

ETAPA DE ELABORARE 2

RAPORT PRIVIND ACTIVITATEA 2.1 / 2.2

Part 2.1 Descrierea algoritmilor

**Metodologie pentru sinteza de detectori cu structuri impuse
Reprezentări în planul timp-frecvență**

**Part 2.2 Cerințele tehnice ale sistemului
Stabilitatea sistemului în condiții de zgomot**

- ☐ **Raport științific în extenso**
- ☐ **Proces verbal de avizare internă**

RAPORTUL ȘTIINȚIFIC ȘI TEHNIC ÎN EXTENSO

CUPRINS

1. Introducere.....	4
2. Detectarea cu structură liberă.....	5
2.1. Ipoteze simple.....	6
2.2. Ipoteze compuse.....	8
2.3. Exemple de probleme de detectare.....	10
3. Detectarea cu structură impusă.....	12
3.1. Prezentarea problematicei.....	13
3.2. Construirea clasei de detectori $\mathcal{D}$	16
3.3. Alegerea și optimizarea criteriului $c(d, \mathcal{A}_n)$	20
3.4. Optimizarea unui ansamblu de criterii de contrast.....	25
3.5. Experimentări în planul timp-frecvență.....	29
4. Concluzii.....	34
Bibliografie.....	35



MONITORIZAREA NIVELULUI DE OBOSEALĂ MENTALĂ
A PERSONALULUI CARE ACTIVEAZĂ INFRASTRUCTURA CRITICĂ

Colectiv de realizare :

Echipa Coordonator – Universitatea Ovidius din Constanta

Echipa Partener – UTI GRUP SA

FUSURI DE SOMN

contact@fusuridesomn.ro; www.fusuridesomn.ro
0735 150 657 0722 604 071

O metodologie pentru sinteza de detectori cu structuri impuse Reprezentări în planul timp-frecvență. Stabilitatea sistemului în condiții de zgomot

1 Introducere

Problema detectării unui semnal înecat într-un zgomot aditiv este, în mod clasic, definit prin testul statistic al ipotezei următoare:

$$\begin{cases} H_0 : X = B \\ H_1 : X = S + B \end{cases} \quad (1)$$

unde X reprezintă observația, B reprezintă zgomotul aditiv și S semnalul pentru detector, S și B sunt presupuse independente. În practică, semnalul S poate fi cunoscut în totalitate, cu excepția unui vector θ de parametrii aleatori.

Acestei probleme i se asociază de regulă o parte $\{X_0, X_1\}$ a spațiului observațiilor X , dar și o funcție booleană $d(x)$, numită test de detecție, pe care o folosim prin următoarea ipoteză:

$$d(x) = 0 \text{ și } x \in X_0 \quad (H_0 \text{ presupus adevărat})$$

$$d(x) = 1 \text{ și } x \in X_1 \quad (H_1 \text{ presupus adevărat}).$$

Adoptând acest model, rezolvarea problemei (1) este echivalentă cu a căuta o parte a spațiului de observație, care este optimă în sensul unui criteriu definit, pornind de la obiectivele propuse. Atunci când se dispune de densitățile de probabilitate condiționate ale observației x , aplicarea unui criteriu de decizie, asemănător criteriului Bayes, conduce la o relație a testului de ipoteze (1), care constă în compararea unei variabile $\Lambda(x)$, numită statistică de detecție, cu un prag λ_0 . Împărțirea spațiului X definit mai sus, constă în a distinge realizările x din regiunile X_0 și X_1 , după următoarea regulă:

$$X_0 = \{x \in X : \Lambda(x) < \lambda_0\},$$

$$X_1 = \{x \in X : \Lambda(x) > \lambda_0\}.$$

În acest caz, avem o detecție cu structură liberă, iar funcția de decizie căutată nu este supusă nici unei constrângeri de mediu structural, dar depinde de alegerea unui criteriu de decizie și de cunoașterea legilor de probabilitate condiționate.

În schimb, atunci când cunoașterea apriori a ipotezelor H_0 și H_1 se limitează, la momentele statistice ale unui anumit ordin, metodele descrise anterior se dovedesc a fi inoperante. Una dintre metodele la care se poate recurge în astfel de situații constă în definirea, în prealabil, a naturii statisticii detectorului, lineară sau analitică, apoi în optimizarea parametrilor caracteristici după un criteriu dat. Din această cauză, această metodă este cunoscută ca detectarea cu structură impusă.

Pentru exemplificare, *figura 1* ilustrează noțiunile de detectare cu structură liberă și impusă prezentând cele două partiții $\{X_0, X_1\}$.

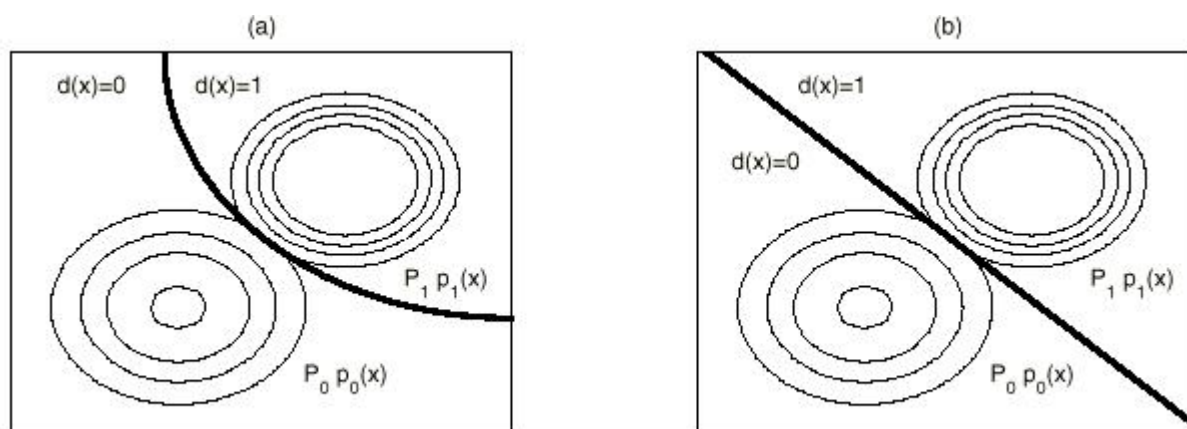


Figura 1. Detectarea unui semnal aleator gaussian dintr-un zgomot alb. Împărțirea spațiului se realizează pentru un detector cu structură liberă (a) și pentru un detector cu structură liniară impusă maximizat cu ajutorul criteriului Fisher (b).

În acest referat complex, ne propunem să abordăm în mod succesiv problemele specifice care apar în alegerea unei statistici de detectare, a unui criteriu optimal, de performanțele pe care sperăm să le atingă detectorul și a unei procesări dedicate optimizării acestui proces, obținând astfel stabilitatea acestuia în condiții de zgomot.

Referatul este organizat astfel: mai întâi, este prezentată part 2.1, dezvoltarea algoritmilor utilizând noțiunea de detectare cu structură liberă, expunând mai multe reguli de decizie, după ce, prezentăm cazul ipotezelor simple și compuse. Part 2.2 este consacrat, exclusiv, definirii unei metodologii pentru realizarea unui detector cu structură impusă și stabilitatea sistemului în condiții

de zgomot. Performanțele obținute sunt demonstrate prin rezolvarea problemei detecției în planul timp-frecvență, cu ajutorul rezultatelor întâlnite în teoria recunoașterii formelor.

2 Detectarea cu structură liberă

Detectarea cu structură liberă își justifică utilizarea datorită faptului că nu există nici o constrângere care să acționeze asupra structurii funcției de decizie căutate. Aceasta rezultă din alegerea criteriului de decizie, care este condiționat simultan de obiectivele propuse și informațiile de care dispunem. Această libertate structurală nu este posibilă decât dacă sunt cunoscute legile de probabilitate ale vectorului observațiilor x , condiționate ipotezelor:

$p_0(x)$: densitatea de probabilitate a vectorului aleator x , condiționat de ipoteza H_0 ,

$p_1(x)$: densitatea de probabilitate a vectorului aleator x , condiționat de ipoteza H_1 .

În această secțiune, vom expune conceptele generale ale detectării cu structură liberă pe care le-am considerat importante pentru continuarea lucrării.

2.1 Ipoteze simple

Într-o primă fază (timp), se presupune că H_0 și H_1 sunt ipoteze simple, ceea ce arată că ele sunt precizate printr-o lege unică de credibilitate condiționată. Criteriile utilizate în elaborarea unui detector cu structură liberă sunt următoarele:

Criteriul lui Bayes

Acest prim criteriu sancționează deciziile eronate și le recompensează pe cele corecte.

Astfel, asociind costul C_{ij} deciziei în favoarea ipotezei H_0 , în timp ce H_1 este efectivă, putem presupune că:

$$C_{00} < C_{01} \quad \text{și} \quad C_{11} < C_{10}.$$

Criteriul lui Bayes constă atunci în minimizarea costul mediu al unei decizii, care este definită prin:

$$C_{\text{mediu}} = \sum_{i,j=0,1} C_{ij} P\{d(x) = i; y = j\}$$

în care $Y = y$ este un boolean care detectează ipoteza pentru care observație $X = x$ constituie o realizare.

Se poate arăta că testul de detectare demonstrat de această regulă de decizie echivalează cu a compara câtul $\Lambda_{iv}(x)$ legilor de plauzibilitate (verosimilitate), numit raport de plauzibilitate, cu un prag [Tre68], [Sch91], [Poo94], [Don95]:

$$\begin{aligned}\Lambda_{iv}(x) &= \frac{p_1(x)}{p_0(x)} \geq \frac{p_0(C_{10} - C_{00})}{p_1(C_{01} - C_{11})}, \quad \text{pentru } d(x) = 1 \\ \Lambda_{iv}(x) &= \frac{p_1(x)}{p_0(x)} \leq \frac{p_0(C_{10} - C_{00})}{p_1(C_{01} - C_{11})}, \quad \text{pentru } d(x) = 0\end{aligned}\quad (2)$$

în care p_j este probabilitatea apriori a ipotezei H_j .

Astfel, alegerea regulii lui Bayes creează structura testului de detectare, aceea a detectorului cu prag ilustrată în figura 2, cu condiția ca densitățile $p_0(x)$ și $p_1(x)$ să fie cunoscute, ca și costurile și probabilitățile apriori p_0 și p_1 .

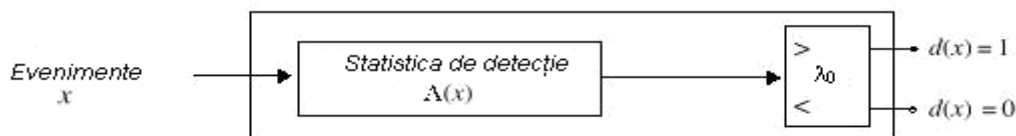


Figura 2. Principiul de detecție

În general, pot fi întâlnite, trei cazuri particulare ale criteriului Bayes, care corespund fiecare unei alegeri speciale de cost. Astfel putem întâlni:

a) *Criteriul de probabilitate aposteriori maxim.*

Acesta constă în alegerea costurilor în așa fel încât să verifice condiția: $C_{10} - C_{00} = C_{01} - C_{11}$.

