

Introduction

On donne ici quelques motivations et définitions pour la segmentation de signaux. On tente ensuite de dégager des éléments de structure commune d'algorithmes de segmentation, qui seront repris dans les diverses fiches. Et on termine en situant brièvement ce document par rapport à l'état de l'art.

1. Introduction à la segmentation de signaux

La **segmentation** est une approche pour le traitement des signaux présentant des non stationnarités « rapides » ou « ruptures », et constitue une première étape possible de traitement en vue de la reconnaissance ou du diagnostic, voire même du codage. Cette approche consiste à considérer que le signal est une succession de segments homogènes, de caractéristiques constantes ou lentement variables, séparés par des transitions brutales où les caractéristiques du signal changent rapidement. Dans ce document, on utilisera indifféremment les termes de « rupture », « saut », ou « changement (brusque) ».

Les algorithmes de segmentation de signaux, qui utilisent essentiellement des modèles paramétriques, sont des algorithmes qui effectuent une telle décomposition d'un signal en segments successifs, et ce avec un degré de précision variable. Certains se préoccupent uniquement de détecter — évidemment avec un retard — qu'un changement s'est produit ; d'autres raffinent la position de la frontière de segment trouvée en estimant, après détection, l'instant auquel la rupture s'est produite ; d'autres enfin vont jusqu'à estimer les paramètres caractéristiques du signal avant et après la rupture. Le formalisme sous-jacent à la construction de ces algorithmes est du type : « *modèle paramétrique où les valeurs des paramètres changent instantanément à un instant inconnu* ». On exclut intentionnellement ici la détection de changement d'ordre, de changement de type de distribution, ...

Cependant, il est important de remarquer que les outils permettant de détecter de telles ruptures — comme les débuts et fins de voisement et de fricatives en parole — permettent aussi de détecter en pratique des « sauts » se produisant non pas instantanément mais en plusieurs échantillons, voire des transitions moins « franches » (telles que les débuts ou les ruptures de pente de certains formants en parole), ou des événements plus fugitifs (tels que certaines ondes particulières dans les signaux biomédicaux ou sismiques). Pour cette raison, et à des fins de comparaison de performances, sont également inclus dans ce document trois algorithmes qui sont plutôt du type *détecteurs de changement de forme* dans un signal que du type détecteur de changement de modèle, mais qui permettent de détecter ce dernier type d'événements.

Enfin, un autre rôle important joué par les détecteurs de ruptures est celui de l'amélioration de la *capacité de poursuite des algorithmes adaptatifs* qui, par essence, ne sont capables de suivre que des variations lentes de

paramètres, et dont on peut améliorer les performances en présence de non stationnarités rapides en leur adjoignant un détecteur de telles ruptures et une procédure de réactualisation de l'algorithme adaptatif après détection.

L'intuition du praticien du traitement de signal conduit à distinguer, dans ce contexte paramétrique, **deux types de ruptures : additives et spectrales**. Il est en effet visuellement évident que les « sauts » ou changements brusques de la valeur moyenne locale dans un signal sont qualitativement différents des changements dans le « degré d'agitation locale » autour d'une moyenne. De plus, on peut caractériser ces deux types de ruptures par leurs effets différents sur une représentation par modèle d'état du signal observé $(Y_k)_k$:

$$\begin{cases} X_{k+1} = FX_k + W_k \\ Y_k = HX_k + V_k \end{cases} \quad (1.1)$$

où X est de dimension n et le signal Y est de dimension m , le plus souvent égale à 1 dans ce document, et où $(W_k)_k$ et $(V_k)_k$ sont deux bruits blancs gaussiens indépendants, de matrices de covariance respectives Q et R . Soit :

$$\mathcal{K}(z) = H(zI - F)^{-1} \quad (1.2)$$

la fonction de transfert de ce système.

Les deux types de ruptures, se produisant à un instant inconnu r , peuvent être caractérisée formellement de la manière suivante :

1. Changements additifs : ces changements sont additifs soit sur l'équation de transition d'état soit sur l'équation d'observation, et sont modélisés soit par :

$$\begin{cases} X_{k+1} = FX_k + W_k + \nu_x \delta_{r, k+1} \\ Y_k = HX_k + V_k \end{cases} \quad (1.3)$$

soit par :

$$\begin{cases} X_{k+1} = FX_k + W_k \\ Y_k = HX_k + V_k + \nu_y \mathbf{1}_{\{k \geq r\}} \end{cases} \quad (1.4)$$

où le symbole de Kronecker $\delta_{r, k}$ représente une impulsion à l'instant r , $\mathbf{1}_{\{k \geq r\}}$ représente un échelon unité à l'instant r , et où ν_x et ν_y sont les vecteurs représentant les « amplitudes » des changements. Il est à noter que le premier type de changement, dans (1.3), se traduit sur l'observation Y par la réponse impulsionnelle du système déterministe (H, F) à l'impulsion qui se produit sur l'état, alors que le deuxième, dans (1.4), se traduit sur Y par un échelon. On peut schématiser ces changements additifs à l'aide de la figure 1.

2. Changements de dynamique (ou spectraux ou encore sauts multiplicatifs) : on en distingue deux catégories ; d'une part, les changements sur les bruits W et V et les changements de la matrice d'observation H , d'autre part les changements de F .

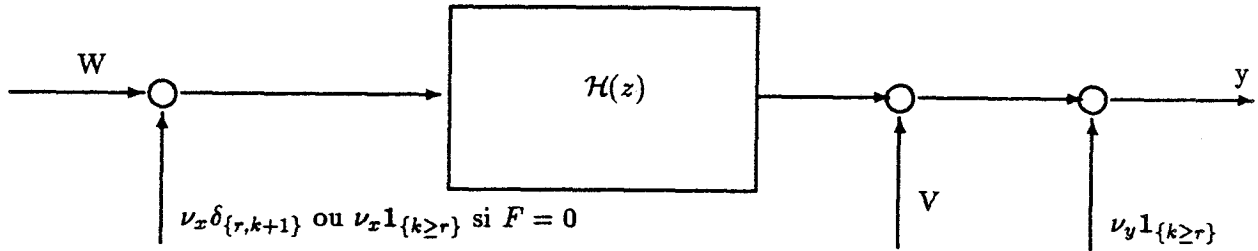


Figure 1. — Changements additifs.

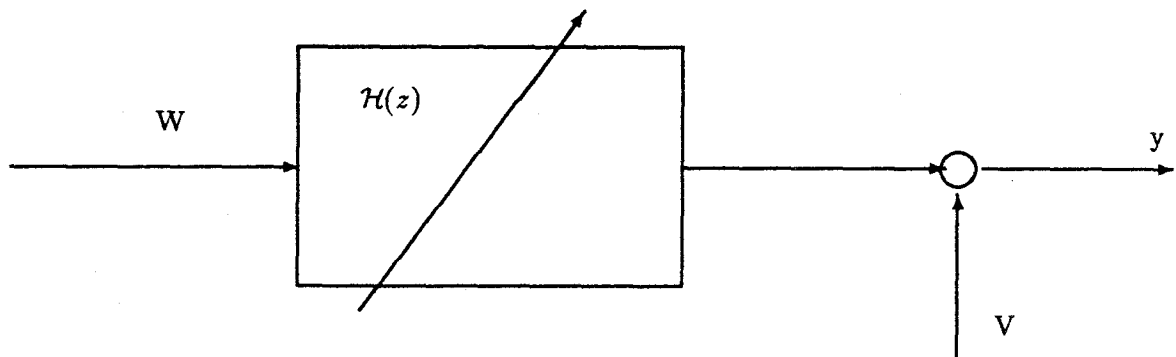


Figure 2. — Changements de dynamique.

Dans la première catégorie, nous rangeons les ruptures modélisables soit par un changement de la covariance des bruits V_k et/ou W_k , soit par un changement de la matrice d'observation (ou de mesure) H , c'est-à-dire par :

$$\begin{cases} X_{k+1} = FX_k + W_k + \mathbf{1}_{\{k+1 \geq r\}} A_w W_k \\ Y_k = HX_k + V_k + \mathbf{1}_{\{k \geq r\}} A_v V_k \end{cases} \quad (1.5)$$

ou par :

$$\begin{cases} X_{k+1} = FX_k + W_k \\ Y_k = (H + \mathbf{1}_{\{k \geq r\}} \Delta H) X_k + V_k \end{cases} \quad (1.6)$$

Lorsque A_v et A_w sont diagonales, les bruits ne subissent qu'un changement d'énergie.

Remarquons que les types de ruptures mentionnés jusqu'ici correspondent en Automatique, à des *pannes de capteurs ou d'actionneurs*, par opposition au deuxième type multiplicatif ci-dessous lié plutôt à des *pannes de composantes du système* proprement dit.

La deuxième catégorie de changements multiplicatifs peut être modélisée par des changements dans la matrice de transition d'état F du modèle d'état ci-dessus, de la façon suivante :

$$\begin{cases} X_{k+1} = (F + \mathbf{1}_{\{k+1 \geq r\}} \Delta F) X_k + W_k \\ Y_k = HX_k + V_k \end{cases} \quad (1.7)$$

Réécrivons la fonction de transfert (1.2) sous la forme :

$$\mathcal{H}(z) = H(zI - F)^{-1} G \quad (1.8)$$

où $G = I$, sauf dans le cas (1.5) où $G = I + A_w$. Les deux

catégories de changements de dynamique peuvent alors être schématisés à l'aide de la figure 2.

Pour un signal $(Y_k)_k$ scalaire (c'est-à-dire $m = 1$), on a également la représentation ARMA :

$$Y_k = \sum_{i=1}^p a_i Y_{k-i} + E_k + \sum_{j=1}^q b_j E_{k-j} \quad (1.9)$$

où (E_k) est un bruit blanc gaussien de variance $R = \sigma^2$, et les a_i et les b_j sont solutions de :

$$HF^p - \sum_{i=1}^p a_i HF^{p-i} = 0$$

$$b_j = H(F^j + a_1 F^{j-1} + \dots + a_j I) G \quad (1.10)$$

et où

$$W_k = GE_k.$$

Alors, l'étude des changements (1.7) revient — à un changement de base près — à l'étude des changements des paramètres de la partie AR, et les changements (1.5) induisent des changements de la partie MA seule, ainsi que le montrent les relations (1.10).

2. Cadre général pour la structure des algorithmes présentés dans ce document

Conformément à la discussion précédente, on convient de désigner les algorithmes présentés dans le présent document par un :

- **M** pour les détecteurs de sauts de moyenne ;

- **D** pour les détecteurs de sauts de dynamique ;
- **F** pour les détecteurs de changements de forme.

2.1 Les algorithmes présentés

La liste des algorithmes présentés dans ce document est donnée dans le **tableau 1**. Ce tableau précise également les laboratoires coordonnateurs X et rédacteurs x pour chaque fiche.

Tous ces algorithmes traitent en ligne des signaux scalaires, à l'exception de GLR qui peut également traiter des signaux multi-capteurs.

Les *exemples d'application* donnés dans chaque fiche sont ceux sur lesquels l'algorithme concerné a été *effectivement* utilisé par le concepteur et/ou les rédacteurs de la fiche.

2.2 Conception d'algorithmes de segmentation

Commençons par quelques remarques sur les fils directeurs qui conduisent à la conception des algorithmes de segmentation de signaux, qu'elle soit ou non fondée sur la modélisation présentée plus haut.

Sur le plan des intuitions, il paraît naturel à beaucoup de chercher à détecter un changement en surveillant de manière adéquate l'**innovation ε_n d'un filtre** — qu'il soit ajusté localement une fois pour toutes ou réactualisé de temps en temps ou en permanence —, en testant :

- un changement de sa moyenne (algorithmes GLR, GLRmod, EPL, DAREP),
- un changement de sa variance (algorithmes FKE-M, SNR, DAREP, DIS2f, Brandt),
- une perte de sa blancheur (algorithme FKE-D).

Une autre idée intuitivement naturelle consiste à **comparer convenablement deux grandeurs estimées respectivement à long terme (ou globalement) et à court terme (ou localement)** sur le signal ; ces quantités comparées peuvent être :

- une énergie résiduelle (algorithmes FKE-M, SNR, DIS2f, Brandt),
- une corrélation résiduelle (algorithme FKE-D),
- des innovations (algorithme DIV).

Du point de vue méthodologique, l'outil de base est le **rapport de vraisemblance généralisé** classique en détection et dans lequel les paramètres inconnus (instant de rupture, paramètres du modèle avant et après rupture) sont estimés par le maximum de vraisemblance. Les algorithmes de ce document qui utilisent cet outil sont tous conçus sous hypothèse gaussienne. Les autres ingrédients possibles sont, au niveau du prétraitement, des intégrateurs ou des opérateurs de dérivation, et, au niveau de la règle de décision, des tests de blancheur ou des mesures de distances spectrales.

2.3 Structure des algorithmes de type M

Ces algorithmes effectuent un premier traitement du signal et travaillent ensuite sur :

1. soit la sortie d'un filtre (passe-bas pour l'algorithme DF),
2. soit la sortie d'un intégrateur (HK),

3. soit l'innovation ε_n d'un filtre (algorithmes EPL, GLR, GLRmod, FKE-M) et leur appliquent, en vue de la détection :

1. un seuillage ((EPL), éventuellement précédé d'une dérivation (DF), ou bien adaptatif (HK), ou bien joint à un compteur (DF)),
2. un double calcul d'énergies résiduelles à long et court terme (FKE-M),
3. une corrélation avec une signature de rupture supposée (GLR et GLRmod).

La détection est suivie d'au moins l'une des actions suivantes :

1. estimation de l'instant de rupture (HK, GLR, GLRmod) ; à défaut, l'instant de rupture estimé est l'instant de détection ;
2. estimation de l'amplitude de la rupture détectée (GLR, GLRmod, FKE-M) ;
3. réactualisation du filtre à l'aide de l'amplitude de saut estimée (GLR, GLRmod, FKE-M) ;
4. redémarrage de l'algorithme de détection : soit immédiatement après la détection si on a un moyen de réactualiser (GLR, GLRmod) ; soit après un temps suffisant pour permettre au filtre de « re-converger », c'est-à-dire au minimum de pouvoir être calculé (DF, HK, FKE-M).

2.4 Structure des algorithmes de type D

Ces algorithmes travaillent essentiellement avec une modélisation paramétrique de type (1.9) autorégressive AR ou autorégressive AR ou autorégressive à moyenne glissante ARMA. Ils effectuent une estimation de tel(s) filtres ajustés à court et/ou long terme selon les cas, sur une ou plusieurs fenêtres de type :

1. fixe (SNR, EPL, DIS2f),
2. de taille fixe glissant de manière continue (DIS2f, DIV, Brandt) ou discontinue (DAREP, DIS2f),
3. croissante (FKE-D, DIV, Brandt)

et calculent :

1. l'innovation ou erreur de prédiction :

$$\varepsilon_n = Y_n - \sum_{i=1}^p \hat{a}_i Y_{n-i} - \sum_{j=1}^q \hat{b}_j \varepsilon_{n-j} \quad (1.11)$$

2. l'énergie résiduelle $\hat{\sigma}^2 = \text{var}(\varepsilon)$ qui sont ensuite testées et/ou comparées à leurs homologues, à l'aide d'une mesure de distance spectrale convenable ; ce qui donne la détection.

Cette détection est éventuellement suivie d'une *estimation* plus fine de l'*instant de rupture* (DIV, Brandt, FKE-D). Un moyen très simple d'y parvenir consiste à remarquer que la « distance » réagit à un changement dans le signal par un changement de sa dérive qui, de nulle devient positive. Il suffit alors d'adjoindre un détecteur CUSUM (voir fiche HK) qui détecte ce changement de dérive et en estime l'instant d'occurrence. La détection est également souvent suivie d'une procédure de redémarrage après détection (en général, temps d'attente de l'ordre de la taille de la fenêtre pour permettre à au moins un des filtres de « converger »).

