

Corrélation multiéchelle et interpolation de la trace warp dans le recalage temporel de signatures vibro-acoustiques

Multi-scale correlation and warp trace interpolation used in the correction of the time-warping of vibro-acoustic signatures

par François LÉONARD

Hydro-Québec (IREQ), 1800 boul. Lionel-Boulet, Varennes, Québec, Canada J3X 1S1

Courriel : leonard@ireq.ca

résumé et mots clés

Les difficultés soulevées lors de la comparaison de signatures vibro-acoustiques d'équipements de commutation de haute puissance nous ont amené à développer un nouvel algorithme de recalage temporel. Cette comparaison de signatures est requise pour réaliser la surveillance de ces équipements. Or, ces signatures comportent une suite de transitoires générée par une séquence d'événements électromécaniques qui apparaissent avec une trame temporelle légèrement différente, d'une commutation à l'autre, selon, entre autres, la température et la charge. Cette distorsion temporelle génère une divergence significative entre les amplitudes instantanées des signatures. Le recalage temporel permet la juxtaposition des signatures en déplaçant les éléments de l'une des signatures afin qu'elle soit conforme à la trame temporelle de l'autre signature. De plus, la grandeur de la déviation temporelle, entre la dernière signature et une référence, a une utilité diagnostique. L'algorithme de recalage temporel proposé permet de trouver la relation temporelle entre les événements de deux signatures, même en présence de discontinuités de la trame temporelle. Or, nous démontrons que ces discontinuités ne sont pas traitées adéquatement par l'algorithme de recalage temporel DTW (Dynamic Time Warping) utilisé couramment dans le traitement de la parole. Nous expliquons donc le fonctionnement de ces deux algorithmes et nous présentons des résultats démontrant la plus grande acuité de la corrélation multiéchelle. C'est en partie l'interpolation de la trace warp qui permet d'atteindre cette acuité. Cette trace warp est une fonction décrivant la déviation temporelle entre signatures, rapportant même la présence d'inversions dans l'ordre d'apparition de transitoires.

Corrélation multiéchelle, DTW, recalage temporel, distorsion temporelle, surveillance acoustique, vibro-acoustique, transformateur, changeur de prises.

abstract and key words

Problems that arise when we are comparing the vibro-acoustic signatures of high-voltage apparatus in electrical substations have led to the discovery of a new time-warp algorithm. Such signature comparisons are made in order to monitor trends in

the health of the equipment. However, the signatures collected contain a sequence of transients generated by electromechanical events which appear with a slightly different time frame from one switching operation to the other, depending on the temperature and load, among other factors. Despite being small, this time distortion results in a significant difference between the instantaneous amplitudes of the signatures. This is where time warping comes in useful. It is used for juxtaposing the signatures by displacing the elements of one of the signatures so that it matches the time frame of the other. The magnitude of the time difference between the last signature and a reference has a diagnostic use. The proposed time warp algorithm makes it possible to identify the time relationship between the events of two signatures, even in the presence of discontinuities in the time frame. We show that these discontinuities are not treated appropriately by the dynamic time warping (DTW) algorithm currently employed in speech processing. The paper explains the way these two algorithms work and presents results that bear witness to the increased accuracy of the multi-scale correlation. In part, it is the interpolation of the warp trace that makes such accuracy possible. The warp trace is a function describing the time deviation between the signatures and even including the presence of inversions in the order of appearance of the transients.

Multi-scale correlation, DTW, time-warping, acoustic monitoring, vibro-acoustic, power transformer, tap-changer.

1. introduction

1.1. Contexte de l'application et problématique à résoudre

Dans les statistiques de défauts concernant les équipements dans les postes de distribution, environ la moitié des défauts survenant sur des transformateurs sont attribuables aux changeurs de prises [1]. Rappelons qu'un changeur de prises effectue la régulation de la tension de ligne en sélectionnant le nombre requis de spires dans l'induit du transformateur. Comme le signal vibro-acoustique d'une commutation véhicule une information concernant l'état de santé de l'équipement, ce signal peut être exploité par un système de surveillance. Lorsque capté par un accéléromètre et échantillonné à un taux approprié, ce signal de vibration a un contenu spectral pertinent pouvant s'étendre à plus de 30 kHz. Nous verrons aux chapitres 3.5 et 3.6 que ce signal est traité afin de condenser l'information jugée utile dans un volume plus réduit et sous une forme plus pratique à manipuler. Nous appelons signature le résultat de cette transforma-

tion. Dans notre contexte, la signature apparaît sous la forme d'une enveloppe temporelle ou d'une distribution temps-fréquence. Cette distribution contient usuellement plus d'information que l'enveloppe, requiert plus d'espace mémoire et améliore la discrimination lorsque les événements mécaniques se manifestent par des transitoires de différentes fréquences.