În acest caz, ipoteza reținută maximizează probabilitatea aposteriori, regula de decizie definită prin relația (4.2) putând fi reformulată astfel:

$$\begin{aligned}P_1 p_1(x) &> P_0 p_0(x) \quad \text{pentru } d(x) = 1 \\ P_1 p_1(x) &< P_0 p_0(x) \quad \text{pentru } d(x) = 0.\end{aligned}$$

b) *Criteriul probabilității de eroare minime.*

Acest criteriu constă în atribuirea unui cost zero pentru o decizie corectă și a unui cost unitar pentru o eroare. Detectorul obținut, pe care-l notăm d_{Bayes} , minimizează probabilitatea de eroare P_e astfel:

$$d_{\text{Bayes}} = \arg \min_d P_e(d) \text{ cu } P_e(d) = P\{d(x) \neq Y\}.$$

c) *Criteriul de tip minimax.*

Acest criteriu se substituie celui al lui Bayes atunci când probabilitățile apriori ale ipotezelor sunt necunoscute. Ele implică căutarea probabilității P_0 cea mai defavorabilă, pentru detectorul Bayes, adică a probabilității care maximizează costul mediu minim.

Criteriul Neyman-Pearson

Când costurile și probabilitățile apriori sunt necunoscute, ceea ce nu permite aplicarea regulilor de decizie descrise anterior, suntem obligați să folosim probabilitățile condiționate.

Regula Neyman-Pearson constă în maximizarea probabilității de detectare P_d având stabilită în prealabil probabilitatea de alarmă falsă P_{fa} pentru un nivel dat, P_d și P_{fa} fiind definite astfel:

$$P_d(d) = P\{d(x) = 1 \text{ pentru } Y = 1\}$$

$$P_{fa}(d) = P\{d(x) = 1 \text{ pentru } Y = 0\}.$$

Rezolvarea acestei probleme de optimizare cu constrângere de inegalitate conduce la următoarea soluție [Sch91], [Poo94], [Don95]:

$$\begin{aligned} \Lambda_{iv}(x) &= \frac{p_1(x)}{p_0(x)} \geq \lambda_0 \text{ pentru } d(x) = 1 \\ \Lambda_{iv}(x) &= \frac{p_1(x)}{p_0(x)} \leq \lambda_0 \text{ pentru } d(x) = 0 \end{aligned} \quad (3)$$

În acest caz, pragul λ_0 este fixat prin probabilitatea P_{fa} și poate fi calculat cu ajutorul relației:

$$P_{fa} = \int_{\lambda_0}^{+\infty} p_0(\Lambda(x)) d\Lambda,$$

în care $p_0(\Lambda(x))$ este legea de probabilitate $\Lambda(x)$ și respectă condiția ipotezei H_0 .

În concluzie, strategiile Bayes și Neyman-Pearson sunt aproape identice. Ambele conduc la compararea raportului de plauzibilitate (verosimilitate) cu un prag λ_0 , acest prag depinzând de alegerea regulii de decizie.

2.2 Ipoteze compuse

Până acum, am presupus că legile de probabilitate condiționate sunt cunoscute. Totuși, sunt frecvente situațiile când trebuie să iei o hotărâre, fără a cunoaște în mod explicit un vector de parametrizare θ . Acest lucru conduce la distingerea a trei ipoteze, pe care le vom numi ipoteze compuse: θ aleator cu $p_i(\theta)$ cunoscute, θ aleator cu $p_i(\theta)$ necunoscute și θ determinist necunoscut.

Vom prezenta soluțiile, pentru fiecare din aceste cazuri.

Parametri aleatori, legi cunoscute

Regula Bayes poate fi aplicată dacă sunt cunoscute probabilitățile apriori P_i , legile condiționale $p_i(x)$ și $p_i(\theta)$ și costurile $C_{ij}(\theta)$. Prin aplicarea acestor probabilități ipotezelor de lucru se obține:

$$\frac{\int [C_{01}(\theta) - C_{11}(\theta)] p_1(x/\theta) p_1(\theta) d\theta}{\int [C_{10}(\theta) - C_{00}(\theta)] p_0(x/\theta) p_0(\theta) d\theta} \underset{d(x)=0}{\overset{d(x)=1}{\geq}} \frac{P_0}{P_1} \quad (4)$$

Dacă costurile C_{ij} sunt independente de vectorul de parametrizare θ , în această relație se recunoaște structura testului de detectare Bayes, care a fost prezentat în cadrul ipotezelor simple.

Parametri aleatori, legi necunoscute

Legile condiționale $p_i(\theta)$ sunt foarte rar cunoscute și din această cauză este imposibil de aplicat criteriul Bayes. Pentru această problemă nu există o soluție analitică, de obicei se identifică o lege susceptibilă de a fi urmată de vectorul aleator θ . Dintre parametrii caracteristici acestei legi sunt reținuți aceia care maximizează costul mediu minim al lui Bayes. În consecință, este vorba de o metodă de tip Minimax.

Parametrii determinați necunoscuți

Parametrii legilor de credibilitate sunt presupuși necunoscuți și este imposibil de aplicat regula lui Bayes. Costul mediu nu poate fi evaluat. În general, acest lucru conduce la adoptarea unei rezolvări cu ajutorul criteriului Neyman-Pearson, după cum urmează:

a) *Ipoteză simplă contra ipoteză compusă.*

Se presupune că H_0 este o ipoteză simplă și că H_1 este o ipoteză compusă, $p_1(x/\theta)$ depinde de un vector θ de parametri necunoscuți. Un demers ar consta în căutarea, dacă există, a celui mai puternic test. Acesta este optimal în sensul regulii Neyman-Pearson, dacă oricare ar fi θ : probabilitatea sa de detectare este superioară față de alte teste de detectare, având probabilitate de alarmă falsă fixată și independentă, pentru vectorul de parametri necunoscuți.

Când acest test nu există, θ poate fi estimat prin metoda credibilității maxime, ceea ce conduce la următoarea structură de detectare:

$$\Lambda(x) = \frac{\max_{\theta} p_1(x/\theta)}{p_0(x)} \underset{d(x)=0}{\overset{d(x)=1}{\geq}} \lambda_0.$$

b) *Ipoteză compusă contra ipoteză compusă.*

Condițiile de existență a unui test, sunt extrem de restrictive când H_0 și H_1 sunt ipoteze compuse [Sap90]. Din acest motiv, se recurge adesea, pentru estimarea parametrilor necunoscuți, pentru fiecare dintre ipoteze la metoda credibilității maxime. Testul de detectare echivalează în acest caz cu compunerea lui $\Lambda g(x)$, numit raport de credibilitate generalizat, cu un prag:

$$\Lambda g(x) = \frac{\max_{\theta} p_1(x/\theta)}{\max_{\theta} p_0(x/\theta)} \underset{d(x)=0}{\overset{d(x)=1}{\geq}} \lambda_0 \quad (5)$$

Pragul de detectare λ_0 este fixat prin alegerea probabilității de alarmă falsă. Constatăm că acest detector nu solicită cunoașterea apriori a vectorului θ . Obiectivul următoarei secțiuni este acela de a arăta principiile generale care au fost prezentate anterior, prin rezolvare a două probleme clasice de detectare. În cel de-al doilea exemplu, tratat pe larg, se prezintă o formulare timp-frecvență a statisticii de detectare. Pentru realizarea acest obiectiv, semnalele utilizate sunt presupuse continue, cadrul teoretic fiind favorizat de diversitatea literaturii care tratează detectarea cu structură liberă în planul timp-frecvență [Fla88], [Lem95], [Say95], [Say96a].

Nu trebuie neglijat faptul că această ipoteză necesită recurgerea la calcule matematice specifice, ca de exemplu noțiunea de densitate într-un spațiu de funcții și la teorema lui Radon Nikodym [Poo94].

2.3 Exemple de probleme de detectare

Problema de detectare pe care dorim să o analizăm constă în a determina decizia de prezență sau absență pentru un semnal s înecat în zgomotul B , utilizând una dintre următoarele ipoteze:

$$\begin{cases} H_0 : x(t) = B(t) \\ H_1 : x(t) = s(t) + B(t) \end{cases}, t \in T.$$

Semnalul s de detectare, pe timpul intervalului de observare T , poate fi presupus cunoscut sau aleator, real sau complex.

Semnal cunoscut perfect. Primul caz avut în vedere presupune că semnalul s este determinat și perfect cunoscut, iar zgomotul B este de tip alb, gaussian. În aceste împrejurări, se arată că statistica de detectare optimă, în sensul raportului de credibilitate, este definită prin produsul scalar al lui X și s [Poo94]:

$$\Lambda(X) = \int_0^T X(t)s^*(t)dt \quad (6)$$

Din relația (6) rezultă, că acest receptor, numit filtru adaptat, constituie un sistem liniar care maximizează raportul semnal/zgomot [Nor63], [Lem95].

Acest exemplu constituie un caz interesant în care alegerea unei reguli de decizie, ca aceea a lui Bayes, conduce la același rezultat ca optimizarea unui detector cu structură impusă de maximizare a unui criteriu de contrast, de regulă raportul semnal/zgomot [Tre68].

Semnal cu parametrii necunoscuți. În majoritatea problemelor de detectare pe care le-am întâlnite, faza inițială a semnalului rămâne dificilă de interpretat. În consecință, semnalul este presupus ca fiind uniform distribuit pe intervalul $[0, 2\pi]$, [Tre68], [Lem95].

Condiționat de ipoteza H_1 , în acest caz notăm:

$$x(t) = s(t; \phi_0) + B(t), \text{ cu } s(t; \phi_0) = s(t)e^{j\phi_0} \quad (7)$$

în care Φ_0 constituie singurul parametru necunoscut al problemei.

Cu ajutorul ipotezelor referitoare la vitezele de modulație a amplitudinii și fazei ale semnalului s , se arată că dezvoltarea raportului de credibilitate conduce la o statistică de detectare echivalentă cu [Tre68], [Lem95]:

$$\Lambda(x) = \left| \int_0^T x(t) s^*(t) dt \right|^2 \quad (8)$$

care corespunde ieșiri dintr-un filtru adaptat cu detector de anvelopă cvadratică.

Aplicând relația lui Moyal, se vede că această statistică de detectare poate fi reformulată astfel, pentru domeniul timp-frecvență:

$$\Lambda(x) = \iint_{t \in T, v} W_x(t, v) W_s^*(t, v) dt dv \quad (9)$$

Fiind vorba de un produs scalar între reprezentările Wigner-Ville ale observării și semnalul care trebuie detectat, structura de detectare (4.9) corespunde unui filtru adaptat timp-frecvență, prin analogie cu filtrul adaptat prezentat anterior.

În general corespondența Weyl [Koz92], a cărei relație Moyal constituie un caz particular, asociată oricărei statistici de detectare cvadractice, este dată de:

$$\Lambda(x) = \iint_{t, t'} x(t) a(t, t') x^*(t') dt dt' \quad (10)$$

cu următoarea formulare timp-frecvență echivalentă:

$$\Lambda(x) = \iint_{t, v} W_x(t, v) a_{TF}(t, v) dt dv \quad (11)$$

unde a_{TF} este definit astfel:

$$a_{TF}(t, v) = \int_{\tau} a\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi v\tau} d\tau \quad (12)$$

În orice statistică de detectare cvadratică, o formulare a observării în termeni de sumă ponderată a componentelor reprezentării Wigner-Ville, permite o interpretare a analizei timp-frecvență datorită configurației detectorilor căutați.

Acesta nu este singurul avantaj al acestei metode. Anumiți autori au arătat, că formularea timp-frecvență (11) permite în unele cazuri simplificarea implementării unui detector cvadratic, în special atunci când localizările temporale și în frecvență ale semnalului de detectat sunt necunoscute [Lem95], [Say95], [Kro97].

3 Detectarea cu structură impusă

Până în prezent, proprietățile statistice ale semnalului și zgomotului sunt bine determinate.