On peut définir des variantes de ces algorithmes en modifiant :

- le nombre des fenêtres,
- la position des fenêtres,
- le choix des méthodes d'identification à l'intérieur de chacune des fenêtres (car elles peuvent influencer le comportement de la règle de décision),
- le choix de la mesure de distance entre les modèles ainsi estimés.

2.5. Structure des algorithmes de type F

Les algorithmes de type F ont pour objectif de détecter une rupture de forme dans une suite de signaux déterministes éventuellement présegmentés.

Par définition, la forme d'un profil est invariante par translation et changement d'échelle sur les abscisses et les ordonnées. Le terme de « forme » est parfois confondu avec celui de « gabarit » où les paramètres d'échelle, et en particulier de largeur, sont fixés. Les algorithmes de type F peuvent être paramétriques : ils dépendent alors d'hypothèses *a priori* et n'utilisent pas toute l'information contenue dans la forme — voir par exemple la fiche GAB.

Plus couramment, les algorithmes de type F sont non paramétriques : ils sont conçus initialement soit pour détecter un changement de gabarit (fiche MFC), soit pour détecter directement un changement de forme (fiche MFR). Cependant, une reconnaissance de gabarit peut devenir une reconnaissance de forme en adjoignant un algorithme d'ajustement des paramètres d'échelle. De la même façon, une reconnaissance de forme peut être restreinte à une reconnaissance de gabarit si l'on fixe les paramètres d'échelle.

3. Le contexte

Ce document présente essentiellement le savoir-faire du GDR 0134 « Traitement du Signal et Images » dans le domaine de la segmentation de signaux. Il se situe dans le contexte des multiples travaux sur la détection de ruptures de modèles menés dans des domaines très divers.

Le lecteur intéressé par l'état de l'art sur ce sujet du point de vue de la statistique mathématique pourra consulter [Desh83], [Nikif83], [Zacks83], [Telk87], [Krish88], ainsi que [Shir78] pour des outils méthodologiques et [Bass82] pour des références supplémentaires.

Le point de vue adopté dans le présent document, et que l'on peut qualifier de « mixte » entre Traitement du Signal et Automatique, est décrit en [Will76] et [Bass88], ainsi que dans l'ouvrage [LNCIS] qui inclut de plus les résultats d'optimalité statistique les plus généraux existant actuellement et tirés de [Desh83]. Un livre [BasNik] unifiant ces trois points de vue est à paraître. En ce qui concerne le Traitement du Signal proprement dit, dans le livre [Cohen87] des algorithmes de segmentation sont présentés comme des outils de base pour le traitement de signaux biomédicaux ; [André88] et [DiFra90] montrent l'intérêt de la segmentation automatique en traitement de la parole continue, pour la reconnaissance et le codage respectivement, et [Nikif86] et [Nikif89] pour le traitement des

signaux sismiques. Le point de vue de la détection de pannes plus proche de l'Automatique est synthétisé — outre [Will76], [Bass88] et le livre [LNCIS] cités plus haut — en [Frank90], [Iserm84] et dans l'ouvrage collectif [Patt89]. Une méthode générale et systématique pour concevoir un algorithme de détection de changements — ou de validation de modèle — associé à tout algorithme récursif d'identification paramétrique, est décrite en [Ben-v87]. Enfin, le problème de l'utilisation d'un algorithme de détection de ruptures pour améliorer la capacité de poursuite d'un algorithme adaptatif est traité par exemple en [Fav84], [Wahn90] et dans l'article de synthèse [Ljung90].

De nombreux travaux existent aussi dans le domaine de l'économétrie comme en témoignent le numéro spécial [Broem82], et les livres [Broem87] et [Poir76].

Nota : Chaque fiche donne les points d'entrée bibliographiques préférentiels pour l'algorithme correspondant. Sont données ici des références d'ordre général, soit sur le problème de la détection de rupture lui-même, soit sur l'intérêt de la segmentation comme première étape de traitement de signaux en vue de la reconnaissance et du codage, soit sur le problème connexe de la détection de pannes.

Ces références générales sont dotées d'une étiquette nominative, contrairement à celles des fiches qui sont numérotées.

Bibliographie

- [André88] R. ANDRÉ-OBRECHT (1988). A new statistical approach for the Automatic segmentation of continuous speech signals. *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP-36, n° 1, pp. 29-40.
- [Bass82] M. BASSEVILLE (1982). A survey of statistical failure detection techniques. In *Contribution à la détection séquentielle de ruptures de modèles statistiques*. Thèse d'État, Université de Rennes 1, France.
- [LNCIS] M. BASSEVILLE and A. BENVENISTE, eds. (1986). *Detection of Abrupt Changes in Signals and Dynamical Systems*. Lecture Notes in Control and Information Sciences, LNCIS 77, Springer-Verlag, New York.
- [Bass88] M. BASSEVILLE (1988). Detecting changes in signals and systems — A survey. *Automatica*, vol. 24, pp. 309-326.
- [BasNik] M. BASSEVILLE and I. V. NIKIFOROV (1992). *Detection of abrupt changes : theory and applications*. International Series in Information and Systems Science, Prentice-Hall, Englewood Cliffs, N.J. à paraître.
- [Ben-v87] A. BENVENISTE, M. BASSEVILLE and G. MOUSTAKIDES (1987). The asymptotic local approach to change detection and model validation. *IEEE Trans. Automatic Control*, vol. AC 32, n° 4, pp. 583-592.
- [Broem82] L. D. BROEMELING (1982). *Jal Econometrics*, vol. 19, special issue on structural change in econometrics.
- [Broem87] L. D. BROEMELING and H. TSURUMI (1987). *Econometrics and structural change*. Marcel Dekker, New York.
- [Cohen87] A. COHEN (1987). *Biomedical signal processing, vol. 1 ; Time and frequency domain analysis ; vol. 2 : Compression and automatic recognition*. CRC Press, Inc., Boca Raton, FL.

- [Desh83] J. DESHAYES et D. PICARD (1983). *Ruptures de modèles en statistique*. Thèses d'État, Université de Paris-Sud, Orsay, France.
- [Difra90] R. DI FRANCESCO (1990). Real-time speech segmentation using pitch and convexity jump models : application to variable rate speech coding. *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP 38, n° 5, pp. 741-748.
- [Fav84] G. FAVIER and A. SMOLDERS (1984). Adaptive smoother-predictors for tracking maneuvering targets. *Proc. 23rd Conf. Decision and Control*, Las Vegas, NV, pp. 831-836.
- [Frank90] P. M. FRANK (1990). Fault diagnosis in dynamic systems using analytical and knowledge based redundancy — A survey and new results. *Automatica*, vol. 26, pp. 459-474.
- [Iserm84] R. ISERMANN (1984). Process fault detection based on modeling and estimation methods — A survey. *Automatica*, vol. 20, pp. 387-404.
- [Krish88] P. R. KRISHNAIAH and B. Q. MIAO (1988). Review about estimation of change points. In *Handbook of statistics* (P. R. Krishnaiah, C. R. Rao, eds.), vol. 7, Elsevier, New York, pp. 375-402.
- [Ljung90] L. LJUNG and S. GUNNARSSON (1990). Adaptation and tracking in system identification — A survey. *Automatica*, vol. 26, n° 1, pp. 7-22.
- [Nikif83] I. V. NIKIFOROV (1983). *Sequential detection of abrupt changes in time series properties*. Nauka, Moscow. (En Russe.)
- [Nikif86] I. V. NIKIFOROV and I. N. TIKHONOV (1986). Application of change detection theory to seismic signal processing. In *Detection of abrupt changes in signals and dynamical systems* (M. Basseville, A. Benveniste, eds.). Lecture Notes in Control and Information Sciences, LNCIS 77, Springer-Verlag, New York, pp. 355-373.
- [Nikif89] I. V. NIKIFOROV, I. N. TIKHONOV and T. MIKHAILOVA (1989). *Automatic on-line processing of seismic station data — Theory and applications*. Far Eastern Dept of USSR Academy of Sciences, Vladivostok, USSR. (En Russe.)
- [Patt89] R. PATTON, P. FRANK and R. CLARK, eds. (1989). *Fault diagnosis in dynamic systems — Theory and application*. International Series in Systems and Control Engineering, Prentice-Hall, Englewood Cliffs, N.J.
- [Poir76] D. J. POIRIER (1976). *The econometrics of structural change*. North-Holland, Amsterdam.
- [Shir78] A. N. SHIRYAEV (1978). *Optimal Stopping Rules*. Springer-Verlag, New York.
- [Telk87] L. TELKSNYS, ed. (1987). *Detection of changes in random processes*. Optimization Software, Inc., Publications Division, New York.
- [Wahn90] E. WAHNON and N. BERMAN (1990). Tracking algorithm designed by the asymptotic local approach. *IEEE Trans. Automatic Control*, vol. AC 35, n° 4, pp. 440-443.
- [Will76] A. S. WILLSKY (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, vol. 12, pp. 601-611.
- [Zacks83] S. ZACKS (1983). Survey of classical bayesian approaches to the change-point problem : fixed sample and sequential procedures of testing and estimation. In *Recent advances in statistics : papers in honor of Herman Chernoff on his sixtieth birthday* (M. H. Rizvi, J. Rustagi, D. Siegmund, eds.). Academic Press, New York, pp. 245-269.

Auteurs de la partie « Segmentation de Signaux »

Michèle BASSEVILLE, IRISA/CNRS, Campus de Beau-
lieu, 35042 Rennes Cedex, tél. : 99 84 71 00.

Christian DONCARLI, Marie-Françoise LUCAS,
LAN/ENSM, 1, rue de la Noë, 44072 Nantes Cedex, tél. :
40 37 16 00.

Denis DE BRUCQ, Olivier COLOT, LACIS/ITEPEA,
BP 119, 76134 Mont-St-Aignan Cedex, tél. : 35 14 66 17.

Abdelahad BENSALLAM, GAPSE/ENSEEIH, 2, rue
C. Camichel, 31071 Toulouse Cedex, tél. : 61 58 82 81.

Hervé RIX, Éric THIERRY, LASSY/I3S, 41, bd Napo-
léon III, 06041 Nice Cedex, tél. : 93 21 79 55.

François KAUFFMANN, Université de Caen, Dept.
Maths, 14032 Caen Cedex, tél. : 31 45 56 02.

Type	Algorithme	Abbréviation	IRISA	GAPSE	LAN	LASSY	LACIS	U. Caen
M	<i>Dérivée filtrée</i>	DF	X					
M	<i>Page-Hinkley</i>	HK	X		x			
M	<i>GLR-Willsky</i>	GLR	X		x			
M	<i>GLR modifié</i>	GLRmod	X					
M	<i>Kalman étendu</i>	FKE-M			X			
D	<i>SNR</i>	SNR					x	X
D	<i>EPL</i>	EPL		X				
D	<i>Paquets</i>	DAREP					X	
D	<i>DIS2f</i>	DIS2f	x	X				
D	<i>Kalman étendu</i>	FKE-D			X			
D	<i>Divergence-Hinkley</i>	DIV	X	x				x
D	<i>Brandt</i>	Brandt	x	X				
F	<i>Gabarit</i>	GAB					x	X
F	<i>Corrélation</i>	MFC			x	X		
F	<i>Fct Répartition</i>	MFR			x	X		

x : rédaction X : coordination

Tableau 1. — Liste des algorithmes décrits

Dérivée filtrée (DF)

1. Publications. Exemples

Publication originale

[1] inconnue ! voir méthodes de gradients en détection de contours dans les images.

Points d'entrées bibliographiques conseillés

- [2] M. BASSEVILLE (1981). Edge detection using sequential methods for change in level. *IEEE Trans. Acoust. Speech. Sig. Process.*, ASSP-29, pp. 32-50.
- [3] M. BASSEVILLE (1986). On-line detection of jumps in mean. In *Detection of abrupt changes in signals and dynamical systems* (M. Basseville, A. Benveniste, eds.), chap. 1. Lecture Notes in Control and Information Sciences, LNCIS-77, Springer-Verlag, New York.

Exemples d'applications

- détection de contours dans des images (par traitement ligne par ligne) ;
- traitement de signaux géophysiques ;
- détection de potentiels d'unités motrices dans les EMG ;
- chromatographie.

2. Objectifs

L'algorithme dit dérivée filtrée est de type M et détecte en ligne un changement de moyenne d'amplitude inconnue dans un signal constant par morceaux et perturbé par un bruit blanc.

3. Principe de la méthode

Il repose sur l'idée très intuitive suivante, souvent utilisée en traitement d'images pour la détection de contours. Idéalement, en l'absence de bruit, un changement de moyenne dans un signal est caractérisé localement par une grande valeur absolue de la dérivée (discrète) des observations. Comme l'opérateur de dérivation opère de manière pauvre en présence de bruit, un détecteur plus réaliste consiste à appliquer un filtre — de longueur $2l$ — avant la dérivation et le seuillage. Mais, à cause de ce filtrage, plusieurs alarmes peuvent se produire au voisinage d'un saut. Pour rendre le détecteur plus robuste, on compte le nombre de franchissements du seuil pendant une durée égale à la longueur du filtre, et on ne décide qu'il y a eu saut que si ce nombre excède une valeur $N_D < l$. En pratique, le simple fait de prendre $N_D = 2$ améliore sensiblement la qualité de la détection par rapport à $N_D = 1$ (détecteur « grossier »).

4. Algorithme

Objets manipulés

En reprenant les notations de l'introduction : $n = m = 1$, $F = H = 1$, $W = 0$.

Procédures externes requises

Un filtre élémentaire $\tilde{Y}_n = \sum_{i=-l}^l \gamma_i (Y_{n-i} - \mu_0)$, par exemple :

- *filtre intégrateur* rectangulaire de longueur $2l$ ($\gamma_i = 1$), c'est-à-dire moyenne des points sur une fenêtre de longueur $2l$;
- *filtre triangulaire* de même taille, $\gamma_{-l+k} = \gamma_{l-k}$ ($0 \leq k \leq l$).

Séquencement des opérations

- dans le cas du filtre intégrateur, le résultat du filtrage et de la dérivation $Z_n = \tilde{Y}_n - \tilde{Y}_{n-1}$ est alors :

$$Z_n = Y_{n+l} - Y_{n-l} \quad (2.1)$$

c'est-à-dire la différence de 2 points décalés ;

- dans le cas du filtre triangulaire,

$$Z_n = (Y_n + \dots + Y_{n+l-1}) - (Y_{n-1} + \dots + Y_{n-l}) \quad (2.2)$$

c'est-à-dire la différence de 2 moyennes locales décalées.

On compte le nombre de fois où $|Z_n|$ excède un seuil λ pendant une durée l , et on détecte si ce nombre excède N_D (en pratique, 2). L'instant de rupture estimé est l'instant de détection.

5. Quelques caractéristiques de la méthode

- *précautions d'emploi* : après une alarme, on ne redémarre le calcul du détecteur qu'après un nombre l d'observations ;
- *avantages* : simplicité, faible coût ;
- *inconvenients* : il est moins bon (efficacité, robustesse au niveau de bruit) que l'algorithme HK de type somme cumulée de Page-Hinkley, et ce d'autant plus que le niveau de bruit est élevé ;
- *complexité* : cet algorithme est de faible complexité.

