Kullback 情報量による非定常時系列の分類

石井直宏・岩田彰・鈴村宣夫
情報工学科
(1978年 9月 9日 受理)

Classification of the Nonstationary Time Series by Kullback's Information Measure

Naohiro Ishii, Akira Iwata and Nobuo Suzumura
Department of Information Engineering
(Received September 9, 1978)

Kullback information plays an important role as a measure of measuring the difference between two probability densities. In this paper, we apply Kullback information in the frequency domain to evaluate the difference between two spectral densities. It is proved here that Kullback information measure is equivalent to other measure under some conditions such as likelihood ratio of residuals of autoregressive model or spectral error measure. Using Kullback information as a measure of the segmentation and the classification of the nonstationary EEG data, we classify the sleep stages of human subject during all night. To recognize the sleep stages which consist of 6 stages in the human subject, first we segment the nonstationary data to make stationary subsequences, and next extract the features from the segmented sequences. Then we classify the EEG data by computing the distance between the features of the segmented sequence and the typical EEG pattern. These segmentation and classification are based on the calculation of Kullback information. The results of the sleep stages obtained by computer comparatively agree with those by a physician. This shows that Kullback information is useful to the analysis of the nonstationary time series. Finally, we compare as a measure Kullback Information with other measures, such as Divergence, Bhattacharyya Distance.

1. まえがき

非定常な時系列から意味のある情報を抽出すること，あるいは非定常な時系列を作り出す系の状態を推定すること，およびこれらの系の状態を予測することなどの問題が多くの分野で見い出される。この種の非定常な時系列の取り扱い方として，定常な小区間に分割し，その上で各小区間の特徴パラメータを抽出する方法や[1]はじめに定常な時系列モデルへ適合させ，順次モデルのパラメータを適応的に変化させる方法[2]などがある。

本論文は非定常時系列の具体的なものとして，睡眠脳波時系列を取り上げ，非定常時系列の分類という問題をパタン識別の立場から検討した。時系列のパタン識別において，標準パタンと未知パタンの比較ができるように非定常時系列を定常小区間に区分化することが必要である。定常小区間に区分化することは，隣り合う小区間の境界をつける問題である。この問題は隣り合う小区間の確率分布の差違を検出する問題に等しい。Kullback 情報量は広くシステムの次数の推定[3]や非線形システムの同定[4]に適用されている。さらに Kullback 情報量はマルコフ過程における変数変換によって，正規過程と非正規過程を結びつける不変な量[5]として重要である。

本論文は非定常時系列の定常時系列への区分化と区分

化された時系列のパタンの分類のため，Kullback情報量を導入した。ここでは Kullback 情報量がスペクトル情報の推定のため展開された最尤法による適合誤差尺度[6] さらには最尤法による予測残差の尤度比[7] と等価となることを明らかにした。また Kullback 情報量は，自己回帰モデルの予測残差から算出されたスペクトル誤差量[8] と等価となることを示した。次に非定常時系列として，本論文の睡眠脳波時系列の定常小区間の区分化のために Kullback 情報量と等価な最尤法による残差の尤度比を採用した。このとき，睡眠段階に特有な脳波を標準パタンとして決め，このパタンと区分化された脳波のパタンの間で残差の尤度比を計算し，睡眠段階の分類を行なった。最後に Kullback 情報量と Bhattacharyya Distance について尺度としての比較を行なった。

2. 自己回帰モデルとパワ・スペクトル

はじめに自己回帰モデルを導入する。いま時刻を表わす整数の集合を $\mathrm{I\!I}=\{\cdots-1,0,1,2\cdots\}$ とし，定常確率過程 $\{X_t, t\in \mathrm{I\!I}\}$ の仮定を一般性を失うことなく，

$$E(X_t)=0$$

とおく。このとき m 次の自己回帰過程は式(1)のように表現される。

$$X_t+\alpha_1X_{t-1}+\alpha_2X_{t-2}+\cdots\cdots+\alpha_mX_{t-m}=\xi_t \tag{1}$$

$\{X_t\}$ が定常正規過程であれば，式(1)において $\{\xi_t\}$ が独立な定常正規過程となる。ここで $\{\alpha_i\}$ は方程式(1)の係数を表わす。

脳波時系列などのように準定常な時系列，すなわち短時間区間内では定常とみなされる時系列から，特徴パラメータを抽出する必要が生じる。たとえば，この短時間区間内の時系列へ自己回帰モデルを適合させ，特徴となるパラメータを抽出することになる。以下に自己回帰モデルとパワ・スペクトル密度の関係を示す。時刻 t = n の予測誤差，あるいは予測残差を e_n とすると

$$e_n=x_n-\hat{x}_n \tag{2}$$

ここで $\hat{x}_n$ は式(1)で $x_{n-1}, x_{n-2},\cdots\cdots x_{n-m}$ を代入した時の x_n の値，および x_n は観測した時系列の値である式(1)を式(2)へ代入し，Z 変換すると

$$E(Z)=(1+\sum_{k=1}^{m}\alpha_kZ^{-k})\,X(Z)\equiv A(Z)X(Z) \tag{3}$$

を得る。ここで $E(Z)$ は e_n の Z 変換を表わし，式(1)の ξ_n は e_n に含まれるものとする。$\{X_t\}$ のパワスペクトルを $P(\omega)$ で表わすと

$$P(\omega)=|X(e^{j\omega})|^2 \tag{4}$$

で与えられるでここで $X(e^{j\omega})$ は，$X(Z)$ に $Z=e^{j\omega}$ を代入して得られる。

3. Kullback 情報量によるスペクトル密度の差違の評価

Kullback 情報量は 2 つの密度関数の差違を測る尺度として用いられて来た。具体的には，自己回帰モデルの次数の推定の基礎となり，またパタン認識における特徴抽出のための尺度として適用された。

Kullback 情報量 $I(1:2)$ は次式で定義される[9]

$$I(1:2)=\int f_1(x_1\cdots x_k)\,\log\frac{f_1(x_1\cdots x_k)}{f_2(x_1\cdots x_k)}\,dx_1dx_2\cdots dx_k \tag{5}$$