Structura detectorului nu necesită o cunoaștere prealabilă determinată, dar rezultă din aplicarea unui criteriu de decizie și de cunoaștere $p_0(x)$ și $p_1(x)$. Totuși, într-un mare număr de situații (date meteorologice, electroencefalogramme), nu se dispune de aportul pe care-l dă cunoașterea legilor probabilității condiționale, ceea ce face imposibilă determinarea raportului de credibilitate. Obiectul acestei secțiuni este prezentarea unei metodologii pentru elaborarea unui test de detectare, care să ofere anumite garanții în termenii definiți de probabilitate de eroare, atunci când nu se dispune decât de un ansamblu de n realizări notate $(x^{(i)}, y^{(i)})$. Într-un astfel de context, demersul adoptat în mod uzual constă în căutarea într-o clasă de detectori, definiți în prealabil, a unei soluții optime în sensul unui criteriu care nu depinde decât de aceste date. Această metodă reprezintă detectarea cu structură impusă.

3.1 Prezentarea problematicii

Fie A_n un ansamblu de n eșantioane $(x^{(i)}, y^{(i)})$. Se presupune că aceste date sunt realizări ale n perechi aleatoare (X^i, Y^i) independente, distribuite identic, urmărind legea $P\{(X;Y)\}$ care este necunoscută. Fie D o clasă de teste de detectare $d: R^L \rightarrow \{0;1\}$, în care L este dimensiunea vectorului aleator x . Fiind dată A_n , sinteza unui detector optimal cu structură impusă constă în a căuta în ansamblul D dat, un detector care să minimizeze probabilitatea de eroare condiționată următoare:

$$P_e(d_n) = P\{d_n(x, A_n) \neq Y / A_n\} \quad (13)$$

În această relație, $d_n(x, A_n)$ reprezintă un detector D determinat pe baza a n elemente (realizări) de A_n , la finalul unui proces care se numește instruire.

Problema minimizării lui $P_e(d_n)$ nu poate fi rezolvată pentru că legea $P\{(X,Y)\}$ este necunoscută. Se definește, în general, un criteriu $c(d_n, A_n)$ a cărui structură nu depinde decât de elementele A_n și a cărui optimizare conduce la un test de detectare care oferă anumite avantaje pentru termenii de probabilitate de eroare.

Ca exemplu, probabilitatea de eroare empirică este unul dintre acele criterii care posedă proprietățile căutate:

$$P_{\text{emp}}(d_n, A_n) = \frac{1}{n} \sum_{j=1}^n \delta(d_n x^{(i)} - y^{(i)}) \quad (14)$$

În practică, optimizarea lui P_{emp} se dovedește, uneori, a fi dificilă. Aceasta justifică numeroasele alternative care au fost propuse [Dud73], [Dub90], [Dev96].

Fie d_{Bayes} detectorul Bayes obținut prin minimizarea probabilității de eroare atunci când legile de probabilitate cu ipotezele H_0 și H_1 sunt bine determinate. Fie $d_{n,\text{opt}}(x, A_n)$ detectorul optim în sensul unui criteriu dat $c(d_n, A_n)$. După cum se arată în *figura 3*, eroarea de modelare definită cu ajutorul relației (15) este caracterizată prin eficacitatea lui $d_{n,\text{opt}}(x, A_n)$:

$$E_{\text{mod}}(d_{n,\text{opt}}) = P_e(d_{n,\text{opt}}) - P_e(d_{\text{Bayes}}) \quad (15)$$

pe care o putem reformula astfel:

$$E_{\text{mod}}(d_{n,\text{opt}}) = \underbrace{[P_e(d_{n,\text{opt}}) - \inf_{d \in D} P_e(d)]}_{E_{\text{estimativ}}} + \underbrace{[\inf_{d \in D} P_e(d) - P_e(d_{\text{Bayes}})]}_{E_{\text{aproximativ}}}, \quad d \in D \quad (16)$$

Reformularea definiției lui E_{mod} pune în evidență influența a două surse de eroare de natură diferită: eroarea de estimare și eroarea de aproximare. Prima notată E_{estim} , reflectă în mod unitar proprietățile criteriului ales $c(d_n, A_n)$ și eficacitatea procesului de învățare care i-a fost asociat, pentru un ansamblu de date A_n și o familie de teste de detectare D . Cea de a doua, notată E_{aprox} , nu depinde decât de alegerea inițială a lui D .

Minimalizarea erorii de modelare se bazează pe căutarea unui compromis între acești doi termeni antagonici: mărirea numărului de teste D conduce la o creștere a lui E_{estim} , în timp ce E_{aprox} descrește și invers. Pentru a ne convinge de existența acestui compromis, care a fost studiat de Vapnik și Chervonenkis în [Vap71] și care va fi abordat în capitolul 5, să presupunem că D regrupează ansamblul funcțiilor măsurabile $d: \mathbb{R}^L \rightarrow \{0;1\}$ și că $c(d, A_n)$ reprezintă probabilitatea de eroare empirică.

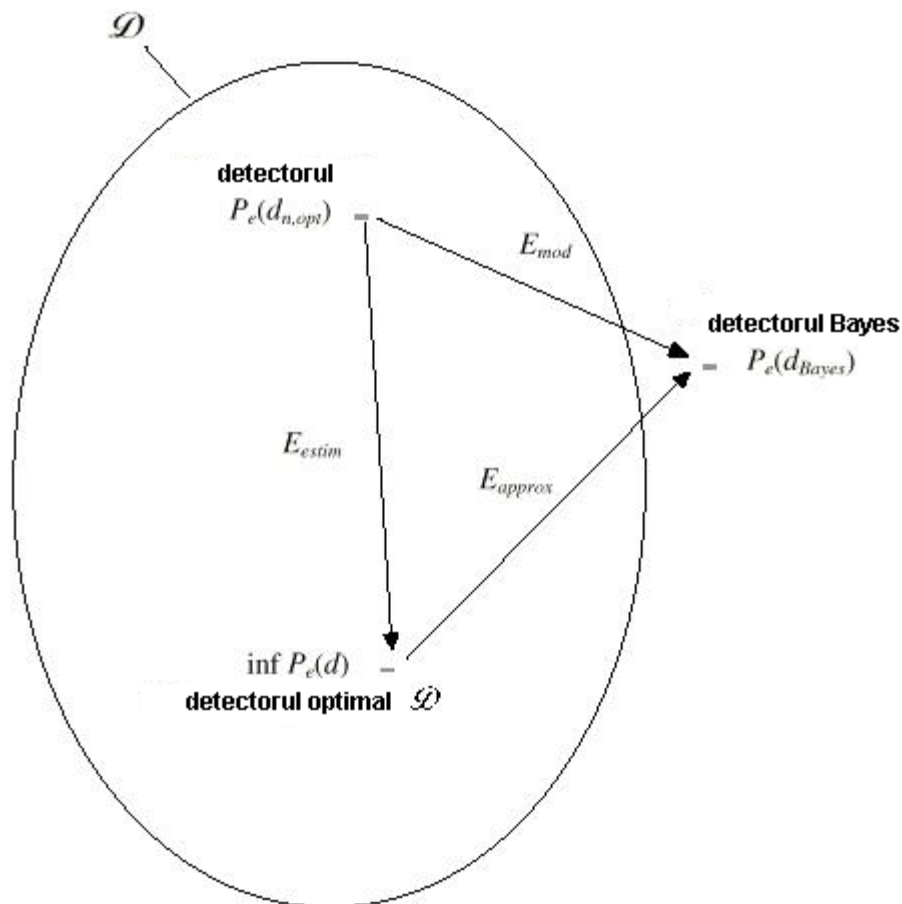


Figura 3. Realizarea detectorului optimizat al clasei $\mathcal{D}$.

În acest context, oricare ar fi legea $P\{(X,Y)\}$, se poate determina un test de detectare a cărui probabilitate de eroare empirică să fie nulă și al cărui rezultat să fie diferit pentru orice realizare $X=x$ care nu aparține lui A_n :

$$d_n(X, A_n) = \begin{cases} y^{(i)} \text{ și } X = x^{(i)}, i = 1, \dots, n \\ 0 \text{ sau } 1 \text{ aleatoare} \end{cases}$$

Acest detector determină, anularea erorii de aproximație în defavoarea erorii de estimare.

De remarcat faptul că, cazul dual ce constă în maximizarea lui E_{approx} și în anularea lui E_{estim} se poate observa dacă restrângem pe $\mathcal{D}$ la un test de detectare unică.

În continuare sunt prezentate problemele care apar în timpul realizării unui detector cu structură impusă:

1. Alegerea clasei D care autorizează o eroare de aproximare nulă, și nu se regăsește în cazul prezentat la secțiunea 4.3.2.

2. Căutarea unor criterii $c(d, A_n)$ plauzibile și algoritmi de instruire pentru minimizarea erorii de estimare (secțiunea 4.3.3).

3. Căutarea automată a unui compromis între eroarea de aproximare și eroarea de estimare (capitolul 5).

Recurgerea la rezultate ale teoriei învățării și ale recunoașterii formelor, cu scopul de a determina elementele necesare găsirii răspunsului la aceste întrebări, constituie originalitatea acestei lucrări.

3.2 Construirea clasei de detectori D

Se disting cel puțin două tipuri de metode pentru rezolvarea unei probleme de detectare.

Prima metodă vizează detectarea cu structură liberă, pentru care relația testului rezultă din aplicarea unui criteriu de detecție și din cunoașterea legilor de probabilitate ale observațiilor, condiționate de ipoteze. A doua metodă, numită detectarea cu structură impusă, presupune definirea în prealabil a unei clase D de teste de detectare, fiind reținut numai testul care este optim sensului pentru un criteriu dat. Obiectivul acestei secțiuni este acela de a alege clasa D de teste de detectare.

a) *Structuri lineare*. Testul de detectare cel mai simplu constă în a presupune că statistica $\Lambda(x)$ este o funcție lineară a componentelor observării:

$$\Lambda(x) = \sum_{i=1}^L a_i X[i] - \lambda_0 \begin{matrix} d(x)=1 \\ \geq \\ d(x)=0 \end{matrix} \begin{matrix} 0 \\ \\ 0 \end{matrix} \quad (17)$$

Acest tip de detectare prezintă avantajul de a fi legat direct noțiunii de filtrare, ceea ce ușurează interpretarea. Totuși, nu oferă soluții optime decât pentru o clasă redusă de probleme.

Uneori, în cazul când H_0 și H_1 urmează unele legi gaussiene, este strict obligatoriu ca acestea să aibă aceleași proprietăți statistice de ordinul doi [Don95].

Cu scopul de a răspunde satisfăcător unui număr cât mai mare de probleme, se poate avea în vedere posibilitatea îmbogățirii clasei D, prin adăugarea la testul de detectare definit prin relația (17), a informației referitoare la înmulțirile încrucișate ale componentelor $x(i)$ ale vectorului x .

Se obține în acest caz o familie de detectori bazați pe o statistică de detectare cvadratică:

$$\Lambda(x) = \sum_{i=1}^L a_i X[i] + \sum_{i=1}^L \sum_{j=1}^L b_{ij} X[i] X[j] - \lambda_0 = a^T X + X^T b X - \lambda_0 \quad \begin{matrix} d(x)=1 \\ \geq \\ d(x)=0 \end{matrix} \quad (18)$$

Aceste teste pot conduce, în anumite situații, la un caz de optim în sensul criteriului Bayes atunci când ipotezele H_0 și H_1 urmează legile gaussiene de parametri distincți [Don95], presupunând că s-ar putea ajunge la a și b .

În general, se poate demonstra că prin intermediul anumitor ipoteze bazate pe $p_0(x)$ și $p_1(x)$, statistica detectorului Bayes admite divizări de aceeași natură cu cele prezentate anterior.