6. Références complémentaires et remarques

L'utilisation d'un « compteur » pour rendre la détection plus robuste, que l'on retrouve dans les algorithmes EPL et SNR, est connue depuis très longtemps en contrôle de qualité sous le nom de *run sum test* [4]. Le filtre triangulaire y est également utilisé. Cette idée est également utilisée en chromatographie.

- [4] S. W. ROBERTS (1966). A comparison of some control charts procedures. *Technometrics*, vol. 8, pp. 411-430.

Test CUSUM de Page-Hinkley (HK)

1. Publications. Exemples

Publication originale

- [1] E. S. PAGE (1954). Continuous inspection schemes. *Biometrika*, 41, pp. 100-115.

Points d'entrées bibliographiques conseillés

- [2] D. V. HINKLEY (1971). Inference about the change-point from cumulative sum tests. *Biometrika*, 58, pp. 509-523.
- [3] M. BASSEVILLE (1981). Edge detection using sequential methods for change in levels. *IEEE Trans. Acoust. Speech. Sig. Process.*, ASSP-29, pp. 32-50.
- [4] M. BASSEVILLE (1986). On-line detection of jumps in mean. In *Detection of abrupt changes in signals and dynamical systems* (M. Basseville, A. Benveniste, eds.), chap. 1. Lecture Notes in Control and Information Sciences, LNCIS-77, Springer-Verlag, New York.

Exemples d'applications

- contrôle de qualité ;
- détection de contours dans des images (par traitement ligne par ligne) ;
- traitement de signaux géophysiques et biomédicaux ;
- prendre une décision avec un indicateur qui réagit à une rupture par un changement de sa dérive moyenne : voir l'algorithme « divergence » dans la fiche DIV ; voir également la fiche FKE-D.

2. Objectifs

Le test de somme cumulée CUSUM dit de Page-Hinkley est de type M et détecte **en ligne** un changement de moyenne d'amplitude inconnue dans un signal $(Y_n)_n$ constant par morceaux et perturbé par un bruit blanc. En particulier, Y_n peut être l'innovation d'un filtre linéaire.

3. Principe de la méthode

Dans le cas (idéal) où les moyennes du signal avant et après rupture sont connues, le rapport de vraisemblance entre les hypothèses « il n'y a pas eu de saut » et « il y a eu un saut avant l'instant courant » revient à un **intégrateur avec seuil adaptatif**, ce dernier venant de la maximisation par rapport à l'instant de saut.

Dans la situation — plus réaliste dans un contexte temps-réel — où l'on ne connaît ni l'amplitude du saut ni son sens (augmentation ou diminution), on suppose connaître la valeur moyenne du signal avant le saut, quitte à l'estimer en ligne. La solution proposée consiste alors à activer deux détecteurs en parallèle, pour détecter des augmentations (resp. des diminutions) de moyenne, dans lesquels on

introduit une notion d'amplitude minimum de saut à détecter ν_m .

4. Algorithme

Objets manipulés

En reprenant les notations de l'introduction : $n = m = 1$, $F = H = 1$, $W = 0$. Soit μ_0 la valeur — éventuellement estimée — de l'état constant X.

Procédures externes requises

Filtre élémentaire — moyenne glissante par exemple — pour estimer la moyenne avant rupture.

Séquencement des opérations

Pour une augmentation de moyenne :

$$T_0 = 0$$

$$T_n = \sum_{k=1}^n \left(Y_k - \mu_0 + \frac{\nu_m}{2} \right)$$

$$M_n = \max_{0 \leq k \leq n} T_k$$

détection quand : $g_n = M_n - T_n \geq \lambda$. (2.3)

Pour une diminution :

$$U_0 = 0$$

$$U_n = \sum_{k=1}^n \left(Y_k - \mu_0 - \frac{\nu_m}{2} \right)$$

$$m_n = \min_{0 \leq k \leq n} U_k$$

détection quand : $g_n = U_n - m_n \geq \lambda$. (2.4)

L'alarme est donnée par le premier des deux détecteurs qui déclenche. L'estimateur de l'instant de saut est alors le dernier instant auquel le maximum (resp. minimum) courant a été atteint.

5. Quelques caractéristiques de la méthode

- *précautions d'emploi* : pour implémenter ce détecteur dans le cas de moyennes connues, μ_0 avant et μ_1 après le saut, il suffit de remplacer ν_m par $|\mu_1 - \mu_0|$. Il est à noter que ceci peut constituer une alternative intéressante à la détection de sorties de visées, couramment utilisée en surveillance dans l'industrie ;
- *avantages* : simplicité, faible coût. En outre, ce détecteur apparaît expérimentalement très efficace et robuste au bruit et convient dans de très nombreuses situations. Il est

meilleur (efficacité, robustesse au niveau de bruit) qu'un algorithme de type « dérivée filtrée » — voir fiche DF — et ce d'autant plus que le niveau de bruit est élevé. L'introduction d'une notion d'amplitude minimum de saut à détecter induit une moins grande sensibilité au choix du seuil, grâce au fait que les choix de l'amplitude et du seuil sont de fait découplés en pratique. Des résultats théoriques et expérimentaux sur les performances de ce détecteur sont décrits en [3] et [4]. Voir aussi [5] ci-dessous.

- *inconvenients* : l'algorithme n'est pas optimal. On peut l'améliorer en incorporant, à la place de l'amplitude minimum forfaitaire ν_m , un estimateur par maximum de vraisemblance de l'amplitude du saut, ce qui est le détecteur optimal (voir la fiche GLR) ;

- *complexité* : cet algorithme est de faible complexité, car la variable de décision g_n peut être calculée de manière récursive :

$$g_n = \left(g_{n-1} + Y_n - \mu_0 - \alpha \frac{\nu_m}{2} \right)^+ \quad (2.5)$$

où $(x)^+ = \max(x, 0)$ et où $\alpha = 1$ pour une augmentation de moyenne et $\alpha = -1$ pour une diminution.

6. Références complémentaires et remarques

On peut concevoir des règles de décision de type CUSUM : $g_n \geq \lambda$, où g_n est de la forme indiquée plus haut, dans des situations beaucoup plus complexes que le cas scalaire gaussien indépendant considéré ici, et en particulier dans le cas ARMA multidimensionnel. De tels détecteurs sont utilisés dans des stations sismiques pour la détection de raz-de-marée ([Nikif86]). On sait aussi calculer analytiquement les performances théoriques de ces détecteurs CUSUM généraux (voir [5]).

[5] I. V. NIKIFOROV (1986). Sequential detection of changes in stochastic systems. In *Detection of abrupt changes in signals and dynamical systems* (M. Basseville, A. Benveniste, eds.), chap. 7. Lecture Notes in Control and Information Sciences, LNCIS-77, Springer-Verlag, New York.

Algorithme GLR de Willsky (GLR)

1. Publications. Exemples

Publication originale

- [1] A. S. WILLSKY and H. L. JONES (1976). A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Trans. Aut. Control*, 21, pp. 108-112 (version courte parue dans *Proc. CDC*, 1974).

Points d'entrée bibliographiques conseillés

- [2] M. BASSEVILLE and A. BENVENISTE (1983). Design and comparative study of some sequential jump detection algorithms for digital signals. *IEEE Trans. Acoust. Speech. Sig. Process.*, ASSP-31, pp. 521-534.
- [3] A. S. WILLSKY (1986). Detection of abrupt changes in dynamic systems. In *Detection of abrupt changes in signals and dynamical systems* (M. Basseville, A. Benveniste, eds.), chap. 2. Lecture Notes in Control and Information Sciences, LNCIS-77, Springer-Verlag, New York.

Exemples d'applications

- détection de pannes de capteurs en avionique ;
- analyse de signaux ECG ;
- détection de manœuvres de mobiles ;
- surveillance du trafic routier ;
- traitement de signaux géophysiques (signaux de sondage de forages pétroliers).

2. Objectifs

L'algorithme GLR est de type M et détecte et estime en ligne des **changements additifs** sur l'entrée U, la sortie Y ou l'état X de tout système linéaire multivariable excité par des bruits blancs. Il permet d'améliorer la capacité de poursuite du filtre de Kalman en présence de non-stationnarités rapides et peut travailler sur des signaux multi-capteurs.

3. Principe de la méthode

L'idée est que tout changement additif sur l'état ou l'observation d'un système linéaire se traduit par un saut de moyenne sur l'innovation du filtre de Kalman associé au modèle sans rupture. (Rappelons que, dans le cas de l'équation (1.3) de l'introduction, ce saut de moyenne est la réponse impulsionnelle du système déterministe (H, F) à l'impulsion sur l'état.) La procédure de détection, de type rapport de vraisemblance généralisé — d'où le nom GLR (generalized likelihood ratio) — exploite de plus la nature gaussienne des excitations des équations d'état et d'observation du modèle pour rendre explicites certaines maximisations. L'algorithme comprend donc un filtre de Kalman, un détecteur de saut, un estimateur de l'instant r

et de l'amplitude ν du saut, et une procédure de réactualisation du filtre de Kalman après le saut ; le tout étant calculé de manière récursive. La seule chose qui n'est pas récursive est l'estimateur de l'instant de saut, qui est obtenu par recherche exhaustive du maximum du détecteur sur un horizon donné.

4. Algorithme

On donne ici la description détaillée de l'algorithme dans le cas du modèle d'état (1.1) introduit au début du document.

Objets manipulés

Le modèle de rupture est (1.3).

Procédures externes requises

Le filtre de Kalman qui estime l'état du modèle peut être résumé de la manière suivante. La prédiction à un pas, l'innovation ε_k , et l'état estimé sont donnés par :

$$\begin{aligned} \hat{X}_{k+1|k} &= F\hat{X}_{k|k} \\ \varepsilon_k &= Y_k - H\hat{X}_{k|k-1} \\ \hat{X}_{k|k} &= \hat{X}_{k|k-1} + K_k \varepsilon_k \end{aligned} \quad (2.6)$$

où K_k est le gain de Kalman, qui, avec la covariance de l'erreur d'estimation de l'état $P_{k|k}$ et la covariance de l'innovation Σ_k , est donné par :

$$\begin{aligned} K_k &= P_{k|k-1} H^T \Sigma_k^{-1} \\ P_{k+1|k} &= FP_{k|k} F^T + Q \\ P_{k|k} &= (I - K_k H) P_{k|k-1} \\ \Sigma_k &= HP_{k|k-1} H^T + R. \end{aligned} \quad (2.7)$$

Séquencement des opérations :

1) **Estimation de l'instant et de l'amplitude de saut :** la linéarité du modèle (1.1) et le caractère additif de la rupture en (1.3) conduisent à décomposer l'estimateur de l'état et l'innovation en :

$$\begin{aligned} \hat{X}_{k|k} &= \hat{X}_{k|k}^0 + \zeta(k, r) \nu \\ \varepsilon_k &= \varepsilon_k^0 + \rho(k, r) \nu \end{aligned} \quad (2.8)$$

où l'exposant « 0 » désigne les quantités en l'absence de rupture, et où le dernier terme de chaque équation représente l'effet d'une rupture d'amplitude $\nu = \nu_x$ se produisant à un instant $r \leq k$. Les fonctions ζ et ρ sont calculées, récursivement dans une fenêtre, à l'aide de :

$$\zeta(k, r) = K_k \rho(k, r) + F\zeta(k-1, r) \quad (2.9)$$

$$\rho(k, r) = H(F^{k-r} - F\zeta(k-1, r)) \quad (2.10)$$

avec les conditions initiales :

$$\begin{aligned} \zeta(k, k) &= \mathbf{K}_k \mathbf{H} \\ \rho(k, k) &= \mathbf{H} . \end{aligned} \quad (2.11)$$

En d'autres termes, l'innovation ε_k a pour loi :

$$\begin{aligned} \varepsilon_k &\sim \mathcal{N}(0, \Sigma_k) && \text{en l'absence de rupture} \\ \varepsilon_k &\sim \mathcal{N}(\rho(k, r) \nu, \Sigma_k) && \text{après rupture} . \end{aligned} \quad (2.12)$$

Le rapport de vraisemblance L entre ces deux hypothèses est donné par :

$$l(k, r) = -2 \ln L(k, r) = 2 \nu^T d(k, r) - \nu^T \mathbf{C}(k, r) \nu \quad (2.13)$$

où

$$\begin{aligned} \mathbf{C}(k, r) &= \sum_{j=r}^k \rho^T(j, r) \Sigma_j^{-1} \rho(j, r) \\ d(k, r) &= \sum_{j=r}^k \rho^T(j, r) \Sigma_j^{-1} \varepsilon_j . \end{aligned}$$

On remarque que d n'est autre que la corrélation entre l'innovation ε et la signature ρ de la rupture sur l'innovation (voir la fiche FA pour une autre utilisation de la corrélation en détection). L'amplitude du saut est estimée en maximisant cette expression, c'est-à-dire par :

$$\hat{\nu}_k = \mathbf{C}^{-1}(k, r) d(k, r) . \quad (2.14)$$

Le logarithme du rapport de vraisemblance généralisé pour un r fixé est alors obtenu en remplaçant l'amplitude par son estimateur :

$$l(k, r) = d^T(k, r) \mathbf{C}^{-1}(k, r) d(k, r) \quad (2.15)$$

r est estimé en recherchant le maximum de l dans une fenêtre de longueur M :

$$\hat{r}_k = \arg \max_{k-M+1 \leq r \leq k} l(k, r) \quad (2.16)$$

et l'estimation de ν devient :

$$\hat{\nu}_k = \mathbf{C}^{-1}(k, \hat{r}_k) d(k, \hat{r}_k) . \quad (2.17)$$

La règle de décision l peut alors s'écrire comme une forme quadratique en $\hat{\nu}_k$:

$$l(k, \hat{r}_k) = \hat{\nu}_k^T \mathbf{C}(k, \hat{r}_k) \hat{\nu}_k \underset{H_0}{\overset{H_1}{>}} \lambda . \quad (2.18)$$

Quand une rupture est décidée, l'algorithme fournit les estimations :

$$\hat{\nu} = \hat{\nu}_k, \quad \hat{r} = \hat{r}_k . \quad (2.19)$$

2) Réactualisation du filtre après la détection d'un saut : la procédure proposée en [1] consiste à réactualiser l'estimation des variables d'état et la covariance d'erreur d'estima-

tion de l'état à l'aide de :

$$\begin{aligned} \hat{\mathbf{X}}_{k, \text{réact}} &= \hat{\mathbf{X}}_k^0 + (\mathbf{F}^{k-\hat{r}_k} - \zeta(k, \hat{r}_k)) \hat{\nu}_k \\ \mathbf{P}_{k|k, \text{réact}} &= \mathbf{P}_{k|k}^0 + (\mathbf{F}^{k-\hat{r}_k} - \zeta(k, \hat{r}_k)) \times \\ &\quad \mathbf{C}^{-1}(k, \hat{r}_k) (\mathbf{F}^{k-\hat{r}_k} - \zeta(k, \hat{r}_k))^T . \end{aligned} \quad (2.20)$$

5. Quelques caractéristiques de la méthode

- **précautions d'emploi :** il est à noter que la procédure n'est pas un algorithme travaillant à horizon fini, car à chaque instant le détecteur est calculé en tenant compte de tout le passé du signal ; seule la recherche de l'instant de rupture est bornée dans le temps ;

- **avantages :** le principal avantage est qu'elle fournit un bon estimateur de l'amplitude du saut qui, réinjecté dans le filtre de Kalman, permet de réinitialiser celui-ci après la rupture d'une manière moins grossière qu'au départ, puisque tenant compte de tout le passé du signal avant la rupture ;

- **inconvenients :** le principal est la difficulté de réglage du seuil à appliquer au rapport de vraisemblance généralisé [2] : la loi théorique (χ^2) du détecteur ne sert pas en pratique pour régler le seuil qui, en fait, dépend de la taille de la fenêtre utilisée pour la recherche de l'instant de rupture. De plus, la qualité de la détection dépend sensiblement de ce seuil (voir les commentaires faits à la fin de la fiche suivante). Pour cette raison, une règle de décision modifiée a été introduite (voir GLRmod). Le deuxième inconvénient est le coût ;

- **complexité :** en [2] sont données des indications sur la complexité de cet algorithme — et d'autres du même type — en termes de nombres d'additions et de multiplications par échantillon de signal, et de place mémoire requise, dans le cas où on utilise, pour le calcul du rapport de vraisemblance, l'estimation permanente des caractéristiques par le biais du filtre de Kalman (voir aussi la fiche GLRmod). Cependant, dans tous les exemples d'application cités dans cette fiche sauf le dernier, on utilise seulement le régime permanent de ce filtre (cf. [1]), ce qui permet de calculer les « signatures » de ruptures (2.9)-(2.10) *a priori* et non en ligne comme en [2], et donc de gagner en complexité.