Parmi les divers équipements de commutation, le changeur de prises exhibe des signatures particulièrement complexes puisqu'elles contiennent diverses transitoires, plus ou moins superposés, correspondant à des séquences électromécaniques qui se déroulent en parallèle sur les trois phases. La position relative de chaque transitoire n'est pas exactement identique d'une signature à l'autre. Cette différence peut être à peine perceptible ou être très significative. La figure 1 illustre un exemple où il apparaît un glissement uniforme de la trame temporelle entre deux signatures. Ici, c'est la viscosité de l'huile, affectée par la température, qui est à l'origine de cette variation de la trame temporelle: la signature enregistrée à 23 heures est accélérée de 3,5 % par rapport à celle collectée à 18 heures, causant l'apparition d'une crête là où l'amplitude était faible sur la signature précédente. De plus, des essais de défauts introduits sur des équipements en usine [2,3] démontrent qu'un ressort lâche affecte aussi la trame temporelle.

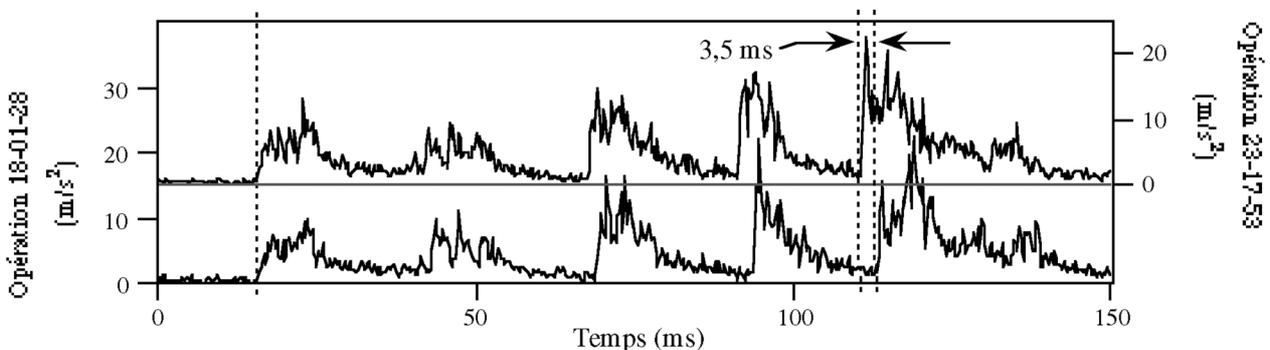


Figure 1. – Enveloppes de l'émission vibro-acoustique de deux changements de prises survenus la même journée mais à des températures différentes.

L'amplitude, le contenu fréquentiel, de même que la position temporelle des transitoires constituent une information utile pour réaliser un diagnostic. Le système de surveillance comparera alors les signatures à différentes époques afin de déceler l'apparition d'un défaut. Cependant, la comparaison des signatures est difficile lorsque les transitoires ne sont pas exactement localisés aux mêmes lieux temporels et n'ont pas la même durée. On parle alors de distorsion de la trame temporelle, et, lorsque l'on compare deux signatures, on parle de déviation temporelle entre celles-ci. Enfin, une différence significative peut mener à un changement de l'ordre d'arrivée de certains événements : la séquence temporelle d'apparition des transitoires n'est plus la même, on dit qu'il y a du repliement temporel.