Enunțul 1. [Dev96] Fie p legea de credibilitate a variabilei aleatoare x . Fie $\{z^{(i)}\}$ un ansamblu ortonormat complet de formațiuni limitate aparținând lui $L_2(\lambda)$, unde λ reprezintă măsura lui Lebesgue asupra lui $\mathbb{R}^L$. Dacă p aparține lui $L_2(\lambda)$, atunci densitățile condiționate aparțin în egală măsură lui $L_2(\lambda)$ și o formă echivalentă $\Lambda_{\text{Bayes}}(x)$ a statisticii de detectare a detectorului Bayes admite următoarea desfășurare:

$$\Lambda_{\text{Bayes}}(X) = P_{1|p_1}(X) - (1 - P_1)p_0(X) = \sum_{i=1}^{+\infty} a_i z^{(i)}(X) \quad (19)$$

unde coeficienții a_i sunt dați de:

$$a_i = \int z^{(i)}(x) \Lambda_{\text{Bayes}}(x) dx.$$

Convergența desfășurării (4.19) considerată în medie cvadratică este:

$$\lim_{k \rightarrow \infty} \int (\Lambda_{\text{Bayes}}(x) - \sum_{i=1}^k a_i z^{(i)}(x))^2 dx = 0.$$

Precizare: Între ansamblurile $\{z^{(i)}\}_i$ care verifică proprietățile enunțate mai sus, pot fi întâlnite polinoamele lui Legendre sau bazele lui Laguerre și Haar [Dev96].

Fie $\mathcal{D}$ clasa detectorilor liniari generalizați definiți prin:

$$d(x) = \begin{cases} 0 & \text{dacă } \sum a_i z^{(i)}(x) \leq 0 \\ 1 & \text{dacă } \sum_{i=1}^k a_i z^{(i)}(x) \geq 0 \end{cases} \quad (20)$$

în care $z^{(i)}(x)$ este o funcție oarecare a variabilei x .

Enunțul 1 stabilește convergența spre 0 a erorii de aproximare definită prin relația (16), în timp ce k tinde spre infinit, iar $\{z(i)\}_{1 \leq i \leq k}$ este un ansamblu ortonormat complet de funcții mărginite $L_2(\lambda)$.

Acest rezultat necesită o anumită cunoaștere apriori a legii de credibilitate a observării X , care este necunoscută prin ipoteză.

Aceasta conduce la renunțarea provizorie la clasa detectorilor liniari generalizați, deoarece ipotezele cerute de enunțul 1 nu sunt satisfăcătoare. Rezultatele prezentate în continuare sunt interesante deoarece urmăresc convergența erorii de modelare în cazul alegerii convenabile a ansamblului $\{z(i)\}_{1 \leq i \leq k}$, fără să facem ipoteze pentru $p_0(x)$ și $p_1(x)$.

b) *Noțiunea de consistență* [Dev96]. De obicei, este imposibil să atingi performanțele detectorului Bayes cu un detector cu structură impusă determinată pornind de la un ansamblu de intervale de dimensiune finită n . Totuși, în clasa $\mathcal{D}$ există o pereche $\{d_{n,\text{opt}}(X, A_n)\}_{n \geq 0}$ de detectori, obținuți prin optimizarea unui criteriu dat A_n , de exemplu probabilitatea de eroare:

$$P_e(d_{n,\text{opt}}) = P\{d_{n,\text{opt}}(x, A_n) \neq Y / A_n\}$$

să poată fi adusă în mod arbitrar aproape de $P_e(\text{dbayes})$, atunci când n tinde spre infinit. Noțiunea care ne interesează se bazează pe această ipoteză.

Definiția 2. Un șir $\{d_{n,\text{opt}}(X, A_n)\}_{n \geq 0}$ de detectori determinați prin optimizarea unui criteriu dat cu A_n este considerată consistentă pentru o lege $P\{(X,Y)\}$ dacă:

$$\lim_{n \rightarrow +\infty} E\{P_e(d_{n,opt})\} = \lim_{n \rightarrow +\infty} P\{d_{n,opt}(X, A_n) \neq Y\} = P_e(d_{Bayes}),$$

și foarte consistentă dacă:

$$\lim_{n \rightarrow +\infty} P_e(d_{n,opt}) = P_e(d_{Bayes}),$$

cu o probabilitate egală cu 1.

Probabilitatea de consistență garantează că probabilitatea de eroare a unui detector, obținut prin optimizarea unui criteriu pe o bază de instruire dată de mărimea n , converge către $P_e(d_{Bayes})$, în timp ce n tinde spre infinit. Șirul $\{d_{n,opt}(X, A_n)\}$, $n \geq 0$ este considerat foarte convergent dacă această proprietate este adevărată, oricare ar fi realizarea lui A_n . În final, se poate distinge situația când proprietatea de consistență nu este verificată decât pentru o lege de probabilitate $P\{(X,Y)\}$ dată, pentru cazul în care ea rămâne independentă față de aceasta. Următoarea definiție se bazează pe caracterul universal al consistenței.

Definiția 3. Șirul $\{d_{n,opt}(X, A_n)\}$, $n \geq 0$ este considerat universal (foarte) consistent dacă el este (foarte) consistent pentru orice lege $P\{(X,Y)\}$.

Majoritatea acestor definiții rezultă din teoria învățării, în continuare putem prezentat rezultatul principal al acestui subcapitol. Acesta rezultă dintr-un enunț generalizat, prezentat și demonstrat în [Dev96] și adaptat problemei care face obiectul acestui subcapitol.

Enunțul 4. Șirul $\{d_{n,opt}(X, A_n)\}_n$ al detectorilor lineari generalizați cu A_n este în mod universal puternic consistentă dacă:

$$k \rightarrow +\infty \text{ și } \frac{k}{n} \rightarrow 0, \text{ când } n \rightarrow +\infty,$$

și dacă funcțiile $z^{(1)}, z^{(2)}, \dots, z^{(L)}, \dots, z^{(k)}$, corespund polinoamele $X[1], X[2], \dots, X[L], X[1], X[2], \dots$. Se reamintește că $X[j]$ corespunde vectorului aleator X .

Acest enunț introduce o clasă particulară a detectorilor lineari generalizați cu condiția ca mărimea bazei de instruire să crească mai repede decât gradul polinoamelor. În consecință, în continuare testele de detectare utilizate vor fi de forma:

$$d(z) = \begin{cases} 0 & \text{dacă } \Lambda(z) = a^T z - \lambda_0 \leq 0 \\ 1 & \text{dacă } \Lambda(z) = a^T z - \lambda_0 > 0 \end{cases} \quad (21)$$

unde $Z[1], \dots, Z[L]$, $Z[L+1], \dots$ sunt polinoamele $X[1], \dots, X[L]$, $X[1]X[2], \dots$ și a un vector cu componente reale.

În concluzie, în afară de garanțiile de performanță reprezentate de consistența universală a acestei clase, determinarea unui detector este ușurată datorită apariției unui discriminant linear în spațiul transformărilor z .

Această ipoteză, se întâlnește în teoria recunoașterea formelor și va fi abordată în următoarea secțiune.

3.3 Alegerea și optimizarea criteriului $c(d, A_n)$

a) *Prezentarea ipotezelor.* Studiul anterior a permis caracterizarea unei clase D de detectori care autorizează o eroare de aproximare nulă, în cazul de față acela al structurilor lineare generalizate definite prin Enunțul 4. Problema care se ridică în continuare este aceea de a căuta un criteriu în care optimul corespunde cu minimumul erorii de estimare.

Reamintim că E_{estim} este definit astfel:

$$E_{estim} = P_e(d_{n,opt}) - \inf_{d \in D} P_e(d) \quad (22)$$

cu $d_{n,opt} = \arg \min_{d \in D} c(d, A_n)$,

unde $\arg \max_{d \in D} c(d, A_n)$ (4.23), conform teoriei lui c.

În [Dev96] se demonstrează că probabilitatea de eroare empirică P_{emp} corespunde criteriului c căutat. Aceasta arată că testul D , cel mai performant care se poate obține, plecând de la ipoteza de cunoaștere a unui ansamblu de instruire A_n , este acela care minimizează probabilitatea de eroare empirică a acestui ansamblu.

Într-adevăr, o detectare optimală cu n puncte etichetate $\mathbb{R}^L$ este demonstrată în [Zoh78] prin recurgerea la metode de rezolvare aproximative.

În aceste condiții, determinarea unui local optim ar putea consta în utilizarea unei metode de optimizare clasică.

Deoarece, P_{emp} este o funcție constantă pe segmente și ține cont de natura eșantionată a lui A_n , operația de minimizare este mai dificilă. Pentru a depăși această dificultate, anumiți autori au propus modificarea parțială a acestui criteriu, de exemplu, prin intermediul unei ferestre glisante (lissage) cu scopul de a limita la minimum numărul particularizărilor [Was72], [Dot75], [Fri76], [Wid88], [Skl79].

b) *Criterii de contrast.* O altă metodă constă în adaptare unui criteriu, numit de contrast, care cuantifică caracterul discriminant al unei statistici de detectare $\Lambda(z)$ pentru o ipoteză dată, pentru măsurarea unei distanțe dintre legile de probabilitate $p_0(\Lambda(z))$ și $p_1(\Lambda(z))$. Acest principiu este prezentat în *figura 4*. Unele dintre aceste criterii, ca divergențele lui Jeffreys și Kuelback-Leiber, afinitatea lui Bhattacharya sau chiar distanța lui Kolmozsare [Dev96], necesită cunoașterea lui $p_0(z)$ și $p_1(z)$ [Dev96].

Aceste criterii furnizează informații de calitate despre performanțele care se pot obține de la un detector, deoarece ei sunt strânși legați de eroarea lui Bayes.

Alte criterii necesită o cantitate mai mică de informații, care, se limitează la momentele condiționale de ordinul unu sau doi ale statisticii de detectare $\Lambda(z)$. În acest caz, argumente simple bazate pe considerații geometrice sunt suficiente pentru a justifica definirii lor, față de mulțimea proprietăților teoretice care vor fi prezentate în continuarea acestei prezentări.

Aceste criterii, care ne interesează din cauza informațiilor apriori reduse pe care le solicită, fac obiectul unei literaturi bogate. Din această cauză, în această lucrare nu vor fi prezentate decât cele mai cunoscute contribuții, de exemplu [And73], [Gar80], [Dub90] și [Dev96].

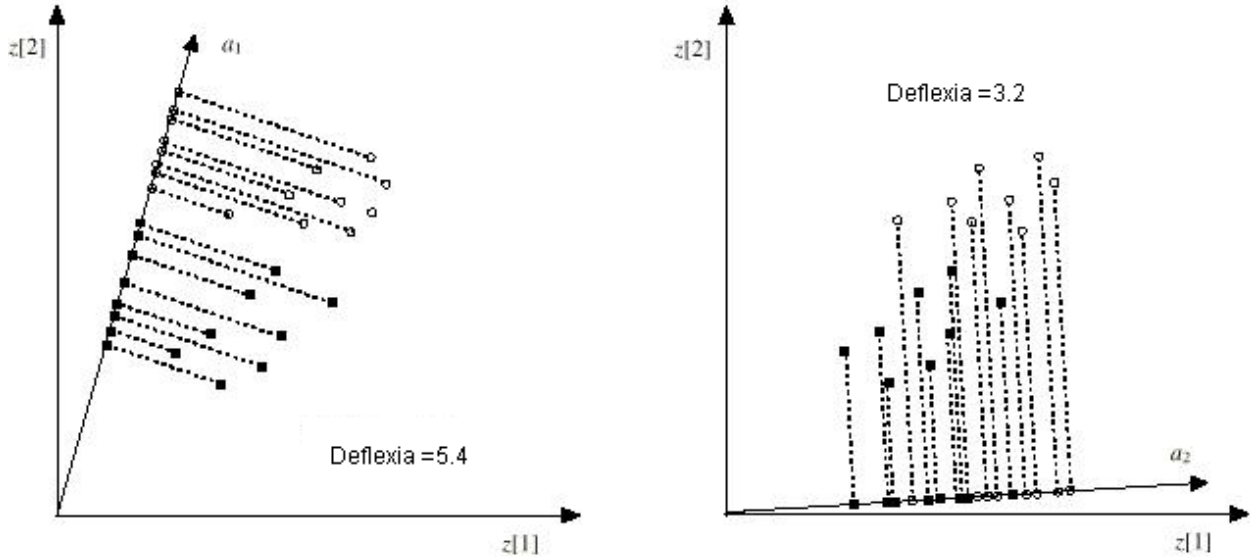


Figura 4. Evaluarea criteriului de contrast compus din deflexie și elementele de statistică asociate vectorilor a_1 și a_2 . Etichetele asociate sunt notate cu $\{\blacksquare, \circ\}$.

