of the expected random values.

I. Introduction

In the past, various methods have been proposed for the automatic V/UV/S classification: some are based on a probabilistic approach, while others are of the deterministic type. Among the probabilistic methods, we recall in particular the one proposed by Atal and Rabiner [1] in which a minimum distance classifier, based on the maximum likelihood criterion, is used.

The method proposed in this correspondence is probabilistic and is based on a Bayesian approach. The maximum *a posteriori* probability criterion is used. In addition, the method presented is adaptive: some pdf's, estimated in a training phase, are updated taking into account the previous decision.

At each step, the classification is based on a vector of measurements $\boldsymbol{x}$ on the segment under examination. It is assumed that the distribution of $\boldsymbol{x}$ is an L-dimensional Gaussian, with an expected value $\boldsymbol{m}$ characterized by a similar distribution. In addition, it is supposed that the sequence of classes constitutes a first-order Markov chain. In order to reduce the risks associated with the adaptive characteristic of the method, the updating is carried out only if the probability of the class chosen is greater than a threshold.

Notation

$\boldsymbol{x}_n$	Vector of measurements on the nth speech segment.
$\boldsymbol{y}_n$	$\boldsymbol{y}_n = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n) = (\boldsymbol{y}_{n-1}, \boldsymbol{x}_n)$.
$\boldsymbol{X}_n$	Random variable. $\boldsymbol{x}_n$ is a determination of $\boldsymbol{X}_n$.
$g(\boldsymbol{x}; \boldsymbol{m}, \boldsymbol{R})$	Multidimensional Gaussian pdf with vector of expected values $\boldsymbol{m}$ and covariance matrix $\boldsymbol{R}$.
$\{C_0, C_1, C_2\}$	Set of classes (V/UV/S, respectively).
s_n	Class to which the nth segment belongs.
S_n	Unknown class to which the nth segment belongs; $S_n \in \{C_0, C_1, C_2\}$.
$\boldsymbol{s}_n$	$\boldsymbol{s}_n = (s_1, \cdots, s_n) = (\boldsymbol{s}_{n-1}, s_n)$.
s_n^*	Classification of the nth segment.
$p_n(\boldsymbol{m}/C_j)$	pdf of $\boldsymbol{m}$ at step n, under the hypothesis that the nth segment belongs to C_j. Note that $p_n(\boldsymbol{m}/C_j) = p(\boldsymbol{m}/s_n = C_j)$.
$p_n(\boldsymbol{x}_n/C_j)$	pdf of $\boldsymbol{X}_n$ at the step n, under the hypothesis that the nth segment belongs to C_j. Note that $p_n(\boldsymbol{x}_n/C_j) = p(\boldsymbol{x}_n/s_n = C_j, \boldsymbol{y}_{n-1})$.
$\boldsymbol{m}_j^{(n)}, \boldsymbol{Q}_j^{(n)}$	Vector of expected values and covariance matrix of $p_n(\boldsymbol{x}_n/C_j)$.
$\boldsymbol{R}_j$	Covariance matrix of $\boldsymbol{X}$ when the segment belongs to the class C_j and the expected value of $\boldsymbol{X}$ is known.
p_{ij}	Transition probability between C_i and C_j. $p_{ij} = \Pr\{s_{n+1} = C_j/s_n = C_i\}$.
p_j	*A priori* probability of the class C_j, $j = 0, 1, 2$.

Manuscript received April 19, 1982; revised September 4, 1986.

The authors are with the Dipartimento di Scienza e Tecnica dell'Informazione e della Comunicazione, Facoltá di Ingegneria, Universitá di Roma "la Sapienza," Rome, Italy.

IEEE Log Number 8612867.

II. Descriptive Model and Algorithm

In the model proposed in this correspondence, it is assumed that the speaker under consideration is initially chosen by a probabilistic mechanism from the set of all possible speakers. The speech signal produced by a specific speaker is supposed to be obtained as follows: a first probabilistic source generates the sequence of classes $\{s_n\}$ and a second probabilistic mechanism, for each class, generates the phoneme, and therefore a particular waveform.