ここで $f_1(x_1\cdots x_k)$ と $f_2(x_1\cdots x_k)$ は各々，クラス 1 とクラス 2 の確率密度関数である。式(5)の左辺の記号 $I(1:2)$ は $\log\{f_1(x_1\cdots x_k)/f_2(x_1\cdots x_k)\}$ の $f_1(x_1\cdots x_k)$ による平均を意味している。ここで，$f_1(x_1\cdots x_k)$ と $f_2(x_1\cdots x_k)$ の各々が正規密度関数 $N(\mu_1, \Sigma_1)$ と $N(\mu_2, \Sigma_2)$ とすると式(5)の Kullback 情報量は式(6)で与えられる。

$$I(1:2)=\frac{1}{2}\log\frac{|\Sigma_2|}{|\Sigma_1|}+\frac{1}{2}t_r\Sigma_1(\Sigma_2^{-1}-\Sigma_1^{-1})+\frac{1}{2}t_r\Sigma_2^{-1}(\mu_1-\mu_2)(\mu_1-\mu_2)^t \tag{6}$$

ここで μ_1, μ_2 は各々平均値を表わし，Σ_1, Σ_2 は各々共分散行列を表わす。また Σ_i^{-1} は Σ_i の逆行列を表わし，t_r は行列のトレースである。

Kullback 情報量はデータ長 T の関数として式(7)のように定義できる[10]。

$$I(1:2)=\frac{1}{2}T^{-1}\left[\log\frac{|\Sigma_2|}{|\Sigma_1|}+t_r\Sigma_1\Sigma_2^{-1}-T\right]+\frac{1}{2}T^{-1}\delta^t\Sigma_2^{-1}\delta \tag{7}$$

ここで

$$\mu_j=[\mu_j(0),\cdots\cdots\mu_j(T-2),\ \mu_j(T-1)],\ j=1,2$$
$$\Sigma_j=\{\sigma_j(s-t),\ s,t=0,1,2,\cdots T-2,\ T-1\}$$
$$\delta=\mu_1-\mu_2$$

を表わす。いまここでデータ長が十分大であれば，次のような関係式が成立する[11]。

$$\lim_{T\to\infty}T^{-1}\log|\Sigma_j|=\int_{-\pi}^{\pi}\log P_j(\omega)\frac{d\omega}{2\pi} \tag{8}$$

Kullback 情報量の時間領域の表示である式(7)へ式(8)を代入すると，式(9)となる[10]。

$$I(1:2)=\frac{1}{2}\int_{-\pi}^{\pi}\left(\frac{P_1(\omega)}{P_2(\omega)}+\log\frac{P_2(\omega)}{P_1(\omega)}-1\right)\frac{d\omega}{2\pi}+\int_{-\pi}^{\pi}\frac{|D(\omega)|^2}{P_2(\omega)}\frac{d\omega}{2\pi} \tag{9}$$

ここで

$$D(\omega)=\int_{-\infty}^{\infty}\delta(t)\,e^{-j\omega t}\,dt$$

であり，式(7)における平均値 μ_1 と μ_2 の差，すなわち

$$\delta(t)=\mu_1(t)-\mu_2(t)$$

のフーリエ変換を表わす。ここで，脳波時系列の情報は交流成分のみで，直流成分は通常帯域フィルタで除去されている，さらに睡眠中の脳波時系列は，せいぜい 4～5 秒間であれば，大まかに定常性と正規性の条件を満足している[12] ことから，

$$\delta(t)=\mu_1(t)-\mu_2(t)\simeq 0 \tag{10}$$

とおくことができる。式(9)と式(10)より，本論文では，Kullback 情報量が

$$I(1:2)=\frac{1}{2}\int_{-\pi}^{\pi}\left(\frac{p_1(\omega)}{p_2(\omega)}+\log\frac{p_2(\omega)}{p_1(\omega)}-1\right)\frac{d\omega}{2\pi} \tag{11}$$

となる。式(3)と式(4)より，次式を定義することができる

$$V(\omega)=\log\frac{p_1(\omega)}{p_2(w)}=\log\left\{\frac{|E_1(e^{j\omega})|^2}{|A_1(e^{j\omega})|^2}\right\}-\log\left\{\frac{|E_2(e^{j\omega})|^2}{|A_2(e^{j\omega})|^2}\right\} \tag{12}$$

式(11)と式(12)より

$$I(1:2)=\frac{1}{2}\int_{-\pi}^{\pi}(e^{V(\omega)}-V(\omega)-1)\frac{d\omega}{2\pi} \tag{13}$$

となり，式(13)の表現はスペクトル情報の抽出のため，最尤法により展開されたスペクトル適合誤差尺度[6] と等価である。

さて，自己回帰モデルの安定性の条件より $A_i(Z)$ のすべての零点が単位円内に存在しなければならない。この条件のもとで次式の展開がなされる。すなわち[13]

$$\begin{aligned}
&\int_{-\pi}^{\pi}\log\{|A_i(e^{j\omega})|^2\}\frac{d\omega}{2\pi}\\
&=\int_{-\pi}^{\pi}\log\{|A_i(e^{-j\omega})|^2\}\frac{d\omega}{2\pi}\\
&=2\,\mathrm{Real}\left(\int_{-\pi}^{\pi}\log\{A_i(e^{-j\omega})\}\right)\frac{d\omega}{2\pi}\\
&=2\,\mathrm{Real}\left(\oint_p\log\left\{A_i\left(\frac{1}{Z}\right)\right\}\frac{dZ}{2\pi jZ}\right.\\
&=2\,\mathrm{Real}\{\log(A_i(\infty))\}\\
&=2\,\mathrm{Real}\,[\log(1)]=0
\end{aligned}$$

が導出できる。ここで $\oint_p$ は Z 平面上の単位円を含む一周積分である。

すなわち $$\int_{-\pi}^{\pi}\log|A_i(e^{j\omega})|^2\frac{d\omega}{2\pi}=0 \tag{14}$$

が成立する。式(14)を適用すると式(12)の積分は

$$\int_{-\pi}^{\pi}V(\omega)\frac{d\omega}{2\pi}=\int_{-\pi}^{\pi}\log\frac{|E_1(e^{j\omega})|^2}{|E_2(e^{j\omega})|^2}\cdot\frac{d\omega}{2\pi} \tag{15}$$