Fie η_i și σ_i^2 experiența matematică și variația condițională a statisticii $\Lambda(z)$. Raportul semnal/zgomot, deflexia și criteriul lui Fisher, ale căror denumiri variază în funcție de autori, sunt definite prin:

$$C_{RSB}(\eta_0, \eta_1, \sigma_0^2, \sigma_1^2) = \frac{(\eta_1 - \eta_0)^2}{\sigma_0^{-2}} \quad (24)$$

$$C_{deflex}(\eta_0, \eta_1, \sigma_0^2, \sigma_1^2) = 2 \frac{(\eta_1 - \eta_0)^2}{\sigma_0^2 + \sigma_1^2} \quad (25)$$

$$C_{Fisher}(\eta_0, \eta_1, \sigma_0^2, \sigma_1^2) = \frac{(\eta_1 - \eta_0)^2}{P_0 \sigma_0^2 + (1 - P_0) \sigma_1^2} \quad (26)$$

Optimizarea acestor criterii poate fi interpretată ca maximizarea dispersiei interipotezei Δ_{inter} efectuată simultan cu minimizarea dispersiei intraipotezei Δ_{intra} , dacă se adoptă următoarele definiții:

$$\Delta_{inter} = (\eta_1 - \eta_0)^2,$$

$$\Delta_{intra} = p \sigma_0^2 + (1 - p) \sigma_1^2, \text{ unde } \{p; 1 - p\} \text{ desemnează un echilibru dat.}$$

Interpretarea lui C_{RSB} , C_{Deflex} , C_{Fisher} este satisfăcătoare în termeni de măsură de contrast între ipotezele H_0 și H_1 , atunci când $p_0(\Lambda(z))$ și $p_1(\Lambda(z))$ sunt determinați în figura 5.

În caz contrar rezultatele obținute pot fi discutabile. În acest caz, se recomandă mărirea gradului polinomial al statisticii de detectare definit prin ecuația (21).

Următorul rezultat teoretic contribuie la justificarea interesului pe care-l prezintă criteriile de contrast care nu depind decât de numerele de ordinul 1 și 2 ale statisticii de detectare $\Lambda(z)$, atunci când nu se dispune de cunoașterea lui $p_0(z)$ și $p_1(z)$.

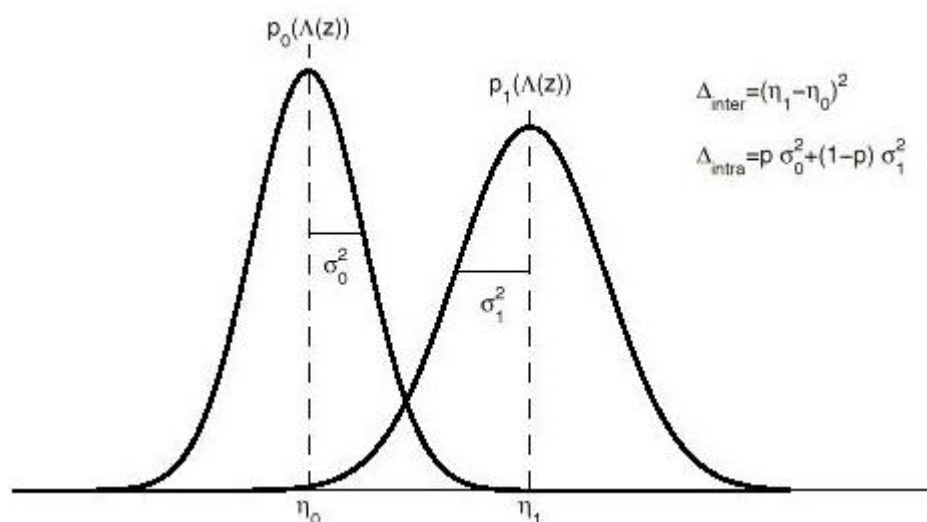


Figura 5. Definiția dispersiei interipotezei și intraipotezei. Principiul criteriului de contrast nu depinde de momentele de ordinul 1 și 2 ale statisticii de detectare.

Enunțul 4.5. [Fuk90] Raportul de credibilitate maximizează orice criteriu de contrast c care nu depinde decât de momente de ordin 1 și 2 ale statisticii de detectare.

Dacă d este un test de detectare definit prin:

$$d(z) = \begin{cases} 0 & \text{și } \Lambda(Z) \leq 0 \\ 1 & \text{și } \Lambda(Z) > 0 \end{cases} \quad (27)$$

Presupunem că statistica de detectare $\Lambda(z)$ este optimală în sensul unui criteriu $c(\eta_0, \eta_1, s_0^2, s_1^2)$

oarecare de tipul:

$$\eta_1 = E\{\Lambda(Z)/H_1\} = \int \Lambda(z)p_1(z)dz$$

$$s_1^2 = E\{\Lambda^2(Z)/H_1\} = \int \Lambda^2(z)p_1(z)dz.$$

Rezultatul final nu este afectat (c rămâne o funcție de η_i și σ_i^2) de imaginea diferitelor criterii prezentate anterior.

Reamintim că:

$$s_1^2 = \eta_1^2 + \sigma_1^2.$$

Calculul variațiilor lui c, notate δc , în funcție de cele ale lui $\Lambda(Z)$ notate $\delta\Lambda(Z)$, conduce la:

$$\delta_c = \frac{\partial c}{\partial s_0^2} \delta s_0^2 + \frac{\partial c}{\partial s_1^2} \delta s_1^2 + \frac{\partial c}{\partial \eta_0} \delta \eta_0 + \frac{\partial c}{\partial \eta_1} \delta \eta_1 \quad (28)$$

unde:

$$\delta \eta_1 = \int \delta \Lambda(z) p_1(z) dz \quad (29)$$

$$\delta s_1^2 = \int 2\Lambda(z) \delta \Lambda(z) p_1(z) dz \quad (30)$$

Utilizând rezultatele (4.29) și (4.30), relația δ_c poate fi reformulată astfel:

$$\delta_c = \int \left[2 \left\{ \frac{\partial c}{\partial s_0^2} p_0(z) + \frac{\partial c}{\partial s_1^2} p_1(z) \right\} \Lambda(z) + \left\{ \frac{\partial c}{\partial \eta_0} p_0(z) + \frac{\partial c}{\partial \eta_1} p_1(z) \right\} \right] \delta \Lambda(z) dz \quad (31)$$

Optimul lui $\Lambda(Z)$ în sensul lui c implică faptul că δ_c este nul, independent de $\delta\Lambda(Z)$. În consecință, termenul [...] este egal cu 0, ceea ce permite stabilirea relației următoare:

$$\Lambda(Z) = \beta_0 \frac{p_0(z)}{\alpha p_0(z) + (1-\alpha)p_1(z)} + \beta_1 \frac{p_1(z)}{\alpha p_0(z) + (1-\alpha)p_1(z)} \quad (32)$$

în care:

$$\alpha = \frac{\partial c / \partial s_0^2}{\partial c / \partial s_0^2 + \partial c / \partial s_1^2} = \frac{\partial c / \partial \sigma_0^2}{\partial c / \partial \sigma_0^2 + \partial c / \partial \sigma_1^2} \quad (33)$$

$$\beta_0 = -\alpha \frac{\partial c / \partial \eta_0}{2\partial c / \partial s_0^2} = -\alpha \frac{\partial c / \partial \eta_0}{2\partial c / \partial \sigma_0^2} \quad (34)$$

$$\beta_1 = (\alpha - 1) \frac{\partial c / \partial \eta_1}{2\partial c / \partial s_1^2} = (\alpha - 1) \frac{\partial c / \partial \eta_1}{2\partial c / \partial \sigma_1^2} \quad (35)$$

Combinarea relațiilor (32) și (27) conduce în final, la următoarea regulă de decizie:

$$\frac{p_1(z)}{p_0(z)} \begin{matrix} \geq \\ \leq \end{matrix} \lambda(\alpha, \beta_0, \beta_1), \quad \begin{matrix} d(x)=1 \\ d(x)=0 \end{matrix}$$

care constă în compararea raportului de credibilitate cu un prag $\lambda(\alpha, \beta_0, \beta_1)$ care depinde de criteriul de contrast c ales. Acest rezultat nu are nici o relevanță în determinarea statistică a lui $\Lambda(Z)$, acesta fiind presupus cunoscut.

Enunțul 5 justifică folosirea criteriilor de contrast care nu depind decât de momentele de ordinul 1 și 2 ale statisticii $\Lambda(Z)$, în timp ce lipsa (relativ) de informație despre ipotezele H_0 și H_1 necesită impunerea structurii detectorului. Detectorii lineari optimi, definiți în sensul unuia dintre aceste criterii, pot conduce la rezultate eronate, după cum prezintă *figura 6*.

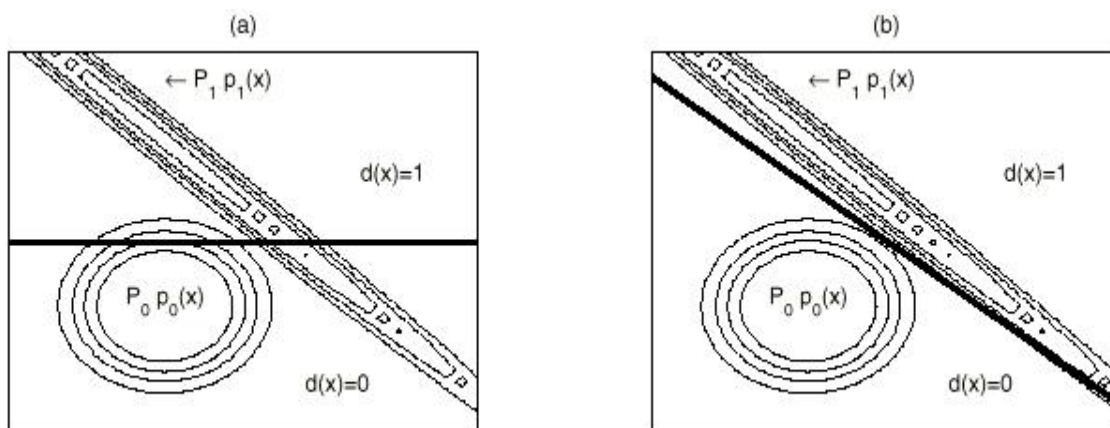


Figura 6. Detectarea unui semnal gaussian aleator cu probabilitatea apriori egală cu 0.7, dintr-un zgomot alb. Partiționarea în timp este realizată de un detector cu structură impusă maximizată (a), iar raportul semnal/zgomot este arătat în figura (b).

3.4 Optimizarea unui ansamblu de criterii de contrast

Studiul prezentat anterior a permis să demonstrăm teoretic interesul pe care îl prezintă clasa δ a criteriilor de separare care nu depind decât de momentele de ordinul 1 și 2 ale statisticii de detectare, pentru sinteza de detectori lineari generalizați. Alegerea unui criteriu rămâne o problemă, deoarece este tratată destul de puțin în literatura de specialitate, soluțiile utilizate fiind raportul semnal/zgomot sau deflexie.