1.2. Solutions possibles et fonctionnalité

Deux solutions peuvent être envisagées afin de comparer les signatures. La première solution consiste à ne comparer que les signatures enregistrées dans de mêmes conditions de fonctionnement, soit à une même charge et une même température : en absence de défaut, ces signatures ont une trame temporelle similaire. Cette solution est coûteuse en allocation de mémoire et pose différents problèmes dont celui de l'incursion dans des conditions de fonctionnement où il n'y a pas de signature de référence enregistrée longtemps après la mise en service du système de surveillance. La seconde solution consiste à réaligner les différents événements temporels ou transitoires de façon à obtenir une comparaison qui ne sera pas, ou peu, affectée par la distorsion temporelle. C'est cette seconde solution qui est abordée dans cet article et illustrée à la figure 2. Dans un premier temps, une estimation de la déviation temporelle est déduite de la comparaison entre la signature et une référence. La signature de référence est constituée usuellement à partir d'une moyenne de signatures [4]. Le résultat de l'estimation de la déviation temporelle est une fonction de transfert temporelle (fig. 3) qui donne la correspondance entre les indices temporels des signatures. Cette fonction est appliquée pour recalcer la nouvelle signature, tout en étant utilisée par le module de diagnostic pour surveiller tout dépassement suspect du gabarit temporel. Ce gabarit peut être construit à partir des tolérances fixées par le manufacturier pour la chronologie des événements mécaniques du changeur de prises, ou encore, être produit à partir de statistiques établies sur des équipements identiques qui sont exempts de défauts.

Précisons que, pour un signal de vibration, le recalage ne peut être estimé à partir du signal temporel brut, ni être réalisé sur un signal temporel brut. En effet, si l'on considère un transitoire unique, c'est cycle après cycle que s'applique le recalage, un peu comme dans une acquisition synchrone sur une machine tournante (order sampling). Les alternances des transitoires s'ajustent l'une après l'autre jusqu'à la fin de l'un des transitoires, laissant l'autre transitoire continuer seul son évolution temporelle. Ici, il

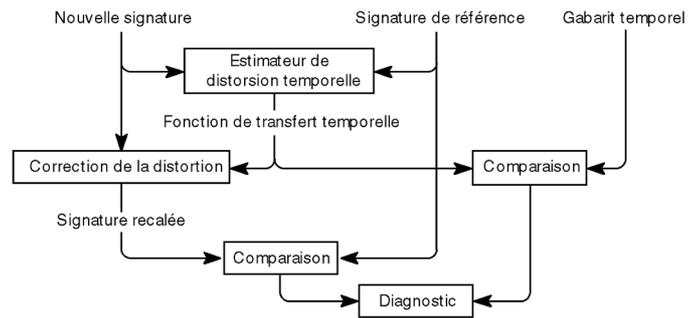


Figure 2 . – Insertion du module d'estimation de la distorsion temporelle dans un système de surveillance où la comparaison de signatures est requise.

ya un paradoxe où, d'une part, si l'on recale les transitoires sur la position des alternances, ils ne peuvent avoir un début et une fin qui correspondent, et, d'autre part, si l'on force la correspondance du début et de la fin des transitoires, l'un de ceux-ci a sa fréquence propre modifiée. Il y a aussi un autre paradoxe : soit la comparaison de deux transitoires débutant et se terminant aux mêmes moments mais ayant des fréquences légèrement différentes. Doit-on recalcer ces transitoires pour que les alternances ou les phases instantanées coïncident ? Cependant, on peut rééchantillonner, allonger artificiellement ou encore tronquer un transitoire afin de contourner ces paradoxes. Or, appliqué sur un signal comportant plusieurs transitoires, le recalage ne peut être réalisé sans dissociation des différents transitoires qui se chevauchent. En effet, il faut extraire chaque transitoire distinctement, modifier la position de chacun et finalement les additionner pour reconstituer le signal. Nous avons envisagé d'utiliser la transformée en vers [5] afin d'extraire ces transitoires (disponible sur le site ftp.ireq.ca/pub/T-F_worms). Cependant, même si cet outil a livré des résultats prometteurs pour certains types de signaux, il nous a été impossible de séparer avec confiance les nombreux transitoires présents dans le signal capté lors d'un changement de prises. Comme nous ne pouvons décortiquer ce signal en une somme de transitoires, le recalage est effectué sur une signature dont on a retiré l'information concernant la phase instantanée des transitoires, et non pas sur le signal temporel brut.

2. Algorithme DTW conventionnel et limitations

Utilisé depuis les années 70 dans le traitement de la parole, l'algorithme « Dynamic Time Warping » (DTW) [6-8] fait appel à de la programmation dynamique et est maintenant employé dans diverses applications tels le traitement d'images ultrasoniques

[9], la reconnaissance de l'écriture [10] et l'interprétation de courbes dans le diagnostic de procédés industriels [11]. Plus proche de notre application, il existe un produit commercial pour le diagnostic de disjoncteurs haute tension [12, 13] qui, comme dans le traitement de la parole, calcule la déviation temporelle mais ne corrige pas la signature afin d'aligner les transitoires deux à deux. Nous verrons que si l'algorithme DTW convient bien pour le traitement de la parole, par contre, il n'est pas adéquat pour la comparaison de signatures vibro-acoustiques. De plus, dans le traitement de la parole, les efforts de recherche [14-16] concernent plus la réduction de l'effort de calcul que l'augmentation de la précision. Or, la surveillance d'équipement de commutation électrique ne requiert pas une analyse en temps réel puisque les commutations sont habituellement distancées de plusieurs minutes et le non-enregistrement de l'une d'elles de temps à autre est sans importance : la contrainte du temps de calcul est moins présente alors que la précision de l'alignement temporel prédomine.