となる。従って Kullback 情報量，ここでは式(13)は次式で与えられる。

$$I(1:2)=\frac{1}{2}\left\{\int_{-\pi}^{\pi}\frac{|E_1(e^{j\omega})|^2}{|E_2(e^{j\omega})|^2}\cdot\frac{|A_2(e^{j\omega})|^2}{|A_1(e^{j\omega})|^2}\cdot\frac{d\omega}{2\pi}-\int_{-\pi}^{\pi}\log\frac{|E_1(e^{j\omega})|^2}{|E_2(e^{j\omega})|^2}\cdot\frac{d\omega}{2\pi}-1\right\} \tag{16}$$

ここで，式(16)の第 2 項に式(8)を適用すると

$$\int_{-\pi}^{\pi}\log\frac{|E_1(e^{j\omega})|^2}{|E_2(e^{j\omega})|^2}\frac{d\omega}{2\pi}=\log\frac{\sigma_1^2}{\sigma_2^2}$$

となる。従って式(16)は

$$I(1:2)=\frac{1}{2}\left\{\int_{-\pi}^{\pi}\frac{|E_1(e^{j\omega})|^2}{|E_2(e^{j\omega})|^2}\cdot\frac{|A_2(e^{j\omega})|^2}{|A_1(e^{j\omega})|^2}\frac{d\omega}{2\pi}-\log\frac{\sigma_1^2}{\sigma_2^2}-1\right\} \tag{17}$$

ここでスペクトル密度 $|E_i(e^{j\omega})|^2$ が独立な定常正規過程という仮定で，

$$|E_i(e^{j\omega})|^2=\sigma_i^2$$

と置くことが出来る。従って

$$I(1:2)=\frac{1}{2}\left\{\left(\frac{\sigma_1}{\sigma_2}\right)^2\int_{-\pi}^{\pi}\frac{|A_2(e^{j\omega})|^2}{|A_1(e^{j\omega})|^2}\frac{d\omega}{2\pi}-\log\left(\frac{\sigma_1}{\sigma_2}\right)^2-1\right\} \tag{18}$$

さらに $[A_2(Z)-A_1(Z)]$ は $A_1(Z)$ に直交することから

$$|A_2(Z)|^2=|A_2(Z)-A_1(Z)|^2+|A_1(Z)|^2 \tag{19}$$

が成立する。したがって式(18)と式(19)より

$$I(1:2)=\frac{1}{2}\left[\left(\frac{\sigma_1}{\sigma_2}\right)^2\left\{1+\int_{-\pi}^{\pi}\frac{|A_2(e^{j\omega})-A_1(e^{j\omega})|^2}{|A_1(e^{j\omega})|^2}\frac{d\omega}{2\pi}\right\}-\log\left(\frac{\sigma_1}{\sigma_2}\right)^2-1\right] \tag{20}$$

が成立する。式(20)の｛　｝の中は自己回帰モデルの予測残差の尤度比に等価である。ところで式(13)の被積分関数は変数

$$d=\log\frac{p_1(\omega)}{p_2(\omega)}$$

とおいたとき，d の正，負の値に対して非対称な関数である。しかるに，このような d の変数に関し対称な関数であることが望まれる。そこで Kullback 情報量 $I(1:2)$ と同様に情報量 $I(2:1)$，すなわち

$$I(2:1)=\int f_2(x_1\cdots x_k)\log\frac{f_2(x_1\cdots x_k)}{f_1(x_1\cdots x_k)}dx_1dx_2\cdots dx_k$$

と定義できる。$I(2:1)$ に関して，上述までの展開を行なうと，

$$I(2:1)=\frac{1}{2}\left[\left(\frac{\sigma_2}{\sigma_1}\right)^2\left\{1+\int_{-\pi}^{\pi}\frac{|A_1(e^{j\omega})-|A_2(e^{j\omega})|^2}{|A_2(e^{j\omega})|^2}\frac{d\omega}{2\pi}\right\}-\log\left(\frac{\sigma_2}{\sigma_1}\right)^2-1\right] \tag{21}$$

となる。Kullback は $I(1:2)$ と $I(2:1)$ を加えた Divergence 尺度を次式 $J(1:2)$ で定義した。

$$J(1:2)=I(1:2)+I(2:1) \tag{22}$$

Divergence 尺度 $J(1:2)$ は，変数に関し対称な関数となる。式 $J(1:2)$ は 2 つのスペクトル密度を分離する

ため展開された Cosh 尺度[7]と等価となる。

式(20)で $\sigma_1=\sigma_2$ という条件で，Kullback 情報量は

$$I(1:2)=\frac{1}{2}\int_{-\pi}^{\pi}\frac{|A_2(e^{j\omega})-A_1(e^{j\omega})|^2}{|A_1(e^{j\omega})|^2}\frac{d\omega}{2\pi}$$

$$=\frac{1}{2}(\delta/\alpha-1) \tag{23}$$

ここで δ/α は自己回帰モデルの予測誤差の尤度比を表わし，α と δ は各々次式で表わされる。

$$\alpha=\sum_{n=-M}^{M}\gamma_{1a}(n)\gamma_{1x}(n)$$

$$\delta=\sum_{n=-M}^{M}\gamma_{2a}(n)\gamma_{1x}(n) \tag{24}$$

ここで γ_{1a} と γ_{2a} は自己回帰モデルの係数の自己相関関数，および γ_{1x} と γ_{2x} はデータの自己相関関数を表わす。

γ_{ia}, γ_{ix} $(i=1,2)$ の i はパワスペクトル密度 $p_i(\omega)$ の i 区間のデータを示す。Mは自己相関の次数を表わす。

4. Kullback の情報量とスペクトル誤差量

睡眠中の脳波時系列は短時間定常時系列の継続したものと考えられる[12]。ここで短時間定常時系列とは，この時間内でスペクトル推定値が有意に変化しないことを意味する。Praetorius らによって，脳波データ解析のための適応的セグメンテーション（区分化）の方法がとられた。適応的区分化で展開された尺度として，スペクトル誤差量がある。本節では，Kullback 情報量とスペクトル誤差量の関係について明らかにする。Kullback 情報量 $I(2:1)$ の積分項の関数は式(11)と同様に次式で表わされる。

$$\int_{-\pi}^{\pi}\left(\frac{p_2(\omega)}{p_1(\omega)}+\log\frac{p_1(\omega)}{p_2(\omega)}-1\right)\frac{d\omega}{2\pi}$$