6. Références complémentaires et remarques

1. Il est important de remarquer que cet algorithme peut être aisément étendu au cas où le modèle de rupture (1.3) ou (1.4) inclut des **profils dynamiques** de rupture plus généraux que des impulsions ou des échelons. Il suffit alors de modifier convenablement (2.9) et (2.10). Voir à ce sujet [3].

2. Pour la détection de manœuvres de mobiles :

[4] J. KORN, S. W. GULLY and A. S. WILLSKY (1982). Application of the generalized likelihood ratio algorithm to maneuver detection and estimation. *Proc. American Control Conference*, Arlington, VA, pp. 792-798.

Algorithme GLR modifié (GLRmod)

1. Publications. Exemples

Publication originale

[1] M. BASSEVILLE and A. BENVENISTE (1983). Design and comparative study of some sequential jump detection algorithms for digital signals. *IEEE Trans. Acoust. Speech. Sig. Process.*, ASSP-31, pp. 521-534.

Points d'entrée bibliographiques conseillés

[2] M. BASSEVILLE (1986). Two examples of application of the GLR method in Signal Processing. In *Detection of abrupt changes in signals and dynamical systems*, LNCIS-77, Springer-Verlag, New York.

Exemples d'applications

- traitement de signaux géophysiques (signaux de sondage de forages pétroliers) ;
- surveillance du rendement du compresseur d'une turbine à gaz.

2. Objectifs

Cette méthode est de type M et a les mêmes objectifs que l'algorithme GLR original de Willsky (voir GLR).

3. Principe de la méthode

Il est le même que pour la méthode originale, dans le cas particulier d'un modèle d'état de type « pente constante + bruit ». La seule modification concerne la règle de décision qui n'est plus basée sur le rapport de vraisemblance mais sur une version lissée de l'estimateur de l'amplitude du saut. La procédure qui en résulte est un filtre passe-bas du signal partout sauf aux instants de sauts de moyenne.

4. Algorithme

Objets manipulés

Le modèle particulier choisi est un modèle d'état de dimension 2 — i.e. pente constante + bruit — et sur lequel se produit un saut de moyenne. La situation de rupture de pente (saut sur la deuxième composante de l'état) peut être envisagée de manière similaire. Le modèle d'état est :

$$\begin{cases} x_{k+1} = x_k + \mu_k + w_k^1 + \nu \delta_{r, k+1} \\ \mu_{k+1} = \mu_k + w_k^2 \\ y_k = x_k + v_k \end{cases} \quad (2.21)$$

où $(w_k^1)_k$, $(w_k^2)_k$ et $(v_k)_k$ sont des bruits blancs, gaussiens, de moyenne nulle et indépendants, dont les variances sont respectivement q^1 et q^2 choisies *a priori* et σ_k^2 estimée en ligne. q^1 et q^2 permettent au filtre de Kalman de suivre les

fluctuations lentes du signal par rapport au modèle de pente constante. En reprenant les notations de l'introduction : $m = 1$, $H = 1$, $R = \sigma^2$, $n = 2$, $Q = \text{diag}(q^1, q^2)$, et :

$$F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \quad (2.22)$$

L'estimateur choisi pour la variance du bruit de mesure est [2] :

$$\hat{\sigma}_{k+1}^2 = \frac{k-2}{k-1} \hat{\sigma}_k^2 1_{\{k>2\}} + \frac{k}{(k+1)(k+2)} (y_{k+1} - \hat{x}_k)^2. \quad (2.23)$$

Séquencement des opérations

Toutes les opérations faites dans l'algorithme GLR original sont à faire ici aussi. On remarquera que, dans le cas particulier d'un modèle de pente nulle, cet algorithme est exactement le détecteur de somme cumulée présenté dans la fiche HK. Pour l'algorithme GLRmod, il faut de plus calculer un test sur une version lissée de l'estimateur de l'amplitude du saut. On a alors à se donner une amplitude minimum de saut à détecter et un seuil, ce qui permet — comme pour l'algorithme HK — de gagner considérablement en robustesse, au sens du découplage entre ces paramètres et la taille M de la fenêtre et aussi d'une plus grande insensibilité de la détection à une variation de seuil [1].

Soit $\bar{\nu}_k$ l'estimation de l'amplitude de saut lissée dans une fenêtre de longueur p ($p < M$) :

$$\bar{\nu}_k = \frac{1}{N_k} \sum_{j=k-N_k+1}^k \hat{\nu}_j. \quad (2.24)$$

La variance empirique de $\hat{\nu}_k$ est calculée par

$$S_k = \frac{1}{N_k - 1} \sum_{j=k-N_k+1}^k (\hat{\nu}_j - \bar{\nu}_k)^2 \quad (2.25)$$

où

$$N_k = p 1_{\{k>p+1\}} + (k-1) 1_{\{k \leq p+1\}}. \quad (2.26)$$

Le test est :

$$T_k = \frac{N_k (\bar{\nu}_k - \nu_m)^2}{S_k} \underset{H_0}{\overset{H_1}{\geq}} \lambda \quad (2.27)$$

où ν_m et λ sont des quantités positives à choisir.

Quand H_1 est décidée, l'algorithme fournit, comme l'algorithme GLR original, les estimations :

$$\hat{\nu} = \hat{\nu}_k, \quad \hat{\theta} = \hat{\theta}_k \quad (2.28)$$

et le filtre est réactualisé comme dans la version originale (voir GLR).

5. Quelques caractéristiques de la méthode

- *précautions d'emploi* : tous les estimateurs du bruit de mesure ne sont pas équivalents dans cet algorithmique ; voir les commentaires faits en [2]. D'autre part, tout comme GLR, cet algorithme n'est pas un algorithme travaillant à horizon fini ;
- *avantages* : le principal avantage est sa grande robustesse vis-à-vis de la simplicité du modèle sous-jacent — par rapport au signal étudié — et vis-à-vis du caractère très hautement non stationnaire des signaux qu'il peut traiter de manière cohérente avec un seul et même réglage de paramètres. Voir aussi les commentaires faits dans la fiche GLR ;
- *inconvenients* : le principal est son coût. Pour cette raison, un détecteur voisin alliant la simplicité du détecteur CUSUM de Page-Hinkley à la précision de l'estimation de l'amplitude du saut et la procédure de réactualisation du filtre de Kalman de l'algorithme GLR, a été étudié en [1] ;
- *complexité* : voir GLR.

6. Références complémentaires et remarques

Les règles de décision des algorithmes GLR et GLRmod, respectivement (2.18) et (2.27), sont toutes deux des formes quadratiques en $\hat{\nu}_k$. Pour le modèle particulier de pente considéré ici et pour l'application aux signaux géophysiques étudiée en [1], leurs performances sont cependant sensiblement différentes en termes de robustesse essentiellement vis-à-vis du réglage de leurs paramètres et vis-à-vis du caractère relativement grossier du modèle choisi. Une explication possible est la suivante : la matrice $C^{-1}(k, \hat{r}_k)$ est théoriquement la matrice de covariance de $\hat{\nu}_k$, mais seulement en présence d'une rupture instantanée et sous l'hypothèse que \hat{r}_k est un estimateur sans biais de l'instant de rupture r . Or, dans les signaux considérés, les sauts de moyenne se produisent non pas instantanément mais en plusieurs échantillons (parfois une dizaine !), et l'estimateur de covariance empirique utilisé en (2.27) semble plus approprié.

Sauts de moyenne d'un ARMA (FKE-M)

1. Publications. Exemples

Publication originale

[1] C. DONCARLI and Y. BIZAIS (1981). Un algorithme de détection-estimation pour le traitement de signaux constants par morceaux perturbés par un bruit coloré. *Congrès AFCET-Automatique*.

Exemples d'applications

- Détection de contours sur une ligne d'image ;
- Détection de recalage d'un capteur de poursuite (saut du point brillant).

2. Objectifs

- Détection et estimation des sauts de moyenne d'un ARMA scalaire.

3. Principe de la méthode

- Estimation des coefficients de l'ARMA et de la moyenne par filtrage de Kalman étendu.
- Détection par double filtrage du carré de la pseudo-innovation issue d'un filtre de Kalman étendu (amélioration du SNR).
- Actualisation de l'estimation et inhibition temporaire du détecteur.

4. Algorithme

Objets manipulés

Signal mesuré $y_k = m_k + Y_k$ avec :

$$Y_k = \sum_{i=1}^p a_i Y_{k-i} + \sum_{i=1}^p b_i E_{k-i} + E_k \quad (2.29)$$

où E_k séquence blanche de variance inconnue R , a_i et b_i inconnus mais constants, ordre p connu, $m_k = m_{k-1} = m$ s'il n'y a pas de saut à l'instant k , $m_k \neq m_{k-1}$ s'il y a un saut à l'instant k .

Objectifs

Estimer a_i , b_i , m_k , et détecter si k est un instant de rupture.

Procédures externes requises

Identification récursive par filtre de Kalman étendu (FKE).

Séquencement des opérations

1. Identification récursive par FKE :

Si k n'est pas un instant de rupture :

$$y_k = \sum_{i=1}^p a_i y_{k_i} + \sum_{i=1}^p b_i E_{k-i} + E_k + d_k \quad (2.30)$$

avec :

$$d_k = d_{k-1} = d = m \left(1 - \sum_{i=1}^p a_i \right). \quad (2.31)$$

On définit un état étendu :

$$X_k^T = [a_1, \dots, a_p, b_1, \dots, b_p, d_k, E_{k-1}, \dots, E_{k-p}]. \quad (2.32)$$

Remarquer d_k en position $2p + 1$ dans X_k .

Équation d'état linéaire :

$$X_{k+1} = FX_k + GE_k \quad (2.33)$$

avec

$$F = \begin{pmatrix} I & 0 \\ 0 & J \end{pmatrix}$$

I = matrice unité $(2p + 1) \times (2p + 1)$

J = matrice de décalage $p \times p$

$$G^T = [0 \dots 0 \ 1 \ 0 \dots 0].$$

L'équation de mesure (2.30) non linéaire est linéarisée autour de la prédiction courante de l'état étendu :

$$Z_k = H_k X_k + E_k \quad (2.34)$$

avec :

$$H_k = [y_{k-1}, \dots, y_{k-p}, E_{k-1|k-1}, \dots, E_{k-p|k-1}, 1, b_{1|k-1}, \dots, b_{p|k-1}]$$

$$Z_k = y_k + \sum_{i=1}^p b_{i|k-1} E_{k-i|k-1}. \quad (2.35)$$

Remarquer le 1 en position $2p + 1$ dans H_k .

Les équations de Kalman (*procédure externe standard*) sont ensuite appliquées au système linéarisé, avec une estimation récursive de R :

$$R_{k|k-1} = u_k R_{k-1|k-2} + (1 - u_k) \varepsilon_k^2 \quad (2.36)$$

$$\varepsilon_k = E_{k|k-1} = Z_k - H_k X_{k|k-1}$$

où $0 < u_k < 1$, et initialisée par la valeur instantanée ε_1^2 .

2. Détection : parallèlement à $R_{k|k-1}$, on calcule :

$$S_{k|k-1} = v_k S_{k-1|k-2} + (1 - v_k) \varepsilon_k^2 \quad (2.37)$$

avec : $0 < v_k < u_k < 1$, et initialisée de la même façon, puis :

$$T_k = \frac{S_{k|k-1}}{R_{k|k-1}} \quad (2.38)$$

u_k et v_k seront définis plus loin. Le principe de la détection consiste à comparer T_k à un seuil. En effet, s'il y a eu un saut, $T_k \gg 1$ car l'horizon de calcul de $S_{k|k-1}$ est plus court que celui de $R_{k|k-1}$. Inversement, s'il n'y a pas de saut, alors $T_k \approx 1$. Des valeurs habituelles du seuil peuvent être 3 ou 4.

3. Actualisation (et initialisation) : si un saut est détecté à l'instant k (ou si $k = 0$), on modifie (ou initialise) la composante d_k du vecteur d'état étendu :

$$d_{k|k}^* = y_k(1 - \Sigma_{a_i|k}) \quad (2.39)$$

et on (ré)initialise l'élément correspondant de la matrice de variance-covariance d'erreur d'estimation :

$$P_{k|k}^* = P_{k|k} + (d_{k|k}^* - d_{k|k})^2. \quad (2.40)$$

Enfin, on inhibe le détecteur pendant quelques points

pour éviter les alarmes multiples :

$$\begin{aligned} u_k &= v_k = v \\ u_{k+1} &= 0,5(u + u_{k+i-1}) \text{ avec } v = 0,5 \text{ et } u = 0,96 \\ v_{k+i} &= v. \end{aligned}$$

5. Quelques caractéristiques de la méthode

- *précautions d'emploi* : le signal doit être à large bande ;
- *avantages* : méthode performante nécessitant très peu de connaissances *a priori* ; améliorant nettement le SNR (grâce à la modélisation du bruit) ;
- *inconvenients* : sensible aux erreurs de structure du modèle ;
- *complexité* : relativement complexe (FKE).

Détection de changements de variance (SNR)

1. Publications. Exemple

Publication originale

[1] F. KAUFFMANN (1987). Analysis of His bundle recordings. *IEEE Conf. Computers in Cardiology*, Louvain, Belgique, 12-15 Sept. 1987, pp. 345-348.

Exemples d'applications

Les techniques simples d'évaluation de rapport signal sur bruit sont efficaces dans de nombreux cas. Elles s'appliquent naturellement dans tous les cas où le rapport signal sur bruit (SNR) est élevé. Dans le domaine biomédical, le traitement de l'électrocardiogramme utilise couramment ces méthodes.

2. Objectifs

Cette méthode permet de détecter en ligne des changements de variance répétés sur l'entrée du modèle. Cette méthode s'apparente aux techniques dites d'évaluation de rapport signal sur bruit. On compare une estimation long terme et court terme de la variance du bruit.

3. Principe de la méthode

Hypothèses : On suppose qu'une modélisation du signal $Y(k)$ est ($\dim W = 2$; $\dim Y = 1$) :

$$\begin{cases} X(k+1) = FX(k) + GW(k) \\ Y(k) = HX(k) \end{cases} \quad (3.1)$$

avec :

$$\begin{aligned} W(k) &= (w_0(k), w_1(k) * h_1(k)) \\ h_1(k) &= 1 \quad \text{si } k \in H_1 \text{ (0 sinon)} \end{aligned}$$

où H_1 une partie de N composée d'une réunion d'intervalles $[a_i, b_i]$ et H_0 son complémentaire.