Soit deux signatures temps-fréquence discrètes $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \dots, \mathbf{a}_N\}$, et $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m, \dots, \mathbf{b}_M\}$, où \mathbf{a}_n et \mathbf{b}_m sont des vecteurs contenant chacun une tranche spectrale de la signature correspondante avec n et m comme indices de localisation temporelle. On désire recalcr \mathbf{B} sur \mathbf{A} . Les nombres N et M sont de grandeurs voisines. La solution cherchée est une fonction de recalage temporel, appelée aussi « dynamic warp function »,

$$m = w(n) \quad (1)$$

qui liera ces deux signatures de façon à minimiser la distance

$$D(\mathbf{A}, \mathbf{B}) = \frac{1}{N} \sum_{n=1}^N d(\mathbf{a}_n, \mathbf{b}_{w(n)}) \quad (2)$$

entre celles-ci où $d(\mathbf{a}_n, \mathbf{b}_m)$ correspond à une distance locale. Cette distance peut être définie de différentes façons, dont certaines très spécialisées comme la distance d'Itakura [7] en reconnaissance de la parole. Dans ce dernier domaine, c'est la distance minimale donnée à l'équation 2 entre un mot enregistré et différents mots de référence qui détermine le mot prononcé. Or, dans l'état de nos connaissances en surveillance vibro-acoustique de changeur de prises, nous ne pouvons reconnaître un transitoire et le rattacher à symptôme utile au diagnostique. C'est plutôt la trajectoire de la fonction de recalage temporel $w(n)$ qui est requise pour la correction de la déviation temporelle et non la reconnaissance d'un transitoire particulier. Dans notre implantation de l'algorithme, nous avons donc utilisé une formulation à la fois simple et plus générale,

$$d(\mathbf{a}_n, \mathbf{b}_m) = \|\mathbf{a}_n - \mathbf{b}_m\|. \quad (3)$$

correspondant à la norme de la soustraction vectorielle entre une tranche temporelle provenant de la signature \mathbf{A} et une tranche provenant de la signature \mathbf{B} .

Les conditions de frontière sont

$$w(1) = 1 \quad (4)$$

pour le point de départ et

$$w(N) = M \quad (5)$$

pour le point d'arrivée. Il est aisé de trouver la fonction de recalage temporel lorsqu'il y a une déviation linéaire de l'échelle temporelle entre les deux signatures. Cette fonction prend alors la forme

$$w(n) = \left(\left(\frac{M-1}{N-1} \right) (n-1) \right) + 1 \quad (6)$$

avec les conditions frontières exprimées aux équations 4 et 5. De façon plus générale, la différence de trame temporelle doit être corrigée par une fonction de recalage discrète fonction de l'indice temporel n . Celle-ci est obtenue localement en partant de l'origine $n = m = 1$ et en progressant tranche temporelle après tranche temporelle. Communément utilisée en traitement de la parole [17], c'est la règle