上式を次式のように変換する。

$$\int_{-\pi}^{\pi}\left(\frac{p_2(\omega)}{p_1(\omega)}-\log\frac{p_1(\omega)+(p_2(\omega)-p_1(\omega))}{p_1(\omega)}-1\right)\frac{d\omega}{2\pi} \tag{25}$$

ここでスペクトル密度関数 $p_1(\omega)$ が $p_2(\omega)$ にほぼ等しいと仮定するなら，式(25)の第2項目はテーラー展開により，次式で表わされることになる。

$$\log\left\{1+\left(\frac{p_2(\omega)-p_1(\omega)}{p_1(\omega)}\right)\right\}=\left(\frac{p_2(\omega)-p_1(\omega)}{P_1(\omega)}\right)$$

$$-\frac{1}{2!}\left(\frac{p_2(\omega)-p_1(\omega)}{p_1(\omega)}\right)^2+\text{高次の項} \tag{26}$$

式(26)の高次の項を無視すると，式(25)は式(26)より式(27)のように表わされる。

$$\int_{-\pi}^{\pi}\left\{\frac{p_2(\omega)}{p_1(\omega)}+\log\frac{p_1(\omega)}{p_2(\omega)}-1\right)\frac{d\omega}{2\pi}$$

$$=\frac{1}{2!}\int_{-\pi}^{\pi}\left(\frac{p_2(\omega)-p_1(\omega)}{p_1(\omega)}\right)^2\frac{d\omega}{2\pi} \tag{27}$$

さらにデータの自己回帰モデルの残差系列のスペクトル密度が平坦と仮定すると，$p_1(\omega)$ は次式で表わされる。

$$p_1(\omega)=\frac{\sigma_1^2}{|A_1(e^{j\omega})|^2} \tag{28}$$

またもう一方のスペクトル密度 $p_2(\omega)$ は

$$p_2(\omega)=\frac{|E_2(e^{j\omega})|^2}{|A_2(e^{j\omega})|^2}=\frac{|E_2(e^{j\omega})|^2}{|A_1(e^{j\omega})|^2} \tag{29}$$

とおく。ここで式(29)の分母が $A_2(e^{j\omega})=A_1(e^{j\omega})$ と仮定されている。式(28)と式(29)を使って，式(27)の積分項の関数は式(30)で表わされる。

$$\int_{-\pi}^{\pi}\left(\frac{p_2(\omega)-p_1(\omega)}{p_1(\omega)}\right)^2\frac{d\omega}{2\pi}$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(1-\frac{E_2^2(e^{j\omega})}{\sigma_1^2}\right)^2 d\omega \tag{30}$$

式(30)は Bodenstein[8] らによって導入されたスペクトル誤差量と等価である。式(28)と式(29)より，

$$\frac{p_2(\omega)}{p_1(\omega)}\cdot\sigma_1^2=E_2^2(e^{j\omega})=|E_2(e^{j\omega})|^2 \tag{31}$$

が成立する。式(31)は自己回帰モデルの残差系列のパワ・スペクトル密度を表わしている。従って，この残差系列の自己共分散関数を使って，近似的に

$$\frac{p_2(\omega)}{p_1(\omega)}\cdot\sigma_1^2=\gamma_0+\sum_{k=1}^{m}\gamma_k\cos(\omega k) \tag{32}$$

と表わされる。ここで $\{\gamma_k\}$ $(k=0,1,\cdots m)$ は自己共分散関数を示す。式(32)を使って，式(30)の積分項の関数は

$$(\sigma_1^2-\gamma_0)^2+2\sum_{k=1}^{m}\gamma_k^2 \tag{33}$$

と表わされる。式(33)は実際の脳波時系列へ適用する際に有用である。

以上までに展開した Kullback 情報量を非定常時系列の具体的例として睡眠脳波時系列へ適用し，一晩中の睡眠段階を計算機により自動的に適用した結果を以下の節に示す。

5. Kullback 情報量による脳波時系列の区分化

睡眠脳波時系列は非定常時系列と見なされる。ここでは，前節までに展開した Kullback 情報量を睡眠脳波時系列へ適用し，一晩中のヒトの睡眠段階の分類を行った。最初に非定常時系列を定常小区間の時系列へ区分化する必要がある。次に区分化された時系列の時系列パタンの特徴抽出による分類を行ない，睡眠段階を決定することになる。ここでは以下のような手順により行なった。

(1) 睡眠脳波時系列を Kullback 情報量により区分化を行なう。

(2) 脳波時系列から典型的な脳波パタンを選択する。

このことは，各睡眠段階にあらわれる特徴的な睡眠脳波をその段階の標準パタンとしている。覚醒（閉眼安静時）段階での α 波（10Hz 前後の波），軽い睡眠状態での紡錘波（12Hz～14Hz 波），深い睡眠状態でのデルタ波（3Hz 前後の波）等が大きな特徴となる。さらに，これらの波の継続時間も判別の際の手がかりとなる。本研究では次の節で示すように7個の標準脳波パタンを選択した。

(3) 標準脳波パタンを自己回帰モデルへあてはめ，その自己回帰モデルの係数と区分化された脳波時系列の間の Kullback 情報量の計算を行なう。

(4), (3)の Kullback 情報量，すなわち，ここでは2つのクラスの距離として，その値が最小となるような標準パタンのクラスを比較に用いた区分化の脳波時系列のクラスと判定する。

(5) 医師が睡眠脳波を判定する際に行なう，20秒間の記録用紙に対応して，その20秒間について(4)の判定をくり返し，その期間の睡眠段階を決定する。

本研究では，脳波時系列の区分化のため Kullback 情報量として，式(23)を用いた。図1に睡眠段階の覚醒時 (Stage W) の脳波時系列の区分化を示す。