Soit $L(z) = (L_0(z), L_1(z))$ la fonction de transfert. On suppose qu'il existe une bande de fréquence $[f_0, f_1]$ dans laquelle le rapport $L_1(e^{i2\pi f})/L_0(e^{i2\pi f})$ est supérieur ou égal à K . On estime alors de deux manières différentes la puissance $V(k)$ du signal dans cette bande de fréquence. Une estimation long terme permettra d'estimer la variance $v_0(k)$ du bruit de la composante w_0 , alors qu'une estimation court terme permettra d'estimer la variance $v_1(k)$ de w_1 . La détection de l'événement $k \in H_1$ se fait grâce à trois tests sur le rapport de ces deux estimations (détection d'un début éventuel d'événement $k \in H_1$, validation, détection de la fin de l'événement).

4. Algorithme

Objets manipulés

Posons :

$$\begin{aligned} d &= \min \{b_i - a_i\} \\ D &= \max \{b_i - a_i\} \\ R &= \inf \{a_{i+1} - b_i\} \end{aligned} \quad (3.2)$$

d et D sont respectivement la durée minimale et la durée maximale des événements et R est la durée minimale entre deux événements.

Séquencement des opérations

Le déroulement du calcul se fait ainsi :

1. estimation de la puissance instantanée du signal $V(k)$
2. estimation long terme $v_0(k)$ de la variance de $w_0(k)$ ($k \in H_0$)
3. estimation court terme $v_1(k)$ de la variance de $w_1(k)$ ($k \in H_1$)
4. règle de décision à l'aide de deux tests $K > K_1 > 1$:
 $SNR(t) \leq K$ ou $\geq K_1$: estimation du début ou de la fin d'un événement.

Plus précisément :

1. On utilise un filtre passe-bande suivi d'une élévation au carré pour estimer la puissance instantanée $V(t)$ du signal dans la bande de fréquence $[f_0, f_1]$.
2. On estime la variance du bruit w_0 grâce à un filtre récursif du premier ordre (de constante de temps de l'ordre de $10D$), avec seuil, par les formules suivantes :

$$b(t) = av_0(t-1) + (1-a)V(t) \quad (3.3)$$

si $(1-x)v_0(t-1) \leq b(t) \leq (1+x)v_0(t-1)$

alors $v_0(t) = \max(v_{0-\min}, b(t))$

si $(1-x)v_0(t-1) \geq b(t)$

alors $v_0(t) = \max(v_{0-\min}, (1-x)v_0(t-1))$

si $b(t) \geq (1+x)v_0(t-1)$

alors $v_0(t) = \max(v_{0-\min}, (1+x)v_0(t-1))$

où x est un paramètre positif compris entre 0 et 1, et $v_{0-\min}$ est la valeur minimale pour l'estimation de v_0 . Les seuils adoptés permettent de borner les variations de v_0 .

3. L'estimation court terme v_1 de la variance de w_1 se fait

par un moyennage sur les L valeurs précédentes (L doit être choisi de l'ordre de d).

4. Décrivons maintenant l'étape de décision comportant 5 états. On teste le rapport des deux estimations

$SNR(t) = \frac{v_1(t)}{v_0(t)}$ ou rapport signal sur bruit par rapport à

deux seuils $1 < K_1 < K$:

état 1 : état réfractaire, aucune détection n'est possible, il dure R échantillons ;

état 2 : recherche d'un début éventuel d'événements si $SNR(t) > K_1$;

état 3 : validation de l'événement si $SNR(t) > K$;

état 4 : recherche de la fin de l'événement si $SNR(t) < K_1$ et si la durée de l'événement est comprise entre d et D ;

état 5 : sauvegarde de l'événement.

Les transitions entre états se font dans l'ordre ci-dessus.

5. Quelques caractéristiques de la méthode

Cette méthode de détection s'applique à tous les signaux ayant une plage de fréquence dans laquelle les signaux ont

des puissances différentes. Le cas le plus simple est la méthode d'évaluation du rapport signal sur bruit : bruit gaussien dans lequel on ajoute des événements de grande puissance. Il suffit de prendre $[0, 0,5]$ comme bande de fréquence. Dans le cas de signaux ayant une composante basse fréquence, les estimations des variances se font après filtrage passe-bas. Dans le cas particulier du signal électrocardiographique le meilleur rapport signal sur bruit a été trouvé dans la bande de fréquence $[3, 30]$ Hz.

Les estimations étant faites pour tous les échantillons, il n'y a pas de période de réinitialisation après la détection d'un début d'intervalle. En particulier cette méthode fonctionne bien dans le cas d'intervalles de courtes durées. De plus cette méthode permet de suivre des évolutions lentes de la variance du bruit v_0 .

6. Références complémentaires et remarques

On peut généraliser cette méthode dans le cas de plusieurs capteurs. Le calcul de la meilleure bande de fréquence commune à toutes les sorties, doit être fait sur toutes les composantes.

Algorithme de l'erreur de prédiction linéaire (EPL)

1. Publications. Exemples

Publication originale

[1] F. CASTANIE et P. SOULE (1983). Méthode de détection de non-stationnarités par analyse des propriétés de franchissement d'un seuil par l'erreur de prédiction linéaire. *Actes du 9^e colloque GRETSI*, pp. 105-109, Nice.

Points d'entrées bibliographiques conseillés

[2] E. DAYMIER et F. CASTANIE (1985). Analyse d'une méthode de détection des sauts de moyenne et de variance. *Actes du 10^e colloque GRETSI*, pp. 131-136, Nice.

Exemples d'applications

- mesures spatiales,
- EMG,
- CND de tubes de générateur de vapeur,
- réponse d'un pont au trafic routier,
- vibrations de moteur.

2. Objectifs

De type D, la segmentation par EPL permet de détecter des changements de paramètres d'un modèle AR excité par un bruit blanc gaussien, sur un signal centré.

3. Principe de la méthode

La méthode consiste à surveiller les innovations (ou EPL) calculées à partir d'une modélisation paramétrique (AR essentiellement) du signal. Les paramètres sont estimés par les moindres carrés exacts (algorithme de Morf).

Le modèle étant construit sur une tranche fixe du signal, la segmentation par EPL consiste à comparer l'EPL ε_n , de variance $\widehat{\sigma}^2$, à un seuil $S_\varepsilon = \lambda \widehat{\sigma}$, λ étant réel. On dit qu'il y a détection lorsque l'EPL ε_n dépasse S_ε , N_D fois consécutives. N_D est appelé *mot de dépassement*. Plus généralement, un « patron » binaire du mot de dépassement peut être utilisé pour la détection de non-stationnarités de type connu.

Actualisation : après la détection, on a le choix entre deux possibilités :

- continuer la segmentation sur le modèle construit au début ; dans ce cas, on doit désactiver le détecteur pendant quelques échantillons (par exemple $2p$) puis on repart, p étant l'ordre du modèle ;
- réactualiser le modèle ; dans ce cas, la nouvelle fenêtre est prise quelques échantillons après la détection précédente.

4. Algorithme

Un modèle AR d'ordre p d'un processus stochastique stationnaire Y_n est défini par :

$$Y_n = \sum_{k=1}^p a_k Y_{n-k} + E_n \quad (3.4)$$

où E_n est un bruit blanc discret de moyenne nulle et de variance σ^2 . Soit \widehat{Y}_n une estimée de Y_n constituée à partir des p derniers échantillons du signal $Y_{n-p}, Y_{n-p+1}, \dots, Y_{n-1}$:

$$\widehat{Y}_n = \sum_{k=1}^p \widehat{a}_k Y_{n-k} \quad (3.5)$$

L'EPL s'écrit alors :

$$\varepsilon_n = Y_n - \widehat{Y}_n \quad (3.6)$$

ou encore :

$$\varepsilon_n = Y_n - \sum_{k=1}^p \widehat{a}_k Y_{n-k} \quad (3.7)$$

Notons que si l'estimation est bonne, ε_n tend vers E_n avec une variance $\widehat{\sigma}^2 \approx \sigma^2$. Cette variance représente donc la puissance de l'EPL, elle doit être plus faible que la puissance du signal σ_Y^2 .

5. Quelques caractéristiques de la méthode

- La segmentation par l'EPL permet la détection des sauts de moyenne, de variance et de paramètres.
- Les valeurs à adopter pour S_ε et pour N_D dépendent de la nature de la non-stationnarité et de son impact sur le signal d'origine. Cependant, N_D restant généralement en dessous de 5 (valeurs typiques : 2, 3 ou 4), le seuil S_ε peut être fixé par la valeur de la *probabilité de fausse alarme* P_{FA} à ne pas dépasser. Celle-ci peut être approchée par :

$$P_{FA} = P(S_\varepsilon)^{N_D} \quad (3.8)$$

où :

$$P(S_\varepsilon) = \text{Prob} (|\varepsilon_n| > S_\varepsilon) \quad (3.9)$$

Sous l'hypothèse gaussienne pour ε_n , cette relation peut s'écrire :

$$P(S_\varepsilon) = \text{Erfc} \left(\frac{S_\varepsilon}{\widehat{\sigma} \sqrt{2}} \right) \quad (3.10)$$

Erfc étant la fonction erreur complémentaire.

- La méthode de l'EPL donne généralement de très bons résultats lorsque le rapport S/B de modélisation $\frac{\sigma_{\hat{Y}}^2}{\sigma^2}$ est assez élevé. Pour chaque type de signaux, on doit chercher le couple (S_ϵ, N_D) optimal.

6. Références complémentaires et remarques

La technique du mot de dépassement est similaire au « compteur » des algorithmes DF et SNR, mais plus exigeante puisqu'on considère des dépassements consécutifs. Voir aussi la fiche DF.

Détection, apprentissage et reconnaissance par paquets (DAREP)

1. Publications. Exemples

Publication originale

[1] D. DE BRUCQ et P. COURTELLEMONT (1987). Apprentissage sur le signal EEG de modélisations autorégressives modifiées. *L'Onde Électrique*, vol. 67, n° 4-5, pp. 69-73.

Exemples d'applications

- Détection de pointe-ondes, de spindles en EEG.

2. Objectifs

- Segmentation d'une observation de longue durée à l'aide d'une modélisation de type autorégressif du signal scalaire, supposé localement stationnaire, par paquets de k observations consécutives.
- Mise en œuvre d'une procédure de détection à double seuillage sur l'innovation, afin de permettre une reconnaissance de la forme la meilleure après détection de changement de modèle.

3. Principe de la méthode

Des modèles autorégressifs sont appris et mis en mémoire. Le modèle retenu est adapté d'un paquet à l'autre. La longueur $k \Delta T$ des paquets d'observations Y_t du signal échantillonné à ΔT , est choisie inférieure à la moitié de la durée de stationnarité locale.

A chaque paquet d'indice L , sont associés des coefficients : $\hat{A}(L) = (\hat{a}_1(L), \dots, \hat{a}_p(L))^T$ d'un modèle autorégressif. Pour $L = 1$ et pour tout apprentissage, ces coefficients sont calculés en résolvant les équations de Yule-Walker. Ensuite par récurrence, à partir des valeurs de $\hat{A}(L-1)$, les coefficients $\hat{A}(L)$ sont adaptés aux observations.

La détection éventuelle est repérée tout d'abord par un dépassement d'un seuil α de l'innovation ε_t tout comme dans l'algorithme EPL, innovation calculée à l'aide des coefficients du paquet précédent. Sous une hypothèse gaussienne, $P(|\varepsilon_t| > \alpha = 2,576) = 0,01$.

Ensuite l'énergie moyenne $\hat{\sigma}_{t+k-1}^2$ de l'innovation sur les k valeurs Y_t suivantes est calculée pour chacun des modèles autorégressifs conservés en mémoire. Si l'énergie moyenne la plus petite est inférieure à un seuil β^2 , le modèle autorégressif correspondant est conservé. Si elle est supérieure au seuil, un nouvel apprentissage est effectué pour aboutir à un nouveau modèle mémorisé.

La méthode est résumée par l'organigramme de la figure 1.

4. Algorithme

Objets manipulés

Données :

- p ordre du modèle autorégressif,
- k longueur des paquets,
- α seuil pour la détection,
- β^2 seuil pour la reconnaissance de formes,
- Y_1, \dots, Y_t observations scalaires ou vectorielles.

Procédures externes requises

Résolution des équations de Yule-Walker.

Séquencement des opérations

$$\begin{aligned} \hat{Y}_t &= \hat{a}_1 Y_{t-1} + \hat{a}_2 Y_{t-2} + \dots + \hat{a}_p Y_{t-p} \\ \varepsilon_t &= Y_t - \hat{Y}_t \\ \hat{\sigma}_t^2 &= \frac{1}{k} \sum_{s=t}^{t+k-1} \varepsilon_s^2 \end{aligned} \quad (3.11)$$

$$Z = \begin{pmatrix} Y_{kL} & \dots & Y_{kL-p+1} \\ \dots & \dots & \dots \\ Y_{kL-k+1} & \dots & Y_{kL-k-p+2} \end{pmatrix}$$

$$\mathcal{K}(L) = \rho \mathcal{K}(L-1) + \frac{1}{\hat{\sigma}_{kL}^2} \left(\frac{1}{k} Z^T Z \right)$$

avec $\rho \in]0, 1]$ et $\mathcal{K}(0) = 0$

$$ES = [\varepsilon_{kL-k+1}, \dots, \varepsilon_{kL}]$$

$$\hat{A}(L) = \hat{A}(L-1) + \frac{1}{\hat{\sigma}_{kL}^2} \frac{ES}{k} Z \mathcal{K}(L)^{-1}$$

5. Quelques caractéristiques de la méthode

- *complexité* : les diverses opérations peuvent être effectuées en parallèle à l'aide de plusieurs processeurs. Le nombre d'opérations est inférieur à un MCR classique ;
- *quelques caractéristiques* : la méthode constitue plus un canevas de traitement qu'un algorithme définitif : chaque partie (apprentissage, adaptation, détection et reconnaissance) peut être modifiée. La durée $k \Delta T$ des paquets et le coefficient ρ d'oubli doivent être convenablement choisis compte tenu de la stationnarité locale des phénomènes étudiés.

Une extension de l'algorithme à une situation multicapteur a été réalisée notamment en détection d'images sur fond bruité et en enregistrement multivoies d'EEG.