$$w(n+1) - w(n) = 0, 1, 2 \text{ si } w(n) \neq w(n-1) \quad (7.a)$$

$$w(n+1) - w(n) = 1, 2 \text{ si } w(n) = w(n-1) \quad (7.b)$$

de concert avec la minimisation de la distance définie à l'équation 2, qui définit le pas de progression $\{0, 1 \text{ ou } 2\}$ de la signature \mathbf{B} . L'équation 7b interdit que l'on retarde d'un échantillon, soit $w(n+1) - w(n) = 0$, deux fois de suite la signature \mathbf{B} par rapport à la signature \mathbf{A} . Lorsque $w(n+1) - w(n) = 2$, on devance d'un échantillon la signature \mathbf{B} par rapport à \mathbf{A} , et, enfin, les deux signatures progressent à un même rythme lorsque $w(n+1) - w(n) = 1$. La figure 3 illustre un exemple didactique d'une trace de recalage. Les trois premiers pas sont illustrés avec des flèches sur cette figure. Le premier pas correspond à un incrément nul de la fonction de recalage, le second pas correspond à un incrément unitaire et le troisième illustre un incrément de 2. La zone grise est celle contenant l'ensemble des trajets possibles. Cette zone grise est bornée par une paire de pentes maximales et une paire de pentes minimales qui, en se croisant, délimitent un losange tel qu'illustré sur la figure 3. Notons que le nombre de trajets possibles est optimal lorsque $M \approx N$ alors qu'il est nul lorsque $M > 2N$ ou pour $2M < N$. Les signatures soumises à l'algorithme doivent donc être alignées grossièrement et tronquées de sorte que les nombres N et M soient de grandeurs voisines. Allonger ou tronquer indûment une signature peut mener à l'élimination du parcours le plus vraisemblable.

Mentionnons qu'à chaque pas sur la signature \mathbf{A} , on évalue l'ensemble des parcours possibles pour ne conserver que le plus court pour chaque point de la signature \mathbf{B} pouvant être rattaché

à ce pas sur **A**. C'est à l'approche de l'autre extrémité, au point d'arrivé (N, M), que tous ces parcours possibles s'éliminent progressivement pour laisser place au parcours minimal qui sera retenu. L'évaluation de tous les parcours possibles est exigeante en effort de calcul. Tels que proposés par Stainhaouer et Carayannis [16], nous avons inséré dans notre algorithme DTW un tampon mémoire ne retenant que les trajets les plus probables (fig. 3). De plus, ce tampon ne se déplace pas de façon rigide le long de la diagonale reliant le point de départ au point d'arrivée : le tampon est recentré à chaque étape sur le nouveau trajet le plus probable. Ce centrage se déplace que d'un élément du tampon à chaque pas pour limiter les instabilités liées à la présence simultanée de deux trajets à la fois très probables et distants. Au départ, la dimension du tampon croît avec la distance, entre la pente maximale et la pente minimale, jusqu'à atteindre la dimension allouée pour celui-ci. Enfin, nous avons introduit dans notre implantation l'option de contourner la règle énoncée à l'équation 7 afin de permettre $w(n+1) - w(n) = 0$ plusieurs fois de suite (voir « DTW règle éq.7 modifiée » aux légendes des figures 11 et 12). Avec cette option en fonction, l'algorithme DTW répond plus rapidement à certaines discontinuités temporelles.

La règle définie aux équations 4, 5 et 7 fonctionne qu'en présence d'une distorsion temporelle ayant une restriction importante sur la pente de progression et la solution de conditions frontières aux extrémités. On peut songer à modifier cette règle. Par exemple, il existe une variante à cet algorithme où est assouplie la contrainte sur les conditions frontières données aux équations 4 et 5 [18]. D'ailleurs, dans notre implantation de la DTW, la position du tampon étant asservie à la trajectoire en dévelop-

pement, le point réel d'arrivé ne donne pas nécessairement au voisinage du point d'arrivée tel que défini à l'équation 5. On peut aussi modifier la règle définie à l'équation 7 afin d'augmenter la pente maximale ainsi que réduire la pente minimale. Permettre localement des sauts de plusieurs incréments de l'indice m soulève un autre problème : lorsqu'un objet corrélera trop bien avec un autre objet d'une durée beaucoup plus courte et présent dans la seconde signature, la distance $d(a_n, b_m)$ donnée à l'équation 3 reste petite si l'on reste accroché à cet objet, ce qui force un blocage de la progression, soit $w(n+1) = w(n)$, jusqu'à ce que le retard équivaille au saut maximum permis. La règle exprimée en 7.b évite ce blocage en poussant en avant la progression de part et d'autre sur les deux signaux. On pourrait aussi pondérer l'accroissement de la longueur du parcours en fonction de la grandeur du saut effectué afin de pénaliser un choix de parcours présentant un tel blocage. De plus, examiner la possibilité de sauts sur plusieurs incréments mène à une croissance géométrique de l'effort de calcul en fonction la grandeur du saut accordé. En effet, en absence de règle, à chaque nouveau pas le nombre de trajets possibles se trouve multiplié de telle sorte que la croissance de l'effort de calcul s'apparente à une progression factorielle !