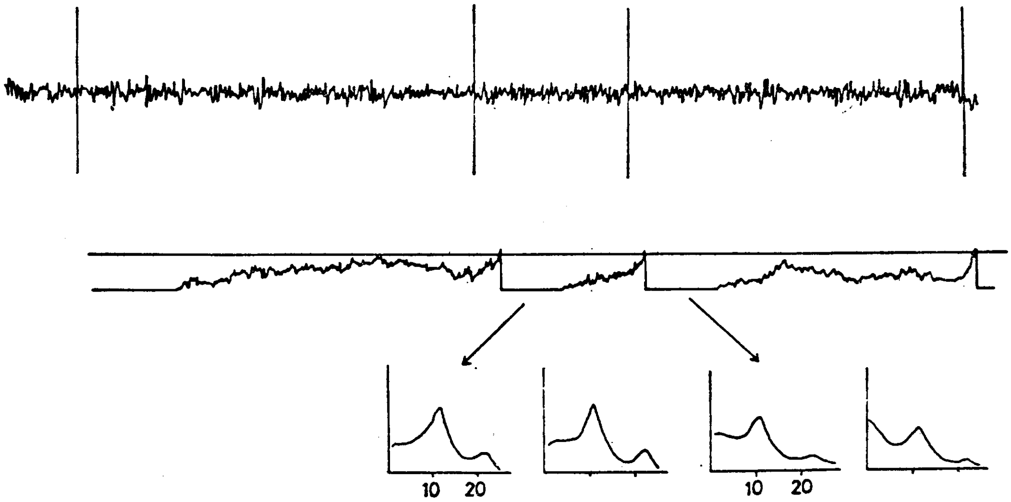

Fig. 1 Segmentations of the EEG during sleep stage W. On the top, the EEG and the vertical line of the segmentation are shown. On the middle, the values of the measure of (23) are indicated, while on the bottom, power spectral densities of the data between the segmentations are shown. The abscissa denotes frequency by *Hz*, while the ordinate denotes gain of power spectral densities by logarithmic scale.

図1の上段は覚醒段階の脳波時系列で，たての線分が小区間に区分化したことを示す。この区分は中の段の式(23)の値が，前もって設定しておいた閾値（threshold，ここでは 0.4 とおいた）を越えた場合，区分化の計算をして決定される。式(23)の値が閾値を越えた次の段階で零の値を2秒間継続する。このことは，この2秒間のデータを正規定常と見なして，自己回帰モデルのあてはめを行っている段階である。次に，この2秒間のデータを最初のクラス1のデータとして，20msec サンプリングごとの新しいデータを加えた2秒間データを第2のクラスのデータと見なす。そして，クラス1とクラス2の間の Kullback 情報量，ここでは式(23)の計算を行ったものである。図1の下段は，最初の2秒間のクラス1のデータのパワ・スペクトルと，クラス1を含む区分化された区間の全体のパワ・スペクトルを示したものである。横軸の数字の単位は，Hzである。クラス1のパワ・スペクトルの方に矢印をつけて示してある。クラス1と区分化された区間全体のパワ・スペクトルの形がほとんど等しい。図2には睡眠段階2の場合の脳波時系列の区分化を示す。

6. Kullback 情報量による脳波時系列の分類

睡眠脳波時系列の標準脳波パタンとして，以下に示す7個のパタンを取り上げた。Kullback 情報量を2つのパタンのクラス間の距離として計算するため，標準脳波パタンの時系列を自己回帰モデルへ適合させ，係数を求めた。以下にその標準パタンを示す。

(1) 高振幅の δ 波（L-δ で示す）
(2) 低振幅の δ 波（S-δ で示す）
(3) θ 波　（θ で示す）
(4) 睡眠段階1の特徴的な脳波パタン

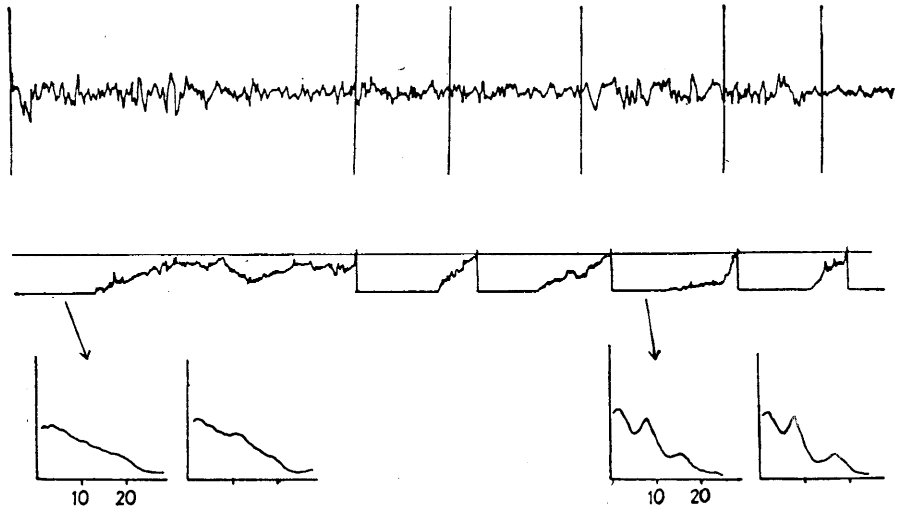

Fig. 2 Segmentations of the EEG during sleep stage 2. Similar representations are shown in Figure 1 are used.

(5) 睡眠段階 REM の特徴的は脳波パタン

(6) α波（α で示す）

(7) 紡錘波（S_p で示す）

上記の7個の標準パタンの自己回帰モデルの係数と区分化された脳波時系列から Kullback 情報量による距離を計算し，その区分化された時系列の脳波パタンのクラスを決定した。さらにその時系列の時間的長さも睡眠段階判別のためのパラメータとした。睡眠段階を決定するためのプログラムのフロー・チャートを図3に示す。この段階の決定の手順は，主に Rechtschaffen と Kales[14] のマニュアルによる。

図3で $\alpha>0$ は α 波の存在を示す。$L\text{-}\delta>0$ は高振幅 δ 波の存在を示す。MAX は観察している 20秒間のデータの中で時間的に一番長い波の長さを表わす。この20秒間は医師が通常睡眠脳波の記録用紙から読みとる最小単位の時間である。図3における LENGTH は δ 波の時間的長さで，100，200 および 500 は各々2秒間，4秒間，および10秒間に対応する（サンプリングが 20msec で 100個なら2秒間ということ。以下同様）。以上述べたような Kullback 情報量による睡眠脳波時系列の区分化と分類により，被験者（24歳，男子）の一晩中の睡眠段階を計算機により判定した。図4に，計算機による睡眠段階の判定結果を示す。睡眠段階は W（覚醒時），Stage 1（睡眠段階1），Stage 2（睡眠段階2）Stage 3（睡眠段階3），Stage 4（睡眠段階4），REM（逆説睡眠段階）の5段階に分かれ，図4の左側の欄各々の数字で表わしたものである。図5は図4と同じデータを記録用紙上から睡眠段階を判定したものである。