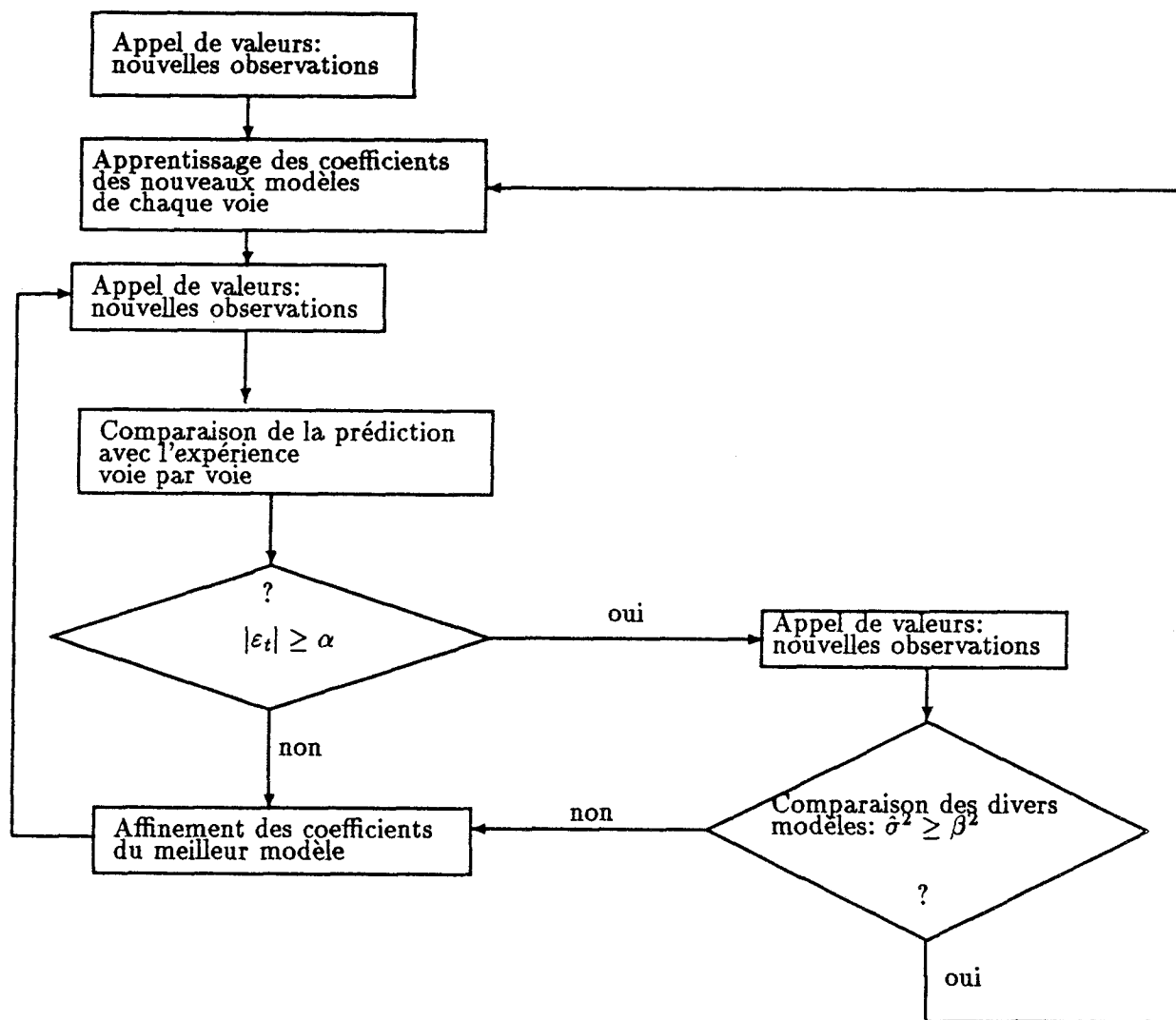


Figure 1. — Apprentissage et détection.

Distances entre deux fenêtres (DIS2f)

1. Publications. Exemples

Publication originale

- [1] G. BODENSTEIN and H. M. PRAETORIUS (1977). Feature extraction from the encephalogram by adaptive segmentation. *Proc. of the IEEE*, vol. 65, pp. 642-652.

Points d'entrées bibliographiques conseillés

- [2] M. MATHIEU (1976). Analyse de l'EEG par prédiction linéaire. Thèse de docteur ingénieur, ENST, Paris.
- [3] A. V. BRANDT (1983). An efficient GLR algorithm for adaptive segmentation of nonstationary signals. *Proc. of EUSIPCO*, Erlangen (RFA), September 12-16, pp. 763-766.
- [4] M. BASSEVILLE (1986). The two-models approach for the on-line detection of changes in AR processes. In *Detection of abrupt changes in signals and dynamical systems* (M. Basseville, A. Benveniste, eds.), chap. 6. Lecture Notes in Control and Information Sciences, LNCIS-77, Springer-Verlag, New York.

Exemples d'applications

- EEG,
- parole,
- EMG,
- vibrations de moteur,
- réponse d'un pont au trafic routier.

2. Objectifs

De type D, la méthode détecte en ligne des changements de paramètre AR.

3. Principe de la méthode

Le principe est celui d'une segmentation à deux modèles. La méthode consiste à construire une distance qui servira d'outil de comparaison entre deux fenêtres du signal. Cette distance est calculée à partir des caractéristiques des modèles AR correspondant aux deux fenêtres. Il y a détection lorsque cette distance dépasse un certain seuil fixé à l'avance.

Les deux fenêtres, de même taille L , sont positionnées suivant la figure 1 où :

- F_0 : fenêtre de référence fixe, choisie après un saut, sinon au début du signal,
- F_1 : fenêtre test glissante, de pas de glissement λ fixé au début de la segmentation. Cette fenêtre est initialement superposée à F_0 .

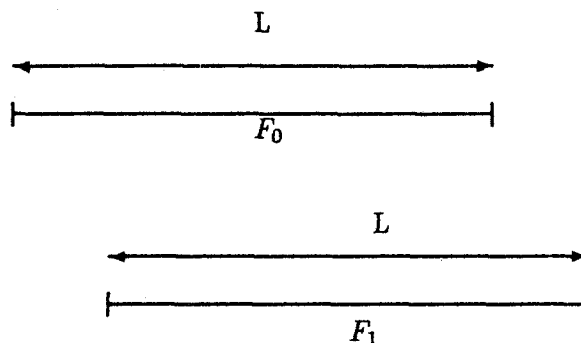


Figure 1. — Disposition des deux fenêtres : F_0 fixe, F_1 glissante.

4. Algorithme

- En général, le glissement de la fenêtre mobile est effectué échantillon par échantillon ($\lambda = 1$). Mais, dans le cas où le nombre d'échantillons du signal est très important, il est possible d'adapter le pas de glissement λ au type de la non-stationnarité (exemple d'une dérive lente).

Si N est le nombre d'échantillons du signal et D_{\max} le nombre maximal de points que peut prendre la distance (dans un logiciel, ce nombre correspond à la taille du tableau qui contient la distance), alors le pas minimal à fixer s'exprime par :

$$\lambda_{\min} = \frac{N - L}{D_{\max} + 1} + 1. \quad (3.12)$$

Dans tous les cas, la recherche de sauts de moyenne sur la distance globale permet de localiser les éventuelles non-stationnarités du signal : supposons qu'il y a eu saut de moyenne au point n_d de la distance, alors l'instant estimé de la non-stationnarité sera au point n_s du signal, avec $n_s = n_d \lambda + L$.

- La méthode d'identification utilisée dans les deux fenêtres est celle de Morf.
- La distance proposée en [3] est basée sur le logarithme du rapport de vraisemblance entre les deux fenêtres du signal à comparer. Cette distance est définie par le rapport des énergies des innovations (ou EPL) correspondant aux deux fenêtres :

$$d_{RV} = \frac{\sum_{i=1}^L (\varepsilon_i^1)^2}{\sum_{i=1}^L (\varepsilon_i^0)^2} - 1. \quad (3.13)$$

• Une autre distance possible est la **distance cepstrale** qui s'exprime en fonction des coefficients cepstraux. Le cepstre est défini comme la transformée de Fourier inverse du logarithme (naturel) du spectre $S(e^{i2\pi\tilde{f}})$. On a :

$$\ln [S(e^{i2\pi\tilde{f}})] = \sum_{k=-\infty}^{+\infty} c_k e^{-i2k\pi\tilde{f}} \quad (3.14)$$

où \tilde{f} est la fréquence normalisée, et c_k sont les coefficients cepstraux. Pour un modèle AR (1.9) d'ordre p , de puissance résiduelle σ^2 , la relation (3.14) devient :

$$\ln \left[\frac{\sigma^2}{\left| 1 - \sum_{k=1}^p a_k e^{-ik2\pi\tilde{f}} \right|^2} \right] = \sum_{k=-\infty}^{+\infty} c_k e^{-i2k\pi\tilde{f}}. \quad (3.15)$$

Si l'on exprime l'équation (3.15) en fonction des transformées en z du spectre et du cepstre, et si l'on dérive ensuite les deux termes de l'équation par rapport à la variable z^{-1} , on peut déduire facilement les coefficients cepstraux c_k en fonction des coefficients autorégressifs a_k :

$$\begin{aligned} c_0 &= \ln(\sigma^2) \\ c_1 &= -a_1 \\ c_{-k} &= c_k \\ kc_k &= -ka_k - \sum_{j=1}^{k-1} (k-j)c_{k-j}a_j \quad \text{pour } k = 2, \dots, p \\ kc_k &= -\sum_{j=1}^p (k-j)c_{k-j}a_j \quad \text{pour } k > p. \end{aligned} \quad (3.16)$$

La distance cepstrale s'écrit alors :

$$d_{\text{ceps}} = (c_0^0 - c_0^1)^2 + 2 \sum_{k=1}^{+\infty} (c_k^0 - c_k^1)^2 \quad (3.17)$$

où c_k^0 et c_k^1 correspondent à chacune des deux fenêtres à comparer.

On montre que l'on obtient une excellente approximation de d_{ceps} si on limite la sommation de l'équation (3.17) aux $2p$ premiers coefficients cepstraux.

5. Quelques caractéristiques de la méthode

- La méthode est non optimale.
- Formulation simple pour (3.13).
- Le paramètre λ influe sur le temps de calcul et la précision de la méthode.

6. Références complémentaires et remarques

- Comme on l'a indiqué en introduction, une variante possible consiste à modifier le type de fenêtre. En [4], la distance cepstrale a été utilisée pour une fenêtre de référence de taille croissante et une fenêtre test glissante de taille fixe comme ici, avec des algorithmes d'identification différents. Voir la fiche DIV pour des précisions et une comparaison avec cet algorithme. En [3], le rapport de vraisemblance est calculé avec une fenêtre croissante aussi ; voir la fiche Brandt.
- Sur les problèmes de distance :
 - [5] A. H. GRAY and J. D. MARKEL (1976). Distances measures for speech processing. *IEEE Trans. Acoust. Speech. Sig. Process.*, ASSP-24, n° 5, pp. 380-391.
 - [6] M. BASSEVILLE (1989). Distance measures for signal processing and pattern recognition. *Signal Processing*, vol. 18, pp. 349-369.

Changements de dynamique d'un ARMA (FKE-D)

1. Publications. Exemples

Publication originale

[1] C. DONCARLI (1983). Un algorithme séquentiel de détection des ruptures d'un modèle ARMA. *GRETSI*.

Exemples d'applications

- Détection de phases d'activité en biomédical (EMG) ;
- Contrôle non destructif en temps réel ;
- Détection de manœuvres de cibles.

2. Objectifs

- Détection séquentielle des ruptures spectrales d'un ARMA scalaire.
- Amélioration des capacités de poursuite des algorithmes d'identification adaptatifs.

3. Principe de la méthode

- Estimation des coefficients de l'ARMA par filtre de Kalman étendu.
- Détection par test récursif de la blancheur de la pseudo-innovation (double filtrage des produits décalés).
- Actualisation par réinitialisation du filtre de Kalman et inhibition temporaire du détecteur.

4. Algorithme

Objets manipulés

Signal mesuré y_k régi par :

$$y_k = \sum_{i=1}^p a_i y_{k-i} + \sum_{i=1}^p b_i E_{k-i} + E_k \quad (3.18)$$

avec :

E_k séquence blanche de variance inconnue R , a_i et b_i inconnus, constants par morceaux, ordre p connu.

Objectifs

Estimer a_i , b_i et détecter si k est un instant de rupture.

Procédures externes requises

Identification récursive par filtre de Kalman étendu (FKE).

Séquencement des opérations

1. Identification récursive par FKE

Si k n'est pas un instant de rupture

$$y_k = \sum_{i=1}^p a_i y_{k-i} + \sum_{i=1}^p b_i E_{k-i} + E_k \quad (3.19)$$

On définit un état étendu :

$$X_k^T = [a_1, \dots, a_p, b_1, \dots, b_p, E_{k-1}, \dots, E_{k-p}] \quad (3.20)$$

Equation d'état linéaire

$$X_{k+1} = F X_k + G E_k \quad (3.21)$$

avec

$$F = \begin{pmatrix} I & 0 \\ 0 & J \end{pmatrix}$$

$I =$ matrice unité $2p \times 2p$
 $J =$ matrice de décalage $p \times p$
 $G = [0, \dots, 0, 1, 0, \dots, 0]$.

L'équation de mesure (3.19) non linéaire est linéarisée autour de la prédiction courante de l'état étendu :

$$Z_k = H_k X_k + E_k \quad (3.22)$$

avec :

$$H_k = [y_{k-1}, \dots, y_{k-p}, E_{k-1|k-1}, \dots, E_{k-p|k-1}, b_{1|k-1}, \dots, b_{p|k-1}] \quad (3.23)$$

$$Z_k = y_k + \sum_{i=1}^p b_i |_{k-1} E_{k-i|k-1}$$

Les équations de Kalman (procédure externe standard) sont ensuite appliquées au système linéarisé non stationnaire, avec une estimation récursive de R :

$$R_{k|k-1} = u_k R_{k-1|k-2} + (1 - u_k) \varepsilon_k^2 \quad (3.24)$$

où $0 < u_k < 1$ et :

$$\varepsilon_k = E_{k|k-1} = Z_k - H_k X_{k|k-1} \quad (3.25)$$

et initialisée par la valeur instantanée ε_1^2 .

2. Détection : parallèlement, on calcule :

$$W_{k|k-1} = u_k W_{k-1|k-2} + (1 - u_k) \varepsilon_k \varepsilon_{k-1}$$

$$S_{k|k-1} = v_k S_{k-1|k-2} + (1 - v_k) \varepsilon_k \varepsilon_{k-1} \quad (3.26)$$

avec $0 < v_k < u_k < 1$, puis :

$$T_k = \frac{S_{k|k-1}}{W_{k|k-1}} \quad (3.27)$$

u_k et v_k seront définis plus loin. Le principe de la détection consiste à comparer T_k à un seuil. En effet, s'il y a eu un saut, la pseudo-innovation n'est plus blanche et $T_k \gg 1$ car l'horizon de calcul de $S_{k|k-1}$ est plus court que celui de $W_{k|k-1}$. Inversement, s'il n'y a pas de saut, alors $T_k \approx 1$. Des valeurs habituelles du seuil peuvent être 3 ou 4.

3. Actualisation (et initialisation) : si un saut est détecté à l'instant k (ou pour $k = 0$), on conserve la valeur courante des paramètres, mais on réinitialise les éléments corres-

pondants de la matrice de variance-covariance d'erreur d'estimation.

Enfin, on inhibe le détecteur pendant quelques points pour éviter les alarmes multiples :

$$u_k = v_k = v$$

$$u_{k+i} = 0,5(u + u_{k+i-1}) \text{ avec } v = 0,5 \text{ et } u = 0,96$$

$$v_{k+i} = v.$$

5. Quelques caractéristiques de la méthode

- *précautions d'emploi* : le signal doit rester à bande large ;

- *avantages* : méthode performante nécessitant très peu de connaissances *a priori* ;
- *inconvenients* : sensible aux erreurs de structure du modèle ;
- *complexité* : relativement complexe (FKE).

6. Références complémentaires et remarques

D'autres variables de décision peuvent être intégrées dans ce schéma (variables S et W normalisées par R) ainsi que d'autres tests (CUSUM sur les variables normalisées, voir fiche HK).

Détecteur dit de Divergence-Hinkley (DIV)

1. Publications. Exemples

Publication originale

- [1] M. BASSEVILLE and A. BENVENISTE (1983). Sequential detection of abrupt changes in spectral characteristics of digital signals. *IEEE Trans. Inf. Theory*, IT-29, pp. 709-724.

Points d'entrées bibliographiques conseillés

- [2] M. BASSEVILLE (1986). The two-models approach for the on-line detection of changes in AR processes. In *Detection of abrupt changes in signals and dynamical systems* (M. Basseville, A. Benveniste, eds.), chap. 6. Lecture Notes in Control and Information Sciences, LNCIS-77, Springer-Verlag, New York.
- [3] R. ANDRÉ-OBRECHT (1988). A new statistical approach for the automatic segmentation of continuous speech signals. *IEEE Trans. Acoust. Speech. Sig. Process.*, ASSP-36, pp 29-40.