Each segment is classified on the basis of a vector of measurements. The classification algorithm is adaptive since, while it proceeds with the classifications, it adapts some parameters of the distributions to the particular speaker, taking into account the previous classifications. Therefore, the distributions depend upon the index n and must be interpreted as being conditioned to the sequence $\boldsymbol{y}_{n-1}$ if considered before the measurement of $\boldsymbol{X}_n$, and to the sequence $\boldsymbol{y}_n$ after the measurement of $\boldsymbol{X}_n$. The initial distributions ($n = 1$) are estimated during a training phase.

For the sake of simplicity (as in [1]), it is assumed that $\boldsymbol{X}_k$ has an L-dimensional Gaussian pdf $g(\boldsymbol{x}; \boldsymbol{m}, \boldsymbol{R})$. In addition, in the present method, the dependency of the expected value $\boldsymbol{m}$ and the covariance matrix $\boldsymbol{R}$ upon the speaker and s_k is taken into account, and $\boldsymbol{m}$ and $\boldsymbol{R}$ are considered as random variables. Let $\boldsymbol{m}_j^{(n)}$ be the vector of the expected values and $\boldsymbol{Q}_j^{(n)}$ the covariance matrix of $p_n(\boldsymbol{m}/C_j)$; it is assumed that

$$p_1(\boldsymbol{m}/C_j) = g(\boldsymbol{m}; \boldsymbol{m}_j^{(1)}, \boldsymbol{Q}_j^{(1)}) \qquad j = 0, 1, 2 \tag{1}$$

where the parameters $\boldsymbol{m}_j^{(1)}$ and $\boldsymbol{Q}_j^{(1)}$ are estimated in the training phase. It is supposed that the covariance matrix $\boldsymbol{R}$ is identical for all speakers and is equal to $\boldsymbol{R}_j$ for class C_j. Therefore, it is assumed that

$$p(\boldsymbol{x}/s_n = C_j, \boldsymbol{m}) = g(\boldsymbol{x}; \boldsymbol{m}, \boldsymbol{R}_j) \qquad j = 0, 1, 2. \tag{2}$$

From (2) and (1), one obtains

$$\begin{aligned} p_1(\boldsymbol{x}/C_j) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_1(\boldsymbol{x}, \boldsymbol{m}/s_1 = C_j) \cdot d\boldsymbol{m} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\boldsymbol{x}/s_n = C_j, \boldsymbol{m}) \cdot p_1(\boldsymbol{m}/C_j) \cdot d\boldsymbol{m} \\ &= g(\boldsymbol{x}; \boldsymbol{m}_j^{(1)}, \boldsymbol{R}_j + \boldsymbol{Q}_j^{(1)}) \qquad j = 0, 1, 2. \end{aligned} \tag{3}$$

Finally, let $P_n(C_j)$ and $P_n(C_j/\boldsymbol{x}_n)$ be the probabilities of the events $\{s_n = C_j\}$ and $\{s_n = C_j/\boldsymbol{X}_n = \boldsymbol{x}_n\}$, respectively; due to the presence of index n, these probabilities are conditioned to $\boldsymbol{y}_{n-1}$. More precisely, one has

$$\begin{aligned} P_n(C_j) &= \Pr\{s_n = C_j/\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{n-1}\} \\ &= \Pr\{s_n = C_j/\boldsymbol{y}_{n-1}\} \\ P(C_j/\boldsymbol{x}_n) &= \Pr\{s_n = C_j/\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\} = \Pr\{s_n = C_j/\boldsymbol{y}_n\} \\ &= \Pr\{s_n = C_j/\boldsymbol{y}_{n-1}, \boldsymbol{x}_n\}. \end{aligned} \tag{4}$$