図4の計算機による結果と図5の医師による結果を比べると細部の点では異なっている個所がみられるが，

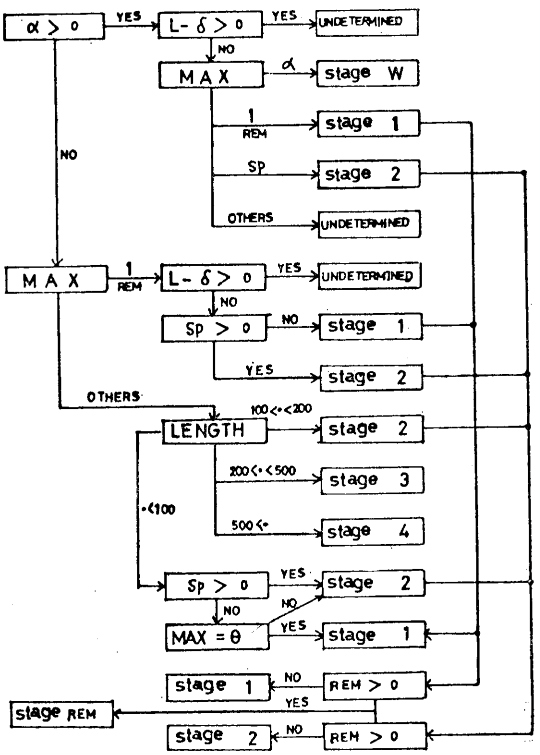

Fig. 3. Flow chart of the program to classify the sleep stages by means of Kullback's information measure.

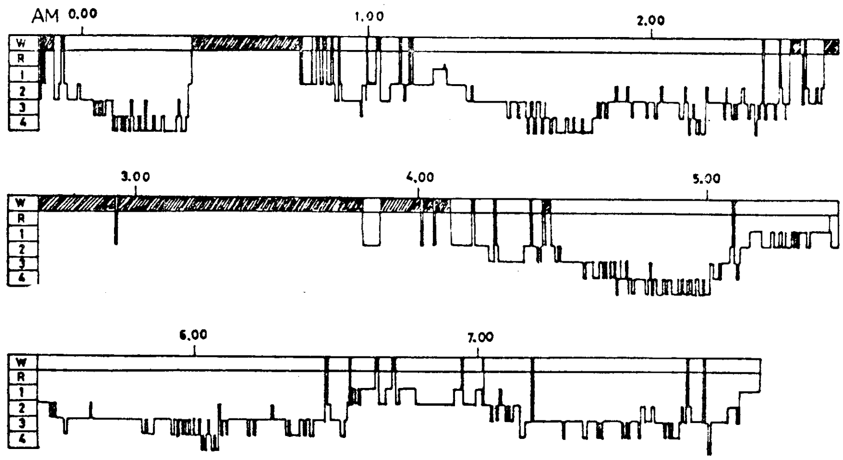

Fig. 4. Representation of the sleep stages during all night determined by computer. The symbols in the left column denotes the sleep stages, i.e. W (wakefulness), R (stage REM), 1 (stage 1), 2 (stage 2), 3 (stage 3), and 4 (stage 4). The abscissa denotes time by hour scale.

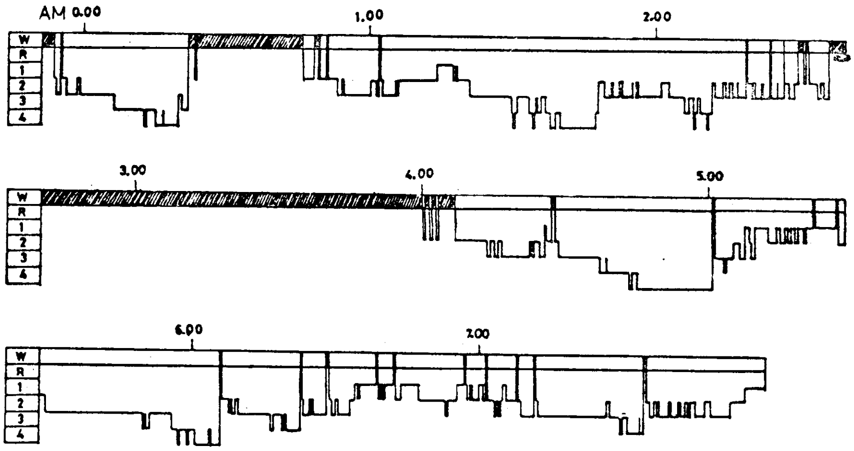

Fig. 5 Representation of the sleep stages during all night determined by a physician. Similar representations as shown in Figure 4 are used.

大まかな判定ではかなり一致していると解釈できる。本研究で導入した Kullback 情報量が時系列をとらえる尺度として，どのような位置づけにあるかを明らかにするため次節に Kullback 情報量，Divergence および Bhattacharyya Distance の間の比較を行った。

7. Kullback 情報量と Bhattacharyya Distanceの比較

パタン識別における特徴の量的な評価として，Kullback 情報量が導入され，また通信工学における信号検出の評価の関数として Kullback 情報量や Bhattacharyya Distance が導入されている。

Bhattacharyya Distance は 2つの確率密度関数 $p_1(x)$，$p_2(x)$ とおいたとき，

$$-\log\int\sqrt{p_1(x)p_2(x)}\,dx \tag{34}$$

で与えられる[15]。$p_1(x)$，$p_2(x)$ が各々正規密度関数 $N(\mu_1,\ \Sigma_1)$，$N(\mu_2,\ \Sigma_2)$ で表わされるとき，式(34)は

$$\frac{1}{8}\mu'\Sigma^{-1}\mu+\frac{1}{2}\log\left(\frac{|\Sigma|}{|\Sigma_1|^{\frac{1}{2}}|\Sigma_2|^{\frac{1}{2}}}\right) \tag{35}$$

で表わされる。ここで

$$\mu=\mu_1-\mu_2,\quad \Sigma=\frac{1}{2}\Sigma_1+\frac{1}{2}\Sigma_2$$

である。平均値が異なるパタンのグループの分類には，式(35)の第1項目が有用である[16]ことが示されているのに対し，第2項はほとんど考慮されていない。しかるに本