Metodologia prezentată în acest subcapitol răspunde acestei probleme și permite determinarea celui mai bun criteriu δ , criteriu pentru care detectorul are o probabilitate de eroare minimă.

a) *Căutarea criteriului optimal.* Fie $\Lambda(Z)$ o statistică lineară generalizată definită prin:

$$\Lambda(Z) = a^T Z - v$$

în care Z este un vector dependent de observarea X și elementele sale fiind numere reale .

Problema sintezei unui detector cu structură impusă, așa cum a fost definită în secțiunea anterioară, constă în căutarea lui a și v optimi în sensul unui criteriu de contrast c care nu depinde decât de următoarele momente:

$$\begin{aligned}\eta_1 &= E\{\Lambda(Z)/H_i\} = E\{a^T Z - v/H_i\} = a^T m_i - v; \\ \sigma_i &= \text{Var}\{\Lambda(Z)/H_i\} = a^T E\{(Z - m_i)(Z - m_i)^T / H_i\} a = a^T \sum_i a \\ m_i &= E\{Z/H_i\}, \quad \sum_i = E\{(Z - m_i)(Z - m_i)^T / H_i\}.\end{aligned}$$

În aceste circumstanțe, derivatele lui c în raport cu a și v trebuie să fie nule, adică:

$$\begin{cases} \frac{\partial c}{\partial a} = \frac{\partial c}{\partial \sigma_1^2} \frac{\partial \sigma_1^2}{\partial a} + \frac{\partial c}{\partial \sigma_2^2} \frac{\partial \sigma_2^2}{\partial a} + \frac{\partial c}{\partial \eta_1} \frac{\partial \eta_1}{\partial a} + \frac{\partial c}{\partial \eta_2} \frac{\partial \eta_2}{\partial a} = 0 \\ \frac{\partial c}{\partial v} = \frac{\partial c}{\partial \sigma_1^2} \frac{\partial \sigma_1^2}{\partial v} + \frac{\partial c}{\partial \sigma_2^2} \frac{\partial \sigma_2^2}{\partial v} + \frac{\partial c}{\partial \eta_1} \frac{\partial \eta_1}{\partial v} + \frac{\partial c}{\partial \eta_2} \frac{\partial \eta_2}{\partial v} = 0 \end{cases} \quad (37)$$

unde derivatele parțiale ale lui η_1 și σ_1^2 sunt date de:

$$\frac{\partial \sigma_i^2}{\partial a} = 2 \sum_i a; \quad \frac{\partial \eta_i}{\partial a} = m_i; \quad \frac{\partial \sigma_i^2}{\partial v} = 0; \quad \frac{\partial \eta_i}{\partial v} = -1 \quad (38)$$

Cu ajutorul acestor rezultate putem rescrie sistemului (4.37) astfel:

$$\begin{cases} 2 \left[\frac{\partial c}{\partial \sigma_0^2} \sum_0 + \frac{\partial c}{\partial \sigma_1^2} \sum_1 \right] a = - \left[\frac{\partial c}{\partial \eta_0} M_0 + \frac{\partial c}{\partial \eta_1} M_1 \right] \\ \frac{\partial c}{\partial \eta_0} + \frac{\partial c}{\partial \eta_1} = 0 \end{cases} \quad (39)$$

Introducând cea de a doua ecuație a sistemului (39) în prima ecuație și notând că a poate fi definit ca un coeficient aproape multiplicativ, ajungem la următorul sistem linear:

$$\left[\alpha \sum_0 + (1 - \alpha) \sum_1 \right] a = (m_1 - m_0) \quad (40)$$

unde:

$$\alpha = \frac{\partial c / \partial \sigma_0^2}{\partial c / \partial \sigma_0^2 + \partial c / \partial \sigma_1^2} \quad (41)$$

Aceste rezultate sunt interesante, deoarece arată că alegerea criteriului de contrast nu intervine în relația lui a decât prin intermediul unui parametru α , care este ușor de optimizat. Este de remarcat că acest rezultat este cuprins între 0 și 1 dacă studiul este limitat la criteriile de contrast ale căror sensuri de variație în raport cu σ_0^2 și σ_1^2 sunt identice ($\partial c / \partial \sigma_0^2, \partial c / \partial \sigma_1^2$, ale aceluiași semnal), proprietate care trebuie satisfăcută.

Termenul constant v , depinde în mod explicit de c datorită oblicității celei de a doua ecuații a sistemului (39). Aceasta nu constituie un obstacol pentru procedura de optimizare la care ne-am referit, influența lui v se anulează prin alegerea pragului de detectare. Fiind dat α , acesta din urmă este determinat astfel încât probabilitatea de eroare $P_e(d_\alpha)$ a detectorului d_α următor, să fie minimă:

$$d_\alpha(Z) = \begin{cases} 0 & \text{dacă } a_\alpha^T - \lambda_0 \leq 0 \\ 1 & \text{dacă } a_\alpha^T - \lambda_0 > 0 \end{cases}$$

unde a_α este soluția sistemului (4.40), iar $\lambda_{0,\alpha}$ pragul de determinat.

În cadrul acestei metode, denumită a criteriului optimal, sinteza unui detector se bazează pe optimizarea comună a parametrului α și a pragului $\lambda_{0,\alpha}$, cu scopul de a minimiza probabilitatea de eroare a lui $P_{e,a}$.

În concluzie, această metodologie conduce la detectorul linear generalizat optimal în sensul celui mai bun criteriu de contrast δ .

Algoritmul corespondent este prezentat în *figura 7*. În practică, probabilitatea de eroare P_e nu poate fi evaluată. Atunci trebuie să dispunem de o estimare, care poate furniza, eroarea empirică P_{emp} pentru ansamblul de instruire A_n . Putem utiliza și alte procedee, printre care amintim metodele “bootstrap” sau cea a validării încrucișate [Fuk90], [Dev96].

În aceste condiții, putem scrie:

$d_{n,opt} = \arg \min \hat{P}_e(d_\alpha, A_n)$, unde $\hat{P}_e$ reprezintă o estimare a probabilității de eroare P_e .

Declararea variabilelor

A_n – baza de instruire a n elemente

$\hat{m}_i$ – estimare a speranței lui Z pe H_i

$\sum_i$ – estimarea matricei de variante covariante a lui Z pe H_i

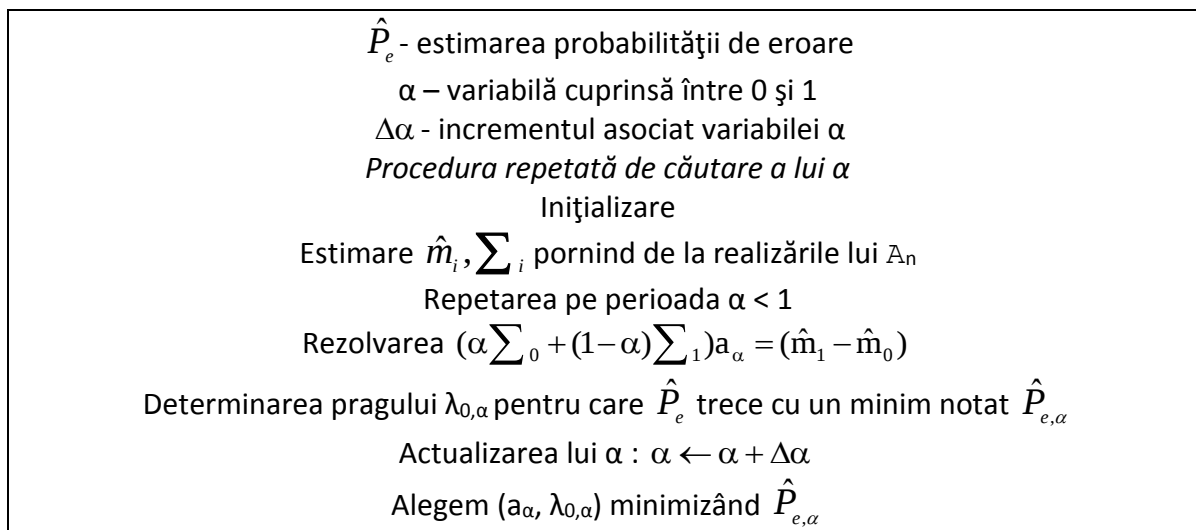


Figura 7. Algoritm pentru determinarea detectorului linear generaliza

b) *Caracteristici.* Detectorul obținut prin metoda criteriului optimal este asemănător ca performanțe cu cei rezultați din maximizarea raportului semnal/zgomot, cu ajutorul deflexiei sau al criteriului Fisher. Aceste proprietăți sunt ilustrate de *figura 8*.

Într-adevăr , C_{RSB} , C_{Deflex} , C_{Fisher} fiind elemente ale lui δ , fiecareia dintre ele îi corespunde o valoare specifică a parametrului α . Astfel, aplicând relația (41) la definiția (24) a raportului semnal/zgomot, ajungem la un rezultat cunoscut:

$$a_{RSB} = \sum_0^{-1} (M_1 - M_0)$$

deoarece $\alpha_{RSB}=1$, presupunând că $\sum_0$ este reversibilă.

În mod analog, se poate stabili, pentru deflexie și pentru criteriul lui Fisher:

$$a_{Deflex} = 2(\sum_0 + \sum_1)^{-1} (M_1 - M_0), \text{ deoarece } \alpha_{Deflex} = 0.5,$$

$$a_{Fisher} = (P_0 \sum_0 + P_1 \sum_1)^{-1} (M_1 - M_0), \text{ pentru } \alpha_{Fisher} = P_0.$$

Aceste relații demonstrează faptul că detectorii obținuți prin maximizarea raportului semnal/zgomot, a deflexiei sau a criteriului Fisher prezintă apriori mai puține avantaje pentru performanța termenilor decât cei rezultați prin metoda criteriului optimal.

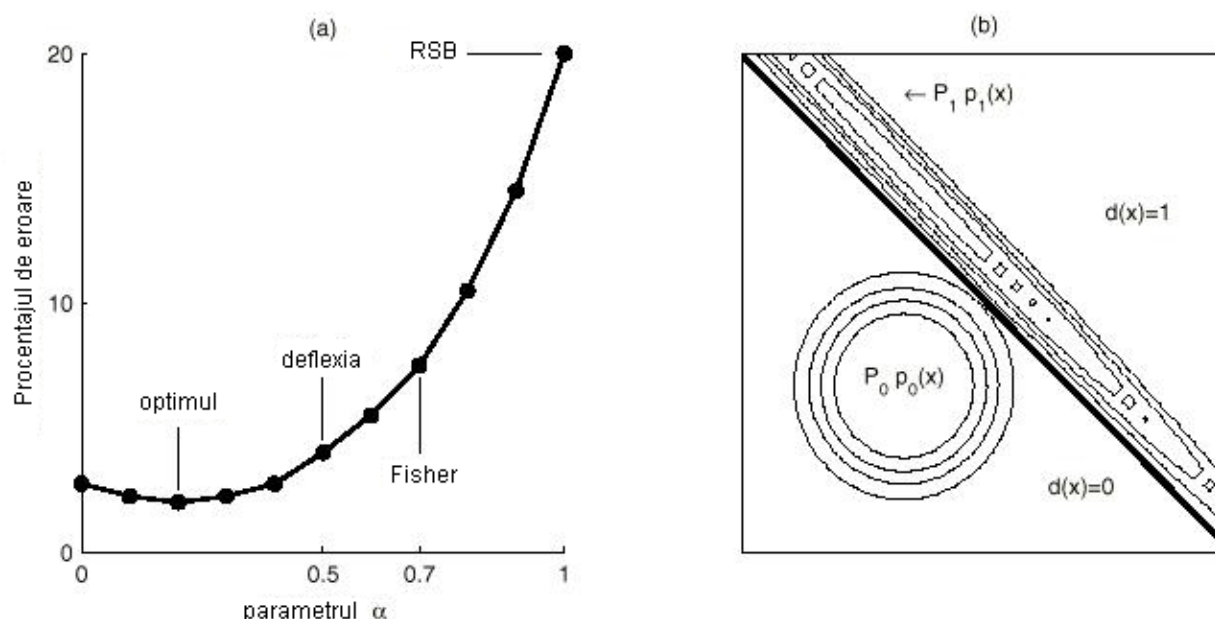


Figura 8. Detectarea unui semnal gaussian cu probabilitatea apriori P_0 egală cu 0.7 dintr-un zgomot alb. (a) Metoda criteriului optimal, care determină parametrul α minimizat cu probabilitatea de eroare și care conduce la un detector performant față de rezultatul obținut prin aplicarea criteriului clasic de maxim. (b) Partiționarea spațiului pentru detectorul optimal prezentat în figura 6.