The classification method proposed is based on the criterion of the maximum *a posteriori* probability. At step n, the class C_k is chosen if

$$P_n(C_k/x_n) = \max_j P_n(C_j/x_n) \quad (5)$$

where

$$P_n(C_j/x_n) = P_n(C_j) \cdot p_n(x_n/C_j)/p(x_n)$$

and

$$p(x_n) = \sum_{h=0}^{2} P_n(C_h) \cdot p_n(x_n/C_h). \quad (6)$$

The classification at step $n + 1$ requires the updating of $P_{n+1}(C_j)$ and $p_{n+1}(x_{n+1}/C_j)$. It is assumed that $\Pr\{s_{n+1} = C_j/s_n, y_n\} = \Pr\{s_{n+1} = C_j/s_n\}$. This corresponds to: suppose that there is some kind of independency of the probabilistic structure of the language from the mechanism of production of the speech signal. In addition, it is supposed that the sequence of classes $\{S_n\}$ is a first-order Markov chain; then $\Pr\{s_{n+1} = C_j/s_n\} = \Pr\{s_{n+1} = C_j/s_n\}$. One can write

$$\begin{aligned} P_{n+1}(C_j) &= \Pr\{s_{n+1} = C_j/y_n\} = \sum_{s_n} \Pr\{s_{n+1} = C_j, s_n/y_n\} \\ &= \sum_{s_n} \Pr\{s_{n+1} = C_j/s_n, y_n\} \cdot \Pr\{s_n/y_n\} \\ &= \sum_{i=0}^{2} \Pr\{s_{n+1} = C_j/s_n = C_i\} \cdot \Pr\{s_n = C_i/y_n\} \\ &= \sum_{i=0}^{2} p_{ij} \cdot P_n(C_i/x_n). \end{aligned} \quad (7)$$

Note that the probabilities p_{ij} are not updated and that the quantities p_{ij}, P_j, and R_j are all estimated in the training phase. As regards $p_{n+1}(x_{n+1}/C_j)$, it is assumed that if m and s_{n+1} are known, X_{n+1} is stochastically independent of $X_1, \cdots, X_n$. Therefore, from (2), one obtains $p_{n+1}(x_{n+1}/C_j, m) = p(x_{n+1}/C_j, m) = g(x_{n+1}; m, R_j)$. Then

$$\begin{aligned} p_{n+1}(x_{n+1}/C_j) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{n+1}(x_{n+1}/C_j, m) \\ &\quad \cdot p_{n+1}(m/C_j, y_n) \cdot dm \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_{n+1}; m, R_j) \\ &\quad \cdot p_{n+1}(m/C_j, y_n) \cdot dm. \end{aligned} \quad (8)$$

The real calculation of (8) requires the determination of $p_{n+1}(m/s_{n+1} = C_j, y_n)$. The pdf $p_{n+1}(m/C_j)$ should be calculated as follows:

$$\begin{aligned} p_{n+1}(m/C_j, y_n) &= p_n(m/C_j, y_{n-1}) \cdot \Pr\{s_n \neq C_j/y_n\} \\ &\quad + p_n(m/C_j, x_n) \cdot \Pr\{s_n = C_j/y_n\}. \end{aligned} \quad (9)$$

Equation (9) makes the calculations considerably complex and will not be used. In this correspondence, the following criterion is proposed:

$$p_{n+1}(m/C_j) = \begin{cases} p_n(m/C_j, x_n) & \text{if } s_n^* = C_j \\ p_n(m/C_j) & \text{if } s_n^* \neq C_j. \end{cases} \quad (10)$$

This criterion is justified by the fact that the pdf to be updated corresponds to the class $C_j = s_n^*$ which has the maximum probability to be the true one. In order to reduce the risks related to the above criterion, the updating will be made only if

$$\max_j \{\Pr\{s_n = C_j/y_n\}\} \geq \underline{a} \quad (11)$$

where the threshold $\underline{a}$ is appropriately chosen.