L'algorithme DTW est acceptable pour le traitement de la parole où il n'y pas de pause au milieu d'un mot et pas de changement dans l'ordre temporel d'arrivée des phonèmes. De plus, on passe progressivement d'un phonème à l'autre sans possibilité de superposition (*i.e.* le son « a » et « u » en même temps). Dans le contexte des changeurs de prises, comme celui des disjoncteurs, nous sommes en présence d'une suite de transitoires où il n'y a pas de continuité entre la fin d'un transitoire et le début d'un nouveau transitoire : on passe subitement d'une séquence spectrale générée par un événement à une autre séquence qui peut venir se superposer ou remplacer la première. L'algorithme DTW n'a pas été conçu pour suivre ce saut subit. Nous verrons qu'il y a alors un glissement, ou délai, le temps que l'algorithme rattrape le saut subit, et ce, à un rythme correspondant à la pente minimale ou maximale. Pire, lorsque l'algorithme est optimisé afin d'examiner un nombre réduit de trajectoires, un saut important dans la distorsion temporelle demande à l'algorithme de prendre un chemin qu'il n'a pas conservé dans son tampon mémoire. Par exemple, dans notre implantation, si le tampon est trop étroit face à une distorsion temporelle subite, l'algorithme décroche et se met à suivre une fausse piste.

En résumé, les limitations de l'algorithme DTW proviennent d'une vision limitée à un sous-ensemble restreint des trajectoires possibles. Nous détaillerons plus loin les conséquences de ces limitations à la lumière des résultats livrés par l'algorithme basé sur la corrélation multiéchelle.

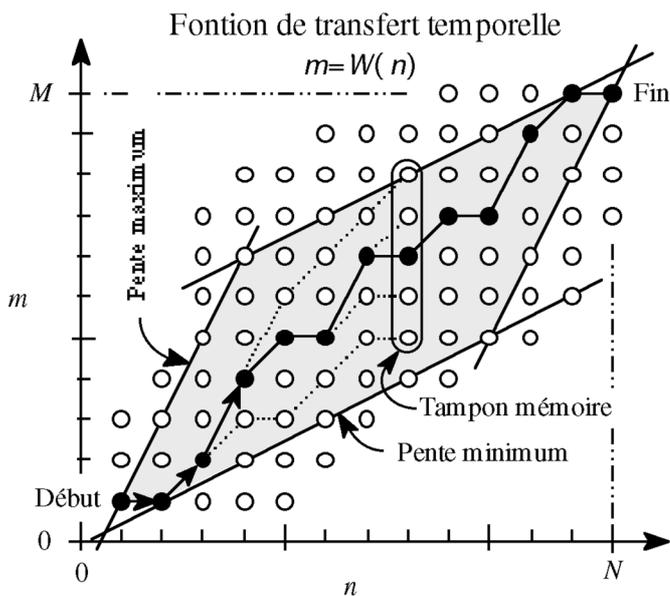


Figure 3. – Progression d'une trace de recalage temporel avec illustration du tampon de mémoire contenant les trajets les plus probables.

3. Recalage temporel par corrélation multiéchelle dyadique

3.1. Mécanisme de fonctionnement menant à la création de la matrice warp

Lorsque l'on cherche à recalcr deux signatures manuellement, on aligne d'abord grossièrement l'ensemble de chaque signature, celles-ci étant prises d'un bloc. Par la suite, on scinde chaque signature en une suite d'objets et on aligne ces objets entre eux. Comme on peut scinder certains de ces objets en sous-objets, on aligne plus finement chaque sous-objet des objets concernés. On prendra aussi soin d'écarter, du processus de comparaison, les objets n'apparaissant pas sur les deux signatures. L'algorithme mis au point s'inspire de cette image mentale que l'on se fait de cette tâche si elle était confiée à un humain, paire de ciseaux en main. La première étape consiste précisément à ajuster les deux signatures globalement par une cross-corrélation où la position du maximum détermine le décalage temporel global. Plutôt que retenir cette valeur de décalage temporel global, nous retenons le résultat de la cross-corrélation entre les signatures et dupliquons celui-ci pour remplir une matrice telle qu'illustrée à la figure 4. Le symbole \otimes représente ici la cross-corrélation. Par la suite, on coupe en deux l'une des signatures et l'on effectue deux cross-corrélations pour remplir les deux portions de matrices correspondantes. Le résultat de chaque cross-corrélation est additionné au résultat de la première cross-corrélation. On recoupe en deux chaque demi-signature et on effectue quatre cross-corrélations, on ajoute les résultats aux résultats précédents et ainsi de suite. La vision de