Exemples d'applications

- segmentation automatique du signal de parole continue ;
- segmentation automatique de signaux biomédicaux (ECG, EEG).

2. Objectifs

L'algorithme dit de divergence-Hinkley est de type D et aborde le problème de la segmentation spectrale de signaux centrés par le biais de la détection en-ligne de changements dans les paramètres d'un modèle AR excité par un bruit blanc gaussien.

3. Principe de la méthode

Cet algorithme est du type général « comparaison entre deux modèles estimés à des endroits différents dans le signal ». Le test calcule une « distance » — la divergence de Kullback, d'où son nom — entre 2 modèles AR « long terme » et « court terme » estimés sur le signal. Cette distance est un indicateur qui réagit à un changement de caractéristiques spectrales du signal par un changement de sa dérive, qui de nulle devient positive. Ce dernier changement est alors détecté par le détecteur de sauts de moyenne dit CUSUM de Page-Hinkley.

4. Algorithme

Objets manipulés

Le modèle AR (1.9) de l'introduction.

Procédures externes requises

Les paramètres autorégressifs sont estimés d'une part à long terme dans une fenêtre de taille croissante (tous les

points du signal depuis le début ou depuis la dernière alarme), d'autre part à court terme dans une fenêtre mobile de taille fixe. Les algorithmes d'estimation dans chacune de ces deux fenêtres sont des *filtres en treillis* implémentant des *méthodes de moindres carrés approchés* pour l'estimation de coefficients autorégressifs : une implémentation séquentielle de l'algorithme de Burg décrite en [2] d'une part, et la méthode dite d'autocorrélation d'autre part.

Séquencement des opérations

A chaque instant n , on estime les 2 modèles AR et on calcule la distance mentionnée plus haut. Le détecteur qui en résulte se résume en la formule (non triviale pour l'intuition !) :

$$w_n = w_{n-1} - \left(\frac{\varepsilon_n^0 \varepsilon_n^1}{\hat{\sigma}_1^2} \right) + \left(1 + \frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} \right) \times \left(\frac{(\varepsilon_n^0)^2}{2 \hat{\sigma}_0^2} - \frac{1}{2} + \frac{\hat{\sigma}_0^2}{2 \hat{\sigma}_1^2} \right) \quad (3.28)$$

où ε_n^0 et ε_n^1 sont les innovations des deux filtres long terme et court terme, et $\hat{\sigma}_0^2$ et $\hat{\sigma}_1^2$ leurs variances estimées.

Comme il est dit plus haut, ce détecteur réagit à une rupture spectrale par un changement de sa dérive qui de nulle devient positive ; on peut alors raffiner la détection et l'estimation de l'instant de rupture en utilisant l'algorithme CUSUM (voir la fiche HK). Il s'agit donc d'introduire encore une sorte de dérive minimum δ ; et de choisir un seuil λ . Le détecteur qui en résulte est alors :

$$W_0 = 0 \quad (3.29)$$

$$W_n = \sum_{k=1}^{k=n} (w_k - \delta) \quad (3.30)$$

$$m_n = \min_{0 \leq k \leq n} W_k \quad (3.31)$$

$$\text{détection quand } W_n - m_n \geq \lambda \quad (3.32)$$

Comme pour les changements additifs, cette notion d'amplitude minimum qui, au premier abord rend l'algorithme plus complexe (il y a 2 paramètres à régler — amplitude minimum et seuil — au lieu d'un seuil unique), est en fait un moyen très puissant pour le rendre *robuste au bruit, au changement de locuteur en parole, ...* L'instant de rupture estimé est le dernier instant où le minimum a été atteint (voir la fiche HK).

5. Quelques caractéristiques de la méthode

- *précautions d'emploi* : après chaque détection, il faut « geler » le détecteur pendant une durée égale à la longueur de la fenêtre d'estimation du deuxième modèle

afin de réinitialiser les 2 filtres. D'autre part, après une détection, on peut choisir de redémarrer soit à l'instant de détection, soit — pour augmenter les performances — à l'instant de rupture estimé (qui lui est antérieur, et donc dans ce cas on ne respecte pas tout à fait le temps réel) ;

- *avantages* : l'usage de méthodes d'identification exactes n'est pas nécessaire pour la mesure de distance choisie entre les deux modèles estimés, ce qui donne un algorithme moins coûteux. D'autre part, cet algorithme est le noyau de base de la procédure de segmentation automatique du signal de parole continue développée à l'IRISA par R. André-Obrecht ; *le réglage des paramètres de cet algorithme de segmentation est indépendant du locuteur* [3] et *robuste au bruit* (ce dernier point a été vérifié sur des enregistrements radio-mobiles). Cet algorithme de segmentation fournit — en tout cas sur du signal de parole — des résultats déjà intéressants même avec un ordre autorégressif faible ($p = 2$) par rapport à l'ordre $p = 16$ plus réaliste du point de vue de la reconnaissance ;

- *inconvenients* : le principal est sa complexité ;

- *complexité* : des indications chiffrées sont données en [3]. Le coût de l'algorithme est lié à la présence de 2 filtres. Il est à noter cependant que ce détecteur fournit déjà des résultats intéressants avec un ordre AR égal à 2, et que dans ce cas sa complexité est faible.

6. Références complémentaires et remarques

Ainsi qu'on l'a indiqué dans les variantes possibles des algorithmes de type *D*, on peut concevoir un algorithme de segmentation spectrale similaire en conservant tous les ingrédients de *DIV* et en ne modifiant que la règle de décision (3.28), que l'on peut par exemple remplacer par la *distance cepstrale* — calculée par la formule (3.17) de la fiche *DIS2f*. On a constaté en [1] sur des signaux simulés que l'algorithme qui en résulte est moins fin que *DIV*.

D'autre part, on renvoie à [DiFra90] pour une utilisation et une variante de *DIV* pour le codage de parole.

Enfin, une comparaison avec l'algorithme *Brandt* sur des signaux de parole continue est décrite en [3].

Algorithme de Brandt (Brandt)

1. Publications. Exemples

Publication originale

[1] A. V. BRANDT (1983). An efficient GLR algorithm for adaptive segmentation of non-stationary signals. *Proc. of EUSIPCO*, pp. 763-766, Erlangen (RFA).

Points d'entrées bibliographiques conseillés

[2] U. A. APPEL and A. V. BRANDT (1984). Performance comparison of two segmentation algorithms using growing reference windows. *Proc. 6th Int. Conf. on Analysis and Optimisation of Systems*, pp. 156-169, Nice, June 19-22.

Exemples d'applications

- EEG,
- parole,
- CND de tubes de générateur de vapeur,
- vibrations de moteur,
- réponse d'un pont au trafic routier.

2. Objectifs

De type D, cet algorithme détecte des changements de paramètres d'un modèle autorégressif excité par un bruit blanc gaussien.

3. Principe de la méthode

L'algorithme de Brandt est un moyen astucieux d'implémenter la méthode du rapport de vraisemblance généralisé en rupture de modèle. Basé sur la modélisation AR, cet algorithme adaptatif construit à chaque itération une distance à partir de trois fenêtres du signal. Une fois qu'un dépassement d'un seuil S_G par cette distance est détecté, la méthode effectue une optimisation pour l'estimation de l'instant de rupture.

Les trois fenêtres prises sur le signal pour effectuer la modélisation AR sont définies comme suit :

- F_1 : fenêtre de référence à mémoire croissante de longueur L_{F_1} ,
- F_2 : fenêtre test glissante de longueur L_{F_2} ,
- F_3 : fenêtre totale regroupant F_1 et F_2 , de longueur L_{F_3} .

La figure 1 illustre le passage de l'itération i à l'itération $(i + 1)$ dans la phase de recherche de non-stationnarité. Chaque itération de l'algorithme est spécifiée par :

- un point t , premier point de la fenêtre F_2 , considéré comme la limite entre les fenêtres courantes F_1 et F_2 ,

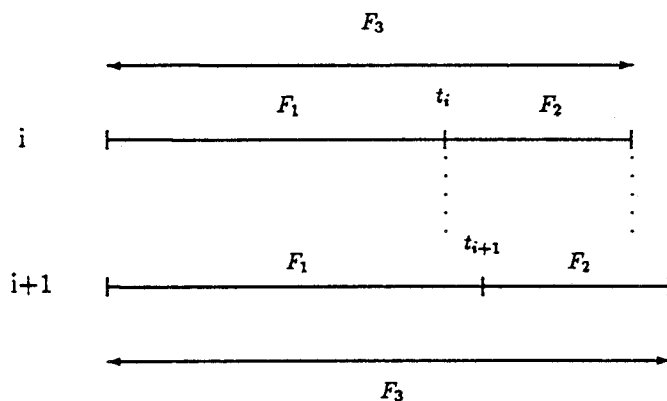


Figure 1. — Recherche de la non-stationnarité.

- le rapport de vraisemblance au point t , appelé aussi métrique, et noté m_t :

$$m_t = H_{F_3} - H_{F_2} - H_{F_1} \quad (3.33)$$

où $H_{F_i} = L_{F_i} \ln(\hat{\sigma}_{F_i}^2)$ représente la quantité de vraisemblance spécifique à la fenêtre F_i , $\hat{\sigma}_{F_i}^2$ étant la puissance résiduelle. m_t représente alors le double du logarithme du rapport de vraisemblance entre la fenêtre F_3 d'une part et les deux fenêtres F_1 et F_2 d'autre part.

Le test du GLR consiste alors à comparer m_t à un seuil S_G fixé arbitrairement au début de l'algorithme. Si $m_t > S_G$, il y a détection, et donc, la non-stationnarité se trouve, *a fortiori*, dans la fenêtre test courante $F_2 = [t, t + L_{F_2} - 1]$. Il sera intéressant de chercher l'instant précis où a lieu cette non-stationnarité, procédure que l'on appellera *optimisation de l'estimation*.

Supposons que la détection a eu lieu à l'instant t_0 avec une métrique m_{t_0} . Une détection optimale à l'instant t_{opt} , doit avoir lieu dans l'intervalle $F_2 = [t_0, t_0 + L_{F_2} - 1]$. La phase d'optimisation est initialisée en prenant $t_{opt} = t_0$. Dans cette phase :

- la fenêtre de référence F_1' devient de longueur fixe L_{F_1} ,
- la fenêtre test devient à mémoire croissante F_2' de longueur L_{F_2} .

La figure 2 illustre une incrémentation dans la boucle d'optimisation. L'optimisation consiste, donc, à chercher l'instant $t \in [t_0, t_0 + L_{F_2} - 1]$ tel que : $m_t > m'_{t_{opt}}$ avec : $m'_{t_{opt}} = H_{F_3} - H_{F_2'} - H_{F_1'}$.

- Comme les fenêtres F_3 et F_3' restent de même longueur, le test d'optimisation devient :

$$H_{F_2'} + H_{F_1'} - H_{F_2} - H_{F_1} > 0 .$$

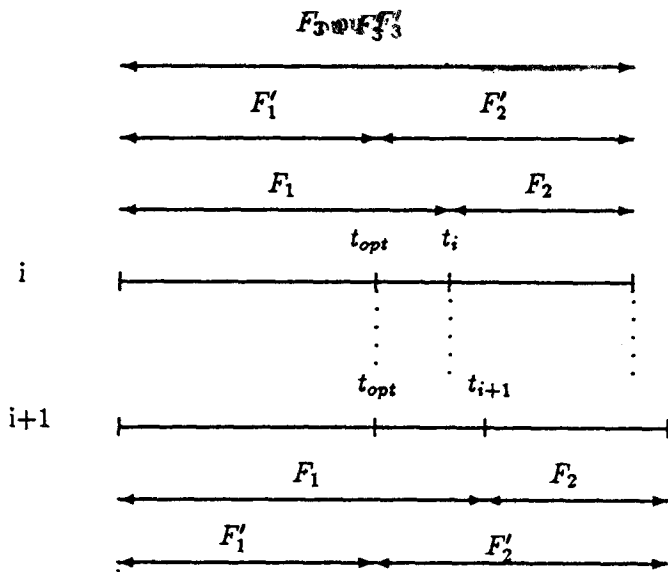


Figure 2. — Optimisation de l'instant de rupture estimé.

Une fois que l'on trouve t , on actualise $t_{opt} = t$ et on continue la phase d'optimisation. Cette phase doit s'arrêter lorsqu'on atteint $t_0 + L_{F_2} - 1$. Le dernier t_{opt} trouvé sera donc considéré comme l'instant optimal où il y a eu non-stationnarité.

A l'instant $t_{opt} + 1$, on réinitialise l'algorithme du GLR pour rechercher d'autres éventuelles non-stationnarités.

4. Algorithme

L'organigramme de la figure 3 schématise les différentes étapes de l'algorithme de segmentation du GLR, N étant le nombre d'échantillons du signal.

5. Quelques caractéristiques de la méthode

- *précautions d'emploi* : le choix approprié du seuil S_G et des tailles des fenêtres présente parfois certaines difficultés,

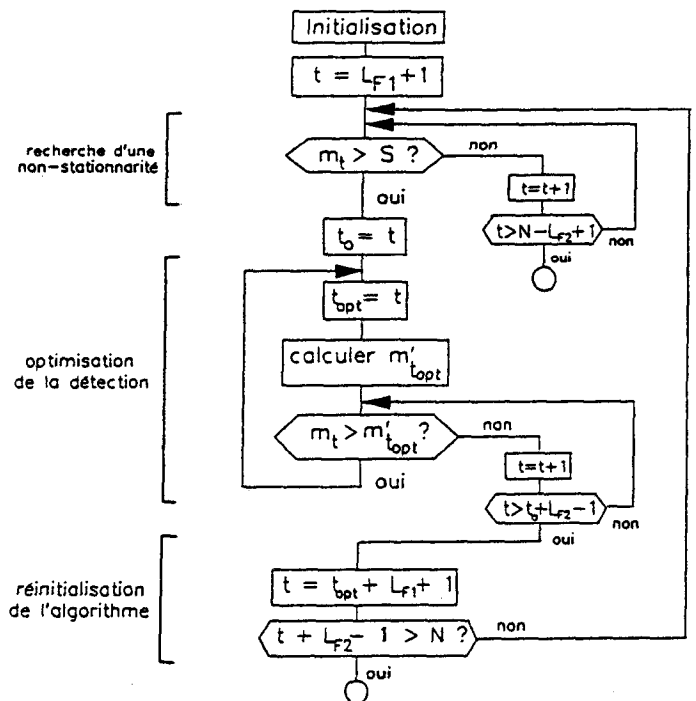


Figure 3. — Algorithme du GLR.

- *avantages* : méthode optimale théoriquement,
- *inconvenients* : le temps de calcul est important,
- *complexité* : 3 modèles AR à calculer à chaque pas.

6. Références complémentaires et remarques

Une comparaison avec l'algorithme DIV a été faite en [3] sur des signaux de parole continue.

- [3] R. ANDRÉ-OBRECHT (1988). A new statistical approach for the automatic segmentation of continuous speech signals. *IEEE Trans. Acoust. Speech. Sig. Process.*, ASSP-36, pp. 29-40.

Détection de motifs par gabarit (GAB)

1. Publications. Exemples

Publication originale

- [1] F. KAUFFMANN (1988). A study of nonstationary phenomena of HRV during 24 hours ECG ambulatory monitoring. *IEEE Conf. Computers in Cardiology*, Washington D.C., 25-28 Sept. 1988, pp. 305-306.

Exemples d'applications

- Détection d'oscillations dans des tachocardiogrammes ;
- Détection de pointes-ondes en EEG.