As an application of (11), suppose, for example, that $s_n^* = C_k$ and $\Pr\{s_n = C_k/y_n\} \geq \underline{a}$. One then has

$$\begin{aligned} p_{n+1}(m/C_k) &= p_n(m/C_k, x_n) \\ &= p_n(x_n/C_k, m) \cdot p_n(m/C_k)/p_n(x_n/C_k) \end{aligned} \quad (12)$$

where $p_n(x_n/C_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_n(x_n/C_k, m) \cdot p_n(m/C_k) \cdot dm$.

As is well known, it is possible to prove that $p_{n+1}(m/C_k)$ is an L-dimensional Gaussian distribution with parameters $m_k^{(n+1)}$ and $Q_k^{(n+1)}$ and one has [2]

$$\begin{aligned} m_k^{(n+1)} &= R_k \cdot [Q_k^{(n)} + R_k]^{-1} \cdot m_k^{(n)} + Q_k^{(n)} \cdot [Q_k^{(n)} + R_k]^{-1} \cdot x_n \\ Q_k^{(n+1)} &= R_k \cdot [Q_k^{(n)} + R_k]^{-1} \cdot Q_k^{(n)} = Q_k^{(n)} \cdot [Q_k^{(n)} + R_k]^{-1} \cdot R_k. \end{aligned} \quad (13)$$

In general, one can write, for any j, referring to $n = 1$, the following updating relations:

$$\begin{aligned} m_j^{(n+1)} &= \Big\{R_j \cdot [Q_j^{(1)} + R_j/n_j]^{-1} \cdot m_j^{(1)} + Q_j^{(1)} \\ &\quad \cdot [Q_j^{(1)} + R_j/n_j]^{-1} \cdot \sum_{i \in N_j} x_i\Big\}\Big/ n_j \\ Q_j^{(n+1)} &= (R_j/n_j) \cdot [Q_j^{(1)} + R_j/n_j]^{-1} \cdot Q_j^{(1)} = Q_j^{(1)} \\ &\quad \cdot [Q_j^{(1)} + R_j/n_j]^{-1} \cdot R_j/n_j \end{aligned} \quad (14)$$

where $N = \{1, 2, \cdots, n\}$, $N_j = \{h \in N: s_h^* = C_j$ and $\Pr\{s_h = C_j/y_h\} \geq \underline{a}\}$, $n_j = \text{Card}(N_j)$, and for $n \to \infty$: $m_j^{(n)} \to (1/n_j) \equiv_i x_i$ and $Q_j^{(n)} \to 0$.

Finally, one has

$$p_{n+1}(x_{n+1}/C_j) = g(x_{n+1}; m_j^{(n+1)}, R_j + Q_j^{(n+1)}) \qquad j = 0, 1, 2. \quad (15)$$

In conclusion, using (7), (14), and (15), one can carry out, at each step, the updating necessary for the classification at the following step.

III. Computation Procedure

Some developments for the reduction of the number of operations required by the algorithm are presented. An extensive description of the procedures used can be found in [3].

As regards $[R_j + Q_j^{(n+1)}]^{-1}$, it is possible to show that

$$\begin{aligned} [R_j + Q_j^{(n+1)}]^{-1} &= [I + (G_j + n_j \cdot I)^{-1}]^{-1} \cdot R_j^{-1} \\ &= R_j^{-1} - [G_j + (n_j + 1) \cdot I]^{-1} R_j^{-1} \end{aligned} \quad (16)$$

where $G_j = [Q_j^{(1)}]^{-1} \cdot R_j$.

If the eigenvalues of the matrix G_j are indicated by a_{ij}, $i = 1, \cdots, L$, and Z_{ij} are the constituent matrices of G_j, one can write

$$[R_j + Q_j^{(n+1)}] = R_j^{-1} - \sum_{i=1}^{L} (a_{ij} + n_j + 1)^{-1} \cdot W_{ij} \quad (17)$$

where $W_{ij} = Z_{ij} \cdot R_j^{-1}$.