Obiectivul următoarei secțiunii este acela de a dezvolta metoda expusă anterior, prin rezolvarea mai multor probleme de detectare cu structură impusă în planul timp-frecvență. Acesta din urmă constituie un spațiu de reprezentare, care ne va permite să realizăm o analiză calitativă asupra configurației structurilor de detectare obținute.

3.5 Experimentări în planul timp-frecvență

Sinteza de detectori prin maximizarea unui criteriu de contrast cu un ansamblu de instruire este o metodă nouă în domeniul de analiză timp-frecvență [Jon95b]. Dacă sunt îndepărtate soluțiile bazate pe extragerea atributelor de reprezentare (poziția unui vârf, încărcătura timp-frecvență a unei componente), majoritatea lucrărilor descriu detectorii lineari care operează cu reprezentarea Wigner-Ville a semnalului observat [Jon95b], [Say96b], [Sam96], [Ric97c], [Dum02]. În general, această teorie este motivată din dorința de a determina un test optimal, atunci când soluționarea problemei face

apel la o statistică de detectare cvadratică [Pic88]. Din acest motiv, într-o primă etapă, ne vom concentrat asupra ansamblului determinat de regulile de decizie definite prin relația (42), pentru a determina criteriul optimal:

$$d(X) = \begin{cases} 0; & (a_{TF})^T W_x - \lambda_0 \leq 0 \\ 1; & (a_{TF})^T W_x - \lambda_0 > 0 \end{cases} \quad (42)$$

unde a_{TF} și λ_0 trebuie să fie determinați, pornind de la un ansamblu de realizări ale ipotezelor care intervin. W_x reprezintă reprezentarea Wigner-Ville a observării X , scrisă sub formă vectorială.

a) *Primul experiment.* Prima problemă pe care o propunem se referă la detectarea unui semnal aleator S înecat într-un zgomot alb, gaussian, centrat și distribuit identic pe ipotezele H_0 și H_1 :

$$\begin{cases} H_0 : X = B \\ H_1 : X = S + B \end{cases} \quad (43)$$

în care $S(k) = k \exp(-0.45k) \sin(0.5k\pi + \theta)$, $k \in \{1, \dots, 16\}$, θ este o variabilă aleatoare, distribuită în mod uniform asupra intervalului $[0, 2\pi]$. Variația lui B este fixată în așa fel încât raportul semnal/zgomot să fie egal cu -3dB. În aceste condiții, în secțiunea 2.3 se arată că testul de detectare optimală este obținut pentru $a_{TF} = W_s$, unde W_s desemnează reprezentarea Wigner-Ville a semnalului de detectat. Reamintim faptul că această structură este numită filtru adaptat timp-frecvență.

Pentru caracterizarea eficienței metodei criteriului optimal pentru acest tip de ipoteză, instruirea detectorului (42) a fost efectuată pe un ansamblu de 2000 de realizări ale fiecăreia dintre ipotezele H_0 și H_1 , transformate într-o primă fază în semnale analitice.

În *figura 9* se folosește ca referința timp-frecvență $\hat{a}_{TF}$, care este asemănătoare reprezentării Wigner-Ville a semnalului S , rezultat care a fost confirmat anterior. *Figura 10* prezintă curbele coeficienților de restituție (COR) ale acestei structuri și ale filtrului timp-frecvență, acesta din urmă furnizând o formă superioară a termenilor realizați.

Compararea acestor caracteristici confirmă raționamentul calitativ făcut privind referința $\hat{a}_{TF}$ și ilustrează eficacitatea metodei criteriului optimal pentru rezolvarea problemei detectorului considerat.

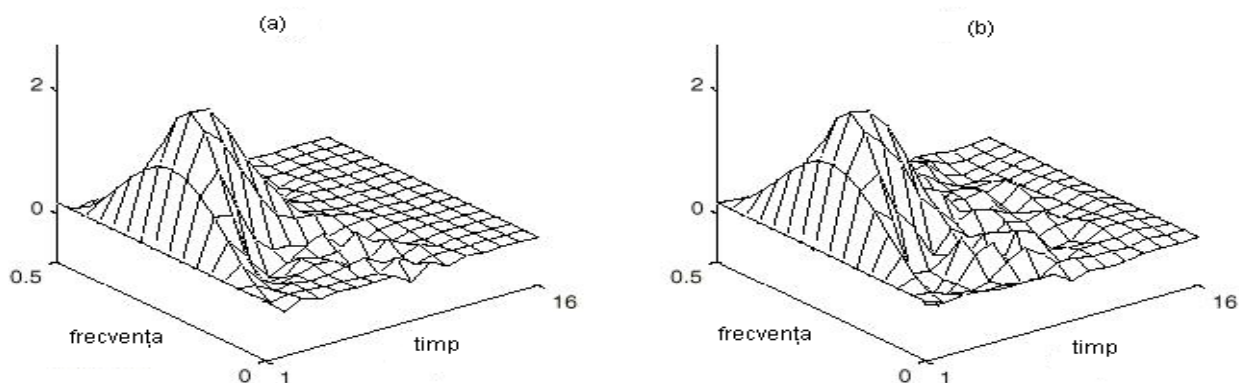


Figura 9. Detectarea unui semnal aleator suprapus unui zgomot alb, gaussian, centrat și distribuit. (a) Reprezentarea Wigner-Ville a semnalului. (b) Configurația detectorului linear care lucrează în planul timp-frecvență pe ipotezele H_0 și H_1 .

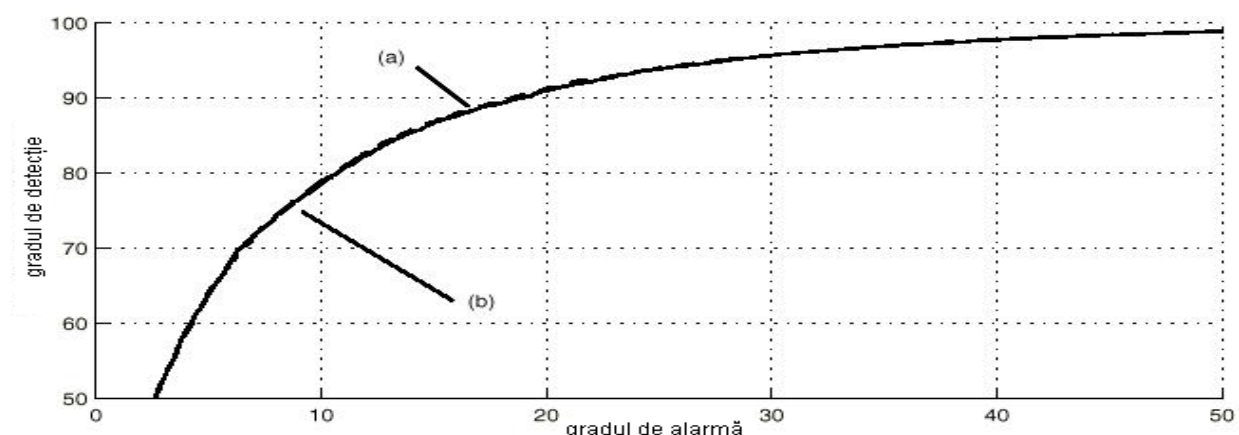


Figura 10. Detectarea unui semnal aleator suprapus unui zgomot alb, gaussian. (a) Curbele COR ale unui filtru adaptat în planul timp-frecvență. (b) Curbele COR ale unui detector linear în planul timp-frecvență, obținute pe baza criteriului optimal.

b) *Al doilea experiment.* A doua problemă studiată este aceea a detectării unui semnal determinat S înecat într-un zgomot alb, gaussian și centrat, a cărui variație depinde de ipoteza H_0 și H_1 considerată:

$$\begin{cases} H_0 : X = \gamma B \\ H_1 : X = S + B \end{cases} \quad (44)$$

unde γ este o realitate dată și $s(k) = k \exp(-0.45k) \sin(0.5k\pi)$, $k \in \{1, \dots, 16\}$. Ca și pentru experimentul precedent, variația lui B este fixată în așa fel încât raportul semnal/zgomot să fie egal cu -3dB .

Scrierea raportului de credibilitate pentru ipoteza (4.44) conduce la următoarea relație a statisticii de detectare [Poo94]:

$$\Lambda(X) = \frac{1-\gamma^2}{2\gamma^2} X^T X + s^T X \quad (45)$$

care trebuie comparat cu un prag λ_0 pentru a putea lua o decizie. Pentru ușurarea analizei ulterioare a statisticii obținute prin metoda criteriului optimal, încă de la început se alege o adaptare pentru formularea timp-frecvență a termenul cvadratic din relației (45). Adoptând o scriere vectorială pentru reprezentarea Wigner-Ville a unui semnal X compus din L eșantioane, separate, se observă că:

$$\frac{1}{L} \sum_{k=1}^{L^2} W_x[k] = \|X\|^2 \quad (46)$$

ceea ce conduce la:

$$\Lambda(X) = \frac{1-\gamma^2}{2L\gamma^2} \sum_{k=1}^{L^2} W_x[k] + \sum_{k=1}^L s[k]X[k] \quad (47)$$

Statistica de detectare definită prin relația (4.45) poate fi rescrisă sub forma:

$$\Lambda(X) = (a_{TF})^T W_x + (a_T)^T X.$$

În acest caz, elementele componenteii timp-frecvență a_{TF} din $\Lambda(X)$ sunt identic egale cu $(1-\gamma^2)/2L\gamma^2$, componenta temporală a_T corespunzând unui semnal δ detectat.

Metoda criteriului optimal a fost folosită pentru optimizarea următorului test de detectare:

$$d(X) = \begin{cases} 0; & (\hat{a}_{TF})^T W_x + (\hat{a}_T)^T X - \lambda_0 \leq 0 \\ 1 & \text{idem} \end{cases} \quad (48)$$

pornind de la un ansamblu de instruire de 2000 realizări pentru fiecare dintre ipotezele problemei (4.44), γ fiind fixat la 4. Referințele temporale $\hat{a}_T$ și timp-frecvență $\hat{a}_{TF}$, rezultat al procedurii de învățare, sunt arătate în *figura 11*.

Compararea curbelor COR reprezentate în *figura 12* confirmă această analiză și ilustrează eficacitatea criteriului optimal.