論文のような，脳波時系列を対象とし，時間領域で見ると，式(35)で平均値が同じパタンのグループを分類する問題となる。周波数領域で考えれば，時系列の振幅の変化の度合いとかスペクトル密度の変化の度合いでは式(35)の第2項目が重要となる。すなわち式(35)で$\mu=0$ の場合で，第2項目を評価関数として検討する必要がある。そこで本節では，はじめに時系列の相関がない場合について，

Kullback 情報量，Divergence，と Bhattacharyya Distance の比較を行なう。比較する時系列の平均値が等しく，各々相関関数値が零の場合，式(6)の右辺の第1項目は n個の独立な観測データに対し

$$I(1:2)=\frac{1}{2}\log\frac{|\Sigma_2|}{|\Sigma_1|}+\frac{1}{2}\mathrm{tr}\ \Sigma_1(\Sigma_2^{-1}-\Sigma_1^{-1})$$
$$=\frac{n}{2}\left(\log\frac{\sigma_2^2}{\sigma_1^2}+\frac{\sigma_1^2}{\sigma_2^2}-1\right) \qquad (36)$$

となる。さらに式(22)の Divergence は $I(1:2)$ と $I(2:1)$ より

$$J(1:2)=\frac{n}{2}\left\{\left(\frac{\sigma_1^2}{\sigma_2^2}+\frac{\sigma_2^2}{\sigma_1^2}\right)-2\right\} \qquad (37)$$

となる。一方 Bhattacharyya Distance を $B(1:2)$とおくと，

$$B(1:2)=\frac{1}{2}\log\left(\frac{\left|\frac{1}{2}\Sigma_1+\frac{1}{2}\Sigma_2\right|}{|\Sigma_1|^{\frac{1}{2}}|\Sigma_2|^{\frac{1}{2}}}\right)$$
$$=\frac{n}{2}\left[\log\left\{\frac{1}{2}(\sigma_1^2+\sigma_2^2)\right\}-\frac{1}{2}\log\sigma_1^2-\frac{1}{2}\log\sigma_2^2\right] \qquad (38)$$

いま比較を容易にするため，

$$\sigma_2=K\sigma_1^2 \qquad (39)$$

とおき，式(39)において Kの値を変えてみる。このとき式(36), (37), (38)で $\frac{n}{2}$の係数を除いて計算すると，

$K=0.25$のとき $I(1:2)=1.613$, $J(1:2)=2.250$, $B(1:2)=0.223$
$K=0.5$のとき $I(1:2)=0.307$, $J(1:2)=0.5$, $B(1:2)=0.059$
$K=1.0$のとき $I(1:2)=0.0$, $J(1:2)=0.0$, $B(1:2)=0.0$
$K=2.0$のとき $I(1:2)=0.193$, $J(1:2)=0.5$, $B(1:2)=0.059$
$K=3.0$のとき $I(1:2)=0.432$, $J(1:2)=1.333$, $B(1:2)=0.143$
$K=4.0$のとき $I(1:2)=0.636$, $J(1:2)=2.25$, $B(1:2)=0.223$

となる。これらの結果より，明らかなように，値そのものは

$$B(1:2)<I(1:2)<J(1:2)$$

であるが，Kを変化させた場合のお互いの比をとると，その比の表現を $\Delta B(1:2)$，$\Delta I(1:2)$ $\Delta J(1:2)$ とするとき，

$$\Delta I(1:2)<\Delta B(1:2)<\Delta J(1:2)$$

となる。すでに上述したようにJ (1:2)はKによる変化分に対し，対称性のある尺度であるが，Bhattacharyya Distance も，ほぼ対称性を有する尺度であることがわかる。ここで相異なる時系列の振幅の比較をした場合，比較の感度という観点から Divergence $J(1:2)$ がすぐれており，次に Bhattacharyya Distance，次に Kullback 情報量 $I(1:2)$ となることがわかる。

次に，相異なる時系列の周波数領域でのスペクトル密度の差違の評価を上記の3つの尺度で比較する。

Kullback 情報量 $I(1:2)$の式(36)を，観測している小区間で定常と見なすならば，

$$I(1:2)=\frac{1}{2}\log\frac{\sigma_2^2|S_2|}{\sigma_1^2|S_1|}+\frac{1}{2}\mathrm{tr}\,\sigma_1^2(S_1)\{\sigma_2^{-2}(S_2)^{-1}-\sigma_1^{-2}(S_1)^{-1}\}$$
$$=\frac{1}{2}\log\frac{\sigma_2^2}{\sigma_1^2}+\frac{1}{2}\log\frac{|S_2|}{|S_1|}+\frac{1}{2}\cdot\mathrm{tr}\left\{\frac{\sigma_1^2}{\sigma_2^2}\frac{(S_1)}{(S_2)}-I\right\} \qquad (40)$$

のように表現できる。(S_1)，(S_2) ここでは各々相関行列 Iは単位行列，$|S_1|$，$|S_2|$ は各々相関行列式を示す。自己回帰モデルにおいて，相隣り合う区間の残差の分散が等しい，すなわちと $\sigma_1^2=\sigma_2^2$ とおいて，相隣り合う区間のスペクトル密度の周波数の変化を検出することができる。このとき式(40)は式(41)で与えられる。

$$I(1:2)=\frac{1}{2}\log\frac{|S_2|}{|S_1|}+\frac{1}{2}\mathrm{tr}\ \{S_1S_2^{-1}-I\} \qquad (41)$$

また Divergence $J(1:2)$ の場合は

$$J(1:2)=\frac{1}{2}\mathrm{tr}\,(\Sigma_1-\Sigma_2)(\Sigma_2^{-1}-\Sigma_1^{-1})$$
$$=\frac{1}{2}\mathrm{tr}\left\{\frac{\sigma_1^2(S_1)}{\sigma_2^2(S_2)}+\frac{\sigma_2^2(S_2)}{\sigma_1^2(S_1)}-2I\right\} \qquad (42)$$

さらに相隣り合う区間のスペクトル密度の周波数の変化を検出するため，$\sigma_1^2=\sigma_2^2$ とおくと，

$$J(1:2)=\frac{1}{2}\mathrm{tr}\ (S_1S_2^{-1}+S_2S_1^{-1}-2I) \qquad (43)$$

となる。最後に Bhattacharyya Distance

$$B(1:2)=\frac{1}{2}\log\frac{\left|\frac{1}{2}\sigma_1^2(S_1)+\frac{1}{2}\sigma_2^2(S_2)\right|}{|\sigma_1^2(S_1)|^{\frac{1}{2}}|\sigma_2^2(S_2)|^{\frac{1}{2}}} \qquad (44)$$