Note from (17) that R_j^{-1} and W_{ij} are computed only once, in the training phase, while the coefficients $(a_{ij} + n_j + 1)^{-1}$ must be updated.

As regards $m_j^{(n+1)}$, it is possible to show that

$$m_j^{(n+1)} = \sum_{i=1}^{L} (a_{ij} + n_j)^{-1} \cdot [v_{ij} + k_{ij} + s_j^{(n+1)}] \quad (18)$$

where $v_{ij} = R_j \cdot Z_{ij} \cdot [Q_j^{(1)}]^{-1} \cdot m_j^{(1)}$, $k_{ij} = R_j \cdot Z_{ij} \cdot R_j^{-1}$, and $s_j^{(n+1)} = \Sigma_{i \in N_j} x_i$. As one can see, the quantities v_{ij} and k_{ij} are calculated only once, in the training phase, while $(a_{ij} + n_j)^{-1}$ and $s_j^{(n+1)}$ must be updated where

$$s_j^{(n+1)} = \begin{cases} s_j^{(n)} + y_n & \text{if } s_n^* = C_j \text{ and } \Pr\{s_n = s_n^*/y_n\} \geq a \\ s_j^{(n)} & \text{otherwise.} \end{cases}$$

TABLE I
CLASSIFICATION PARAMETERS AFTER TRAINING

	Voiced ZC	Voiced LE	Voiced $A1$		Unvoiced ZC	Unvoiced LE	Unvoiced $A1$		Silence ZC	Silence LE	Silence $A1$
Mean	13.40	53.80	−1.45	Mean	63.80	28.70	0.21	Mean	24.20	23.60	−0.78
Var.	6.42	4.49	0.43	Var.	22.90	5.76	0.76	Var.	15.30	7.05	0.32
Norm.	1.00	0.14	0.12	Norm.	1.00	−0.12	0.61	Norm.	1.00	−0.13	0.76
Covar.	0.14	1.00	−0.32	Covar.	0.12	1.00	−0.36	Covar.	−0.13	1.00	−0.38
Matrix	0.12	−0.32	1.00	Matrix	0.61	−0.36	1.00	Matrix	0.76	−0.38	1.00

TABLE II
CLASSIFICATION PARAMETERS AFTER A TEST SENTENCE (6 s DURATION)

	Voiced ZC	Voiced LE	Voiced $A1$		Unvoiced ZC	Unvoiced LE	Unvoiced $A1$		Silence ZC	Silence LE	Silence $A1$
Mean	11.60	50.20	−1.53	Mean	62.80	30.40	−0.16	Mean	31.40	23.90	−0.68
Var.	6.36	4.14	0.43	Var.	22.80	5.13	0.50	Var.	12.80	7.05	0.30
Norm.	1.00	0.10	0.11	Norm.	1.00	−0.10	0.86	Norm.	1.00	−0.18	0.72
Covar.	0.10	1.00	−0.38	Covar.	−0.10	1.00	−0.02	Covar.	−0.18	1.00	−0.42
Matrix	0.11	−0.38	1.00	Matrix	0.86	−0.02	1.00	Matrix	0.72	−0.42	1.00

TABLE III
RESULTS OBTAINED WITH FOUR PARAMETERS (PERCENT) TOTAL NUMBER OF SEGMENTS = 939 (V = 678, UV = 74, S = 187)

In \ Out	Method I V	Method I UV	Method I S	In \ Out	Method II V	Method II UV	Method II S
V	98.50	0.75	0.75	V	95.40	1.90	2.70
UV	1.30	87.80	10.90	UV	1.35	85.15	13.50
S	0.00	5.35	94.65	S	0.00	3.20	96.80

As regards $|R_j + Q_j^{(n+1)}|^{-1/2}$, it is possible to show that

$$|R_j + Q_j^{(n+1)}|^{-1/2} \cong \left\{ |R_j| \cdot \left(1 + \sum_{i=1}^{L} \frac{1}{a_{ij} + n_j}\right)\right\}^{-1/2}. \quad (19)$$

IV. EXPERIMENTAL RESULTS

In this section, results concerning the experimental application of the classification method are presented. The results obtained are more extensively described in [4].