2. Objectifs

Cette méthode permet de détecter des motifs simples d'amplitudes et de durées variables dans un signal peu bruité.

3. Principe de la méthode

Un gabarit décrivant la forme du motif recherché est décrit par un ensemble de paramètres (amplitude minimale et maximale, durée minimale et maximale, ...). La méthode consiste alors à trouver les tranches de signal qui satisfont ce gabarit.

4. Algorithme

Objets manipulés

Description d'un gabarit : nous avons choisi des motifs de type oscillation. Un gabarit composé est constitué de trois gabarits simples. Un gabarit simple permet de détecter une sous-suite du signal presque monotone au sens large. Il est déterminé par le pourcentage minimal de points variant dans le sens recherché, par sa durée minimale et maximale, ainsi que son amplitude minimale et maximale.

Procédures externes requises

Fonction utilisée :

`gabarit_simple (t_debut, t_recherche, long_max, ok)`

Cette fonction calcule le dernier instant dans l'intervalle $[0, t_recherche]$ tel que la sous-suite de $t_recherche$ à t_debut satisfasse le gabarit de longueur maximale

`long_max`. A partir du signal original $s(t)$, on calcule deux autres signaux $1_+(t)$, $1_-(t)$ qui sont les valeurs logiques des tests de monotonie :

$$\begin{aligned} 1_+(t) &= 1 \text{ si } s(t) \geq s(t-1) \text{ (0 sinon)} \\ 1_-(t) &= 1 \text{ si } s(t) \leq s(t-1) \text{ (0 sinon)}. \end{aligned} \quad (4.1)$$

On somme alors simplement les signaux 1_+ ou 1_- dans l'intervalle courant jusqu'à obtenir le bon pourcentage.

Séquencement des opérations

A chaque instant une recherche arrière est faite :

1. Initialiser les durées maximales et minimales des gabarits simples.
2. Recherche du premier gabarit simple de longueur minimale à partir du point courant. S'il n'existe pas, passer au point courant suivant.
3. Recherche du deuxième gabarit simple de longueur minimale à partir du premier point du gabarit précédent. S'il n'existe pas, continuer la recherche du premier gabarit en augmentant sa longueur minimale.
4. Recherche du troisième gabarit à partir du premier point du gabarit précédent. S'il n'existe pas, continuer la recherche gabarit en augmentant sa longueur minimale.
5. Si un motif a été trouvé, alors la recherche se poursuit après la moitié de la longueur du motif trouvé.

5. Quelques caractéristiques de la méthode

- *Complexité* : l'algorithme est de grande complexité. Pour chaque échantillon, le traitement fait est de l'ordre du produit des longueurs des gabarits simples tests. Ainsi la longueur des motifs doit être courte. Les oscillations doivent être dans des bandes de hautes fréquences. Dans nos applications les motifs recherchés après pré-traitement avaient des périodes de l'ordre de 5 à 10 échantillons.

- *Pré-traitements* : la méthode est sensible au bruit de haute fréquence par rapport aux fréquences du motif recherché. Deux méthodes peuvent être utilisées. La première consiste à augmenter le pourcentage maximal de points variants dans le sens opposé. La deuxième consiste à faire un filtrage passe-bas.

Afin d'éviter des temps de calcul prohibitifs, filtrage et décimation sont utilisés pour placer la bande de fréquence du motif recherché dans les hautes fréquences. La précision de la détection est alors diminuée.

Mesure de forme par corrélation (MFC)

1. Publications. Exemples

Publication originale

- [1] A. J. BERNI, D. E. DICK and M. W. LUTTGES (1975). Detection of digitalis toxicity by computerized electrocardiogram monitoring. *IEEE Trans. Biom. Eng.*, vol. 22, pp. 29-34.

Points d'entrées bibliographiques conseillés

- [2] J. M. AGNIER, B. ROUVRAIS, J. L. COATRIEUX, P. Y. ARQUES et R. LE BARS (1981). Présentation de méthodes de reconnaissance des signaux d'électro-myographie. *GRETSI VIII*, pp. 551-556.
- [3] H. RIX and E. THIERRY (1989). Comparison of two similarity criterions in shape classification of transient signals : application to high resolution electrocardiography. *AIPAC'89, IFAC Symposium*, vol. 2, pp. 304-307.

Exemples d'applications

- signaux biomédicaux.

2. Objectifs

L'algorithme MFC (Mesure de Forme par Corrélation) permet de mesurer une différence de forme indépendamment d'un changement d'échelle sur les abscisses et sur les ordonnées. Il est donc du type F.

3. Principe de la méthode

Soit $s(t)$ et $v(t')$ deux signaux échantillonnés. On calcule la valeur maximale dans le plan (β, τ) du coefficient de corrélation entre $s(t)$ et $v(\beta(t' - \tau))$ où τ est le retard de v par rapport à s et où le facteur d'échelle β représente le rapport des largeurs entre s et v .

Pour calculer $v(\beta(t' - \tau))$, il est nécessaire de faire une interpolation entre les instants d'échantillonnage à partir d'un modèle ajusté sur le signal.

Cette méthode peut faire penser au filtrage adapté où l'on cherche à détecter la position d'un signal $s(t)$ noyé dans un bruit blanc en utilisant un gabarit $v(t) = s(t - \tau)$: dans ce cas $\beta = 1$ et on cherche seulement le paramètre τ qui rend maximale la corrélation de $v(t)$ avec le signal bruité. Dans l'algorithme MFC, on estimera la différence de forme entre les deux signaux en mesurant la corrélation lorsqu'elle est maximale par rapport à la fois à τ et à β .

4. Algorithme

Objets manipulés

- un signal $s(t_i)$ ($i = 1, \dots, N$),

- un signal de référence $v(t'_j)$ ($j = 1, \dots, M$),
- la fenêtre pour le calcul du coefficient de corrélation,
- l'indice définissant le choix de la procédure d'interpolation.

Modélisation du signal de référence pour l'interpolation

- modélisation par séries de Fourier,
- modélisation par fonctions splines.

Calcul du coefficient d'intercorrélation

$$u(t) = v(\beta(t - \tau)) \quad (4.2)$$

$$\rho(s, u) = \frac{\int_0^T [s(t) - E(s(t))][u(t) - E(u(t))] dt}{\left[\int_0^T [s(t) - E(s(t))]^2 dt \int_0^T [u(t) - E(u(t))]^2 dt \right]^{1/2}} \quad (4.3)$$

$$d(s, v) = 1 - \max_{(\beta, \tau)} \rho(s, u). \quad (4.4)$$

Pour la recherche du maximum, on peut envisager toute méthode d'optimisation à deux variables (dichotomie, gradient, ...) en prenant garde à l'existence possible de maxima secondaires indésirables.

5. Quelques caractéristiques de la méthode

Différence de forme et de largeur

Si l'application nécessite la mesure d'un écart combinant une différence de forme et une différence de largeur, on imposera un $\beta = 1$.

Influence de la modélisation

L'erreur de modélisation réalisée pour l'interpolation se répercute sur la mesure de la différence de forme. Il peut être utile d'en tenir compte au cours des traitements ultérieurs (prise de décision, classification...).

Réglage des paramètres

- Choix de la modélisation : la modélisation dépend de l'application. Par exemple, en électrocardiographie, il a été montré que pour modéliser le complexe QRS, 6 termes de la série de Fourier suffisaient pour obtenir une erreur de 5 % [1].
- Choix du domaine de variation des paramètres β , τ et du pas de discrétisation : c'est à l'utilisateur de définir ces différents paramètres.

Méthode des fonctions de répartition (MFR)

1. Publications. Exemples

Publication originale

- [1] H. RIX and J. P. MALENGE (1980). Detecting small variations in shape. *IEEE Trans. Syst. Man. Cybern.*, vol. SMC-2, pp. 90-96.

Points d'entrées bibliographiques conseillés

- [2] H. RIX and J. P. MALENGE (1980). Detection of impurity (1 %) at low resolution (0.25). *J. High Resolut. Chromatogr. Commun.*, vol. 4, pp. 172-176.
- [3] H. RIX (1986). Le traitement des formes en chimie analytique. *Anal. Chimica Acta*, vol. 191, pp. 467-472.
- [4] S. JÉSUS (1986). Estimation d'un signal répétitif bruité par sommation synchrone et lissage adaptatif : application à la structure fine du signal cardiaque. Thèse de doctorat, Nice.
- [5] E. THIERRY (1989). Analyse de forme des signaux monodimensionnels : étude théorique et application à la structure fine du signal électrique cardiaque. Thèse de doctorat, Nice.
- [6] H. RIX and E. THIERRY (1989). Comparison of two similarity criterions in shape classification of transient signals : application to high resolution electrocardiography. *AIPAC'89, IFAC Symposium*, vol. 2, pp. 304-307.

Exemples d'applications

- Caractérisation des formes de pics chromatographiques et détection d'impuretés ;
- Spectrométrie Mössbauer ;
- Séparation de cibles radar ;
- Électrocardiographie, signaux biomédicaux.

2. Objectifs

L'algorithme MFR (Méthode des Fonctions de Répartition) permet de donner une **mesure de la différence de forme** entre deux signaux monodimensionnels positifs, intégrables et à support borné : **on suppose donc que les signaux ont été segmentés au préalable**. Cet algorithme est donc du type F.

Un tel algorithme est particulièrement utile dans l'analyse des signaux transitoires répétitifs (ECG, potentiels évoqués, ...) pour segmenter des trains de signaux en détectant un changement de forme des signaux.

Il permet également de faire de la classification des formes puisqu'il fournit une « distance » entre les signaux.

3. Principe de la méthode

L'idée de base est de considérer les signaux positifs à comparer $s(t)$ et $v(t)$, une fois normés par leurs intégrales,

comme les densités de probabilité de deux v.a. X et Y et de rechercher la transformation Φ telle que $Y = \Phi(X)$. Les intégrales une fois normées sont donc des fonctions de répartition S et V : la fonction Φ_{sv} permettant de comparer v à s s'exprime par :

$$\Phi_{sv}(t) = (V^{-1} \circ S)(t) \quad (4.5)$$

où V^{-1} désigne la fonction réciproque de V.

Deux signaux sont dits de même forme s'ils se déduisent l'un de l'autre par translation et changement d'échelle sur les ordonnées et les abscisses. L'égalité de forme se traduit alors par le fait que Φ_{sv} est une fonction affine : toute mesure de l'écart à la linéarité de Φ_{sv} est donc une mesure de la variation de forme de $v(t)$ par rapport à $s(t)$.

La méthode proposée est globale en ce sens qu'elle tire une mesure de différence de forme de la totalité des signaux et non de paramètres estimés à partir de ceux-ci ; de plus elle ne nécessite aucune modélisation *a priori* des signaux.

4. Algorithme

Objets manipulés

- 2 signaux *positifs* à comparer ($s(t_i)$, $i = 1, \dots, N$, $v(t'_j)$, $j = 1, \dots, M$),
- la fenêtre d'intégration,
- la fenêtre de calcul de l'écart,
- l'indice définissant le choix de l'écart.

Procédures externes requises

- segmentation des signaux à comparer,
- élimination de la ligne de base (filtrage ou autre),
- obtention d'un signal positif :
 - valeur absolue,
 - partie positive,
 - carré ;
- utilisation des écarts de forme dans un algorithme de classification des signaux selon leur forme. Par exemple : nuées dynamiques, classification hiérarchique ascendante.

Intégration

Évaluation par une méthode numérique choisie (trapèzes, Simpson, ...) des fonctions de répartition :

$$S(t_i) = \frac{\int_{t_1}^{t_i} s(\tau) d\tau}{\int_{t_1}^{t_N} s(\tau) d\tau} \quad (4.6)$$

$$V(t'_j) = \frac{\int_{t'_1}^{t'_j} v(\tau) d\tau}{\int_{t'_1}^{t'_M} v(\tau) d\tau} \quad (4.7)$$

Calcul de la fonction $\Phi_{sv}(t)$

• Option « pas constant en ordonnées » : on divise l'intervalle $[0, 1]$ en P points équidistants (couramment $P = 100$) et on recherche, en faisant si nécessaire une interpolation linéaire, les valeurs t_h et t'_h définies par :

$$S(t_h) = V(t'_h) = \frac{h}{P}, \quad h = 1, \dots, P. \quad (4.8)$$

On récupère la fonction Φ_{sv} échantillonnée $\Phi_{sv}(t_h) = t'_h$.

• Option « pas constant en abscisse » : on part de P valeurs de $t(t_1, \dots, t_p)$ équidistantes et on détermine les valeurs (t'_1, \dots, t'_p) telles que :

$$S(t_h) = V(t'_h), \quad h = 1, \dots, P. \quad (4.9)$$

Remarque : on peut, pour éviter des interpolations, prendre $P = N$ et utiliser l'échantillonnage de $S(t)$.

On récupère la fonction Φ_{sv} échantillonnée $\Phi_{sv}(t_h) = t'_h$.

Calcul de l'écart de v par rapport à s : $d(s, v)$

- Choix de l'écart parmi n options ; voir ci-après ;
- Voir le sous-programme correspondant (écart n° i) permettant de mesurer l'écart à la linéarité de la courbe Φ .

Symétrisation

• Transposer s et v et calculer $d(v, s)$, ce qui permet d'obtenir un écart symétrique : $d^*(s, v) = \frac{d(s, v) + d(v, s)}{2}$ mesurant la différence de forme entre s et v .

Sous-programmes ÉCARTS

• Écart n° 1 : on calcule la droite des moindres carrés ajustant t'_h en fonction de t_h : $y_h = at_h + b$; l'écart $d_1(s, v)$ est donné par :

$$d_1(s, v) = \left(\frac{1}{P} \sum_{h=1}^P (at_h + b - t'_h)^2 \right)^{1/2} \quad (4.10)$$

a : rapport des largeurs,
 b : séparation temporelle des signaux.

• Écart n° 2 : on estime la tangente à la courbe Φ en calculant la droite des moindres carrés sur une fenêtre glissante de largeur donnée (paramètre à régler). On calcule pour chaque position de la fenêtre l'écart quadratique moyen des ordonnées de Φ par rapport à la tangente, en dehors de la fenêtre. L'écart $d_2(s, v)$ représente le *maximum* des écarts précédemment calculés.

Pour d'autres écarts voir la référence [3].

5. Quelques caractéristiques de la méthode

Différence de forme et de largeur

Dans certaines applications (par exemple la sommation synchrone de signaux répétitifs), il est nécessaire de mesurer ou de détecter à la fois une variation de forme et une variation de largeur des signaux. En utilisant l'écart n° 1, il suffit d'imposer $a = 1$. L'écart combinerait alors un écart de forme proprement dit et un écart de largeur.

Influence du bruit sur la méthode

- Pour un rapport S/B supérieur à 20 dB, la méthode ne présente pratiquement aucun biais et permet la mesure de très petites variations de forme indécélables à l'œil : l'exemple type est la caractérisation de la forme des pics chromatographiques (cf. [2]).
- Pour un rapport S/B inférieur à 20 dB, la méthode est relativement robuste mais nécessite des précautions d'emploi (pour une étude théorique approfondie voir [5]).

Réglage des paramètres

- Fenêtres d'intégration : en principe elles sont définies en terme de pourcentage de la hauteur maximale du signal : un réglage est à faire en fonction du type de signaux et du rapport S/B. Dans ce cas, il est utile de faire la recherche des bornes d'intégration sur un signal préalablement lissé.
- Fenêtre de calcul de l'écart : il est souvent nécessaire pour éviter des effets de bord de supprimer lors du calcul de l'écart une partie des points extrêmes de la fonction Φ .