は式(41)，式(43)と同様にして，周波数の変化を検出するため $\sigma_1^2=\sigma_2^2$ とおくと，

$$B(1:2)=\frac{1}{2}\log\frac{\left|\frac{1}{2}S_1+\frac{1}{2}S_2\right|}{|S_1|^{\frac{1}{2}}|S_2|^{\frac{1}{2}}} \qquad (45)$$

となる。

次に上に展開した周波数の変化を検出するための Kullback 情報量による式(41), Divergence による式(43)式, Bhattacharyya Distance による式(45)を簡単な例によって比較する。

いま,

$$S_1=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix},\ S_2=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \tag{46}$$

とおく。式(46)で ρ は相関係数値である。このとき式(41)より

$$I(1:2)=\frac{1}{2}\log\frac{1}{1-\rho^2}$$

式(43)より

$$J(1:2)=\frac{1}{1-\rho^2}-1$$

式(45)より

$$B(1:2)=\frac{1}{2}\log\frac{1-\frac{1}{4}\rho^2}{\sqrt{1-\rho^2}}$$

となる。

$\rho=0.0$のとき $I(1:2)=0.00$, $J(1:2)=0.00$, $B(1:2)=0.00$

$\rho=0.1$のとき $I(1:2)=0.005$, $J(1:2)=0.01$, $B(1:2)=0.001$

$\rho=0.2$のとき $I(1:2)=0.020$, $J(1:2)=0.04$, $B(1:2)=0.010$

$\rho=0.5$のとき $I(1:2)=0.144$, $J(1:2)0.333$, $B(1:2)=0.040$

$\rho=0.7$のとき $I(1:2)=0.337$, $J(1:2)=3.961$, $B(1:2)=0.103$

となる。これらの結果から，値そのものは

$$B(1:2)<I(1:2)<J(1:2)$$

であるが，ρ を変化させた場合のお互いの比をとると，その比の表現を $\square I(1:2)$, $\square J(1:2)$, $\square B(1:2)$ と表現するとき，ほぼ

$$\square I(1:2)<\square B(1:2)<\square J(1:2)$$

の関係が成立する。このことは，隣り合う時系列のスペクトル密度の周波数のみの変化分を検出する際の感度のよさの尺度の順番を見れば，Divergence, その次に Bhattacharyya Distance, Kullback 情報量と見なすことができる。

8. むすび

非定常時系列は種々の分野で見い出される。いままで定常時系列の解析法や処理法が詳しく研究され，また各分野でそれらの手法が適用されて来た。しかし非定常時系列の取り扱いについては統一的な解析法や処理法が望まれているにもかかわらず，まだ完成されていない現状である。本研究は非定常時系列を定常小区間に自動的に区分化するという試みに対して Kullback 情報量を導入したものである。従来 Kullback 情報量は時間領域でのシステムの設計の基礎として広く使われて来た。しかし周波数領域での検討があまりなされていない。そこで本論文では２つの時系列の周波数および振幅の差違を明確にするような統計量として Kullback 情報量を適用し，従来研究されて来た統計量と比較検討したところ，Kullback 情報量により統一的に把えられることが明らかとなった。さらにパタン識別の分野で，特徴抽出のために導入された Bhattacharyya Distance と比較検討したところ，Kullback 情報量を拡張した Divergence 尺度が一番優れていることが明らかとなった。しかし Divergence 尺度の値を求めるためには，データの前後方向からの計算が必要となり，Kullback 情報量の値を求めるよりも計算時間が長くなる。そこで本論文では，Kullback 情報量を使って，一晩中の睡眠脳波時系列の睡眠段階の判別を行った。その結果，計算機による睡眠段階の判別結果が，医師によるそれらとほぼ等しい結果を得ることができた。

文　献

1) A. Iwata, N.Suzumura and K. Ikegaya: "Pattern classification of phonocardiogram using liner prediction analysis." Medical & Biological Eng. & Comput. Vol 15, No. 4 (1977)

2) 松沢，石井，岩田，鈴村："時間的に可変な時系列モデルによる適応的同定"，信学論(A)，J61-A, 2, (昭53—02)

3) H. Akaike: "Information Theory and Extension of the Maximum Likelihood Principle," 2nd International Symposium on Information Theory, Akademia Kiado, Budapest (1973)

4) 沖田，太田，山口："Kullback の情報量を用いた非線形系の一同定法"，信学論(D)，J60-D, 9 (昭52-09)

5) 石井，岩田，鈴村：情報量による自己回帰モデルの評価と非正規過程への適用，信学論(A)，J61-A, 1(昭53-01)

6) 板倉，斉藤："統計的手法による音声スペクトル密度とホルマント周波数の推定"，信学論(A)，J53-A, 1, (昭45-01)

7) A.H. Gray, Jr and J. Markel: "Distance measures for speech processing," IEEE ASSP, vol 24, 5 (1976)

8) G. Bodenstein and H. M. Praetorius: "Feature

extraction from the electroencephalogram by adaptive segmentation", Proc. IEEE, vol 65, 5 (1977)

9) S. Kullback: Information theory and statistics, John Wiley (1959)

10) R.H. Shumway and A.N.Unger: "Linear discriminate function for stationary time series", J. American Stat. Assoc. vol 69, 348 (1976)

11) V. Grenander and G. Szego, Toeplitz forms and their applications, V. California Press (1958)

12) H. Sugimoto, N. Ishii, A. Iwata, N. Suzumura and T. Tomita: "On the stationarity and normality of the electroencephalographic data during sleep stages", Computer Programs in Biomedicine, vol 8 (1978)

13) J.D. Markel and A.H. Gray, Jr: Linear prediction of speech, Springer (1976)

14) A. Rechtschaffen and A. Kales: A manual of standardized terminology, technigues and scoring system for sleep stages of human subjects, NIH, Maryland (1968)

15) T. Kailath: "The Divergence and Bhattacharyya Distance measures in signal selection," IEEE Trans. Communication Tech., vol COM-15, 1 (1967)

16) K. Fukunaga and T.F. Krile: "A minmum distance feature effectiveness criterion," IEEE Trans. Information Theory, Sept. (1968)