The classification method was experimented on an IBM Series/1 computer. A dynamic microphone was used. The speech signal was filtered at 4.8 kHz and sampled at 10 kHz. It was subdivided into segments of 12.8 ms duration corresponding to $N = 128$ samples. The following measurements were considered:

1) log energy of the signal LE,
2) zero crossing ZE,
3) first linear prediction coefficient $A1$,
4) normalized autocorrelation coefficient between adjacent samples C.

During the training phase, the vector of expected values of x and the associated covariance matrix were estimated for each class on the basis of 1440 speech segments spoken by one female speaker and one male speaker. The *a priori* probabilities of the three classes and the transition probabilities were evaluated on the basis of an analysis of 4936 segments. All of the segments of the training data were classified manually by visual examination of spectrograms and waveforms.

The algorithm was tested on four sentences (939 segments): two were pronounced by a female speaker (F) and the remainder by a male speaker (M). Both speakers (F and M) did not participate in the training. The algorithm proposed by Atal and Rabiner [1] was tested using the same speech materials.

Some experiments were also carried out using only three parameters (excluding C) on three sentences (1419 segments) pronounced by M and F. Table I shows the mean values, the variances, and the covariance matrices $P_j + Q_j$ (normalized) of the three parameters at the end of the training phase, and Table II shows their values at the end of a test sentence of 6 s duration, pronounced by M.

The results obtained using four and three parameters are shown in Tables III and IV, respectively. The misclassification rates are the following (method I denotes the method proposed; method II denotes Atal–Rabiner's method).

		Method I	Method II
With Four Features:	on M	2.15	4.57
	on F	3.95	5.62
	Globally:	3.08	5.1
With Three Features:	on M	2.87	5.1
	on F	5.62	6.87
	Globally	3.8	5.7

These results show that the adaptation is helpful: the global misclassification rate is 3.08 percent compared to 5.1 percent obtained with the nonadaptive method. Moreover, even when using only three parameters, the global misclassification rate is 3.8 percent, always lower than that obtained with method II and with one more parameter. However, the computation is longer in method I. The amount of computation required is discussed in the next paragraph.

TABLE IV
RESULTS OBTAINED WITH THREE PARAMETERS (PERCENT) TOTAL NUMBER OF SEGMENTS = 1419 (V = 1049, UV = 121, S = 249)

In \ Out	Method I V	Method I UV	Method I S	In \ Out	Method II V	Method II UV	Method II S
V	97.70	1.15	1.15	V	95.20	2.00	2.80
UV	0.00	85.90	14.10	UV	0.90	83.40	15.70
S	0.00	5.20	94.80	S	0.50	4.00	95.50

With the method proposed, most of the classification errors occur in the transition V/UV or UV/V, and the examination of the sequence of decisions has shown the uselessness of smoothing after decision. This procedure, however, is useful in the nonadaptive method. Those results may be compared to that obtained with a similar method in which only the *a priori* probabilities of the classes are updated [5].

The algorithm did not show any significant sensitivity to the value of $\underline{a}$ in the range of 0.7 to 0.95. The results reported correspond to this range of values. For $0.7 < \underline{a} < 0.95$, the adaptation algorithm converges to a stable state after about 1.5 s. Clearly, if $\underline{a}$ is decreased, the convergency to a stable state is quicker but less precise, and the error rate increases.

V. Amount of Computation

It is of interest to compare the number of computations (multiplications) required for the application of method I and method II.