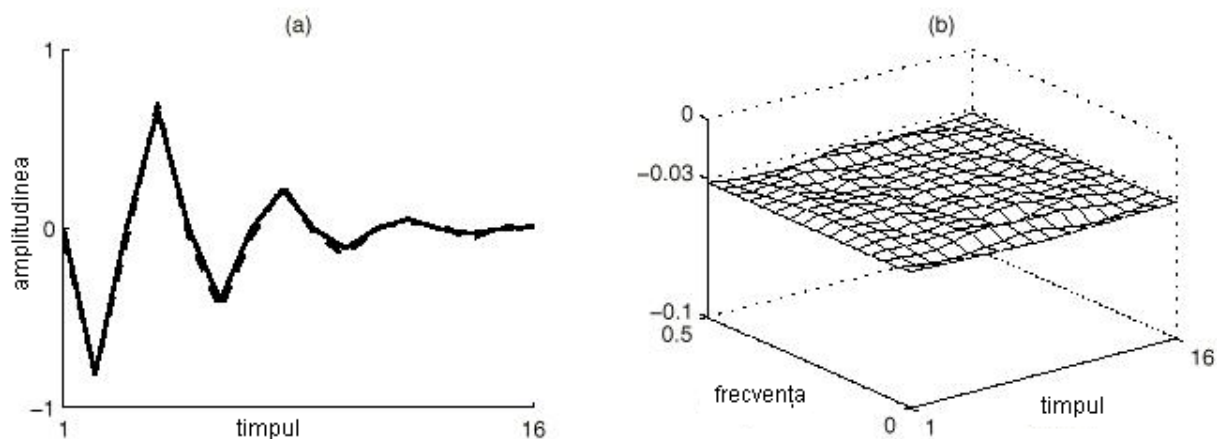


Figura 11. Detectarea unui semnal determinat suprapus unui zgomot alb, gaussian și centrat, a cărei variație depinde de ipoteza H_0 sau H_1 . (a) Compararea semnalului detectat (linie continuă) cu componenta temporală a_T a detectorului obținut prin metoda criteriului optimal (linie punctată). (b) Configurarea componentelor timp-frecvență a_{TF} ale acestui detector.

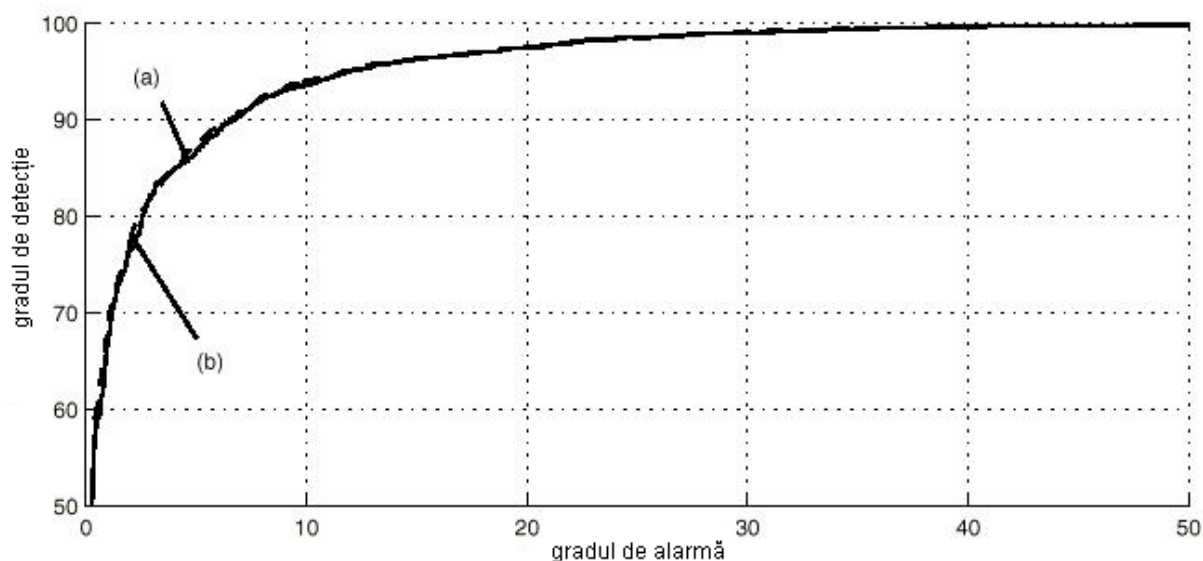


Figura 12. Detectarea unui semnal determinat suprapus unui zgomot alb, gaussian și centrat, a cărei variație depinde de ipoteza H_0 sau H_1 . (a) Curba COR a detectorului optimal conform raportului de credibilitate. (b) Curba COR a detectorului obținut prin metoda criteriului optimal.

3.6 Concluzii

Pentru rezolvarea unei probleme de detectare, pot fi avute în vedere cel puțin două metode.

Prima metodă se referă la detectarea cu structură liberă, pentru care relația testului de detectare rezultă din aplicarea unui criteriu de decizie și din cunoașterea legilor de probabilitate ale observării. Cea de a doua metodă, numită detectare cu structură impusă, necesită definirea în prealabil a unei clase $\mathcal{D}$ de teste de detectare, înainte de a determina testul optim pentru un criteriu dat.

Acest livrabil a fost consacrat, definiri unei metodologii pentru elaborarea de detectori cu structură impusă. În special, s-a arătat că există cel puțin două argumente care pledează pentru o alegere în favoarea clasei $\mathcal{D}$ a detectorilor polinomiali definiți prin enunțul 4.

Primul argument se bazează pe garanțiile de performanță pe care le reprezintă consistența acestei clase. Al doilea argument, mai practic, se bazează pe căutarea unui detector în interiorul acestei clase care echivalează cu optimizarea unui discriminant linear, metodă utilizată în analiza recunoașterii formelor.

După ce s-a demonstrat teoretic interesul pe care îl prezintă, clasa δ a criteriilor de separare, care nu depind decât de momente de ordinul 1 și 2 ale statisticii de detectare, utilizate pentru sinteza unui detector linear generalizat, constatările experimentale au arătat importanța alegerii acestei clase.

Metoda de învățare prezentată în cadrul acestui referat, a permis să se răspundă la această alegere, conducând la determinarea celui mai bun criteriu δ pentru problema analizată, adică obținerea unui detector cu o probabilitate de eroare minimă.

Rezultatele experimentale propuse confirmă, pe de o parte, teoria algoritmilor propuși, în timp ce mărirea ansamblului de învățare este importantă în raport cu gradul statisticii de detectare și creșterea stabilității sistemului în condiții de zgomot.

BIBLIOGRAFIE

- [Ami96] M. G. AMIN. Recursive Kernels for Time-Frequency Signal Representation. *IEEE Signal Processing Letter*, 1996, pg. 16-18.
- [Aug91] F. AUGER. Representation Temps-Frequence des Signaux Non-Stationnaires. *Synthese et Contribution*. 1991, pg. 221.
- [Aug91] F. AUGER. Representation Temps-Frequence des Signaux Non-Stationnaires. *Synthese et Contribution*. 1991, pg. 221.
- [Coh95] L. COHEN. Time-Frequency Analysis. *Englewood Cliffs, Prentice Hall*, 1995.
- [Cun92] G. CUNNINGHAM. Fast Computation of the Wigner Distribution for Finite Length Signal. *IEEE International Conference Signal Processing*, 1992, pg. 177-180.
- [Dev96] L. DEVROYE, G. LUGOSI. A Probabilistic Theory of Pattern Recognition. *New York, Verlage*, 1996.
- [Don95] R. N. DONOUGH, A. WHALEN. Detection of Signals in Noise. *London, Academic Press*, 1995.
- [Dum01] C. DUMITRESCU. Analiza EEG utilizând metoda Matching Pursuit și funcția Gabor. A *XXIX Sesiune de comunicări tehnico-științifice a Institutului pentru Tehnologii Avansate*, București 2001.
- [Dum02] C. DUMITRESCU. Dezvoltarea și implementarea algoritmului DTDF pentru analiza de semnal utilizând reprezentarea în planul timp-frecvență. A *XXX Sesiune de comunicări tehnico-științifice a Institutului pentru Tehnologii Avansate*, București 2002.
- [Dum02a] C. DUMITRESCU. Analiza semnalelor cu ajutorul algoritmului Matching Pursuit și reprezentarea Wigner-Ville. *Raport de activitate, vol. 1, nr. 2, pg.205-220*, 2002.
- [Fla93] P. FLANDRIN. Temps-Frequence. *Paris, Hermes*, 1993.
- [Fuk92] K. FUKUNAGA. Introduction to Statistical Pattern Recognition. *Academic Press*, 1990.
- [Gon98] P. GONCALVEZ, E. PAYOT. Adaptive Equation for Time-Frequency Representation. *Bryce Canyon, Utah, USA*, 1998.
- [Loo35] A. LOOMIS, E. HARVEY. Potentials Rhythms of the Cerebral Cortex during Sleep. *Science*, 1935, pg. 597-598.

- [Lwh92] J. LU, J. WEAVER, D. HEALY. Noise Reduction with Multiscale Edge Representation and Perceptual Criteria, *Proceedings of IEEE-SP International Symposium on Time-Frequency and TimeScale Analysis, Victoria, B.C., pp. 555—558, 1992.*
- [Mar95] W. MARTIN, P. FLANDRIN. Wigner-Ville Spectral Analysis of Nonstationary Processes. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1985, vol. 33, pg. 1464-1469.
- [Mal 63] S. MALLAT, Z. ZHANG. *Matching pursuits with time-frequency dictionaries. IEEE Trans. on Signal Process.*, 12(41), pp. 3397-3415, 1993.
- [Ric97c] C. RICHARD, R. LENGELLE. Joint Time and Time-Frequency Optimal Detection. *IEEE International Symposium on Time-Frequency and Time-Scale Analysis*, 1997.
- [Ric98a] C. RICHARD, R. LENGELLE. Structural Risk Minimization for Reduced-Bias Time-Frequency-Based Detectors Design. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1998, pg. 2397-2400.
- [Sch91] L. SCHARF. Statistical Signal Processing. *USA, Addison-Wesley*, 1991.
- [Skl79] J. SKLANSKY, G. WASSEL. Pattern Classifiers and Trainable Machine. *New-York, Springer-Verlag*, 1979.
- [Vap82] V. VAPNIK. Estimation of Dependences Based on Empirical Data. *New-York, Springer-Verlag*, 1982.
- [Vas91] V. VASEGHI. Advanced Signal Processing. *Wiley* 1998.
- [Vil48] J. VILLE. Theorie et Application de la Notion de Signal Analytique. *Cables et Transmissions*, 1948, pg. 61-74.
- [Was72] G. WASSEL, J. SKLANSKY. Training a one-dimensional Classifier to Minimize the Probability of Error. *Transactions on Systems, Man and Cybernetics*, 1972, pg. 533-541.
- [Wid88] B. WIDROW, M. HOFF. Adaptive Switching Circuits. *Neurocomputing Foundations of Research. Cambridge, MIT Press*, 1988.
- [Wig32] E. WIGNER. On the Quantum Correction for Thermodynamic Equilibrium. *Physical Review*, 1932, vol. 40, pg. 749-759.

- [W91] M. WICKERHAUSER. Lectures on Wavelet Packet Algorithms, *INRIA, Roquencourt, France, Minicourse lecture notes*, 1991
- [Wb98] D. WEI, A. BOVIK. Enhancement of Compressed Images by Optimal Shift-Invariant Wavelet Packet Basis, *J. Visual Commun. Image Represen., Special Issue on High-Fidelity Media Processing*, vol. 9, no. 1, pp. 15-24, Mar. 1998.
- [Woog96] L.J. WOOG. Adapted Waveforms Algorithms for Denoising, Teză de doctorat, *Yale University*, 1996.
- [Zie91] T. ZIELINSKI. Note on Computation of some Bilinear Time-Frequency Signal Representation. *International Conference on Digital Signal Processing*, 1991, pg. 40-44.