It is not difficult to see that the number of computations for method II is

$$C_{\text{II}} = (2L^2) \cdot 3 = 6L^2.$$

In method I, the number of computations for the calculation of $p_n(x_n / C_j)$ is

$$C_{\text{I}}' = (2L^2 + 3) \cdot 3 = 6L^2 + 9.$$

In addition, the number of computations needed for obtaining the *a priori* probabilities of the classes is

$$C_{\text{I}}'' = 3 \cdot 3 = 9.$$

Finally, the number of computations needed for the updating of the parameters of the pdf is

$$C_{\text{I}}''' = L^3 + L^2 + L.$$

The total number of computations needed for applying method I is then

$$C_{\text{I}} = C_{\text{I}}' + C_{\text{I}}'' + C_{\text{I}}''' = L^3 + 7L^2 + L + 18.$$

As one can see, C_{I} is a polynomial in L of the third order, while C_{II} is of the second order. It is possible to reduce C_{I} by appropriately choosing $p_n(x_n / C_j)$ as shown in [6].

VI. Conclusions

In this correspondence, a new method for V/UV/S classification of segments of a speech signal is presented. This method uses a probabilistic pattern recognition approach based upon a Bayesian decision criterion. The method presented introduces two new elements: the hypothesis that the sequence of classes constitutes a first-order Markov chain, and adaptation to the specific speaker of some statistical parameters used for the classification. These two characteristics allow us to obtain a classification procedure which has proved to be promising.

It is appropriate to highlight the fact that the method proposed does not require a final smoothing process of the sequence of classifications: this is due to the use of a Markovian model for the sequence of the classes.

References

[1] B. S. Atal and L. R. Rabiner, "A pattern recognition approach to V/UV/S classification with applications to speech recognition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, June 1976.

[2] T. Y. Young and T. W. Calvert, *Classification, Estimation and Pattern Recognition*. New York: American Elsevier, 1978.

[3] G. Bruno, A. Gilio, and P. Mandarini, "Un metodo di decisione Bayesiano-adattativo per la classificazione V-UV-S di segmenti di un messaggio vocale," Istituto di Comunicazioni Elettriche, Universitá degli Studi di Roma "La Sapienza," Rome, Italy, Internal Rep. 35, Dec. 1980.

[4] G. Bruno, M. D. Di Benedetto, M. G. Di Benedetto, A Gilio, and P. Mandarini, "New results of a Bayesian-adaptive V/UV/S classifier for speech signals," Istituto di Comunicazioni Elettriche, Universitá degli Studi di Roma "La Sapienza," Rome, Italy, Internal Rep. 47, Apr. 1982.

[5] G. Bruno, M. D. Di Benedetto, M. G. Di Benedetto, and A. Gilio, "A Bayesian approach for V/UV classification of segments of a speech signal," in *Proc. 4th FASE Symp. Acoust., Speech*, Venice, Italy, Apr. 1981.

[6] M. D. Di Benedetto, M. G. Di Benedetto, and P. Mandarini, "A new hypothesis on the adaptive approach on V/UV/S classification," Istituto di Comunicazioni Elettriche, Universitá degli Studi di Roma "La Sapienza," Rome, Italy, Internal Rep. 48, 1982.

On a Fundamental Property of the Wigner Distribution

LEON COHEN

***Abstract*—We discuss and illustrate the following fundamental property of the Wigner distribution: the Wigner distribution is not necessarily zero when the signal is zero.**

I. Introduction

The Wigner [1] distribution has been studied and applied with considerable profit in many areas of signal analysis [2]. However, it appears that one of its fundamental properties and consequences is not fully appreciated or realized: the Wigner distribution is not necessarily zero when the signal is zero, and similarly, it is not necessarily zero for those frequencies for which the spectrum is zero.

Although some manifestations of the above property have been previously noted, it has been my experience over the last 20 years that even active workers in the field are surprised at this property of the Wigner distribution. Hence, I thought it would be of value to explicitly discuss this aspect and some of its consequences. Re-

Manuscript received February 4, 1986; revised July 16, 1986.

The author is with Hunter College, New York, NY 10021 on leave at the IBM T. J. Watson Research Center, Yorktown Heights, NY 10598.

IEEE Log Number 8612868.

0096-3518/87/0400-0559$01.00 © 1987 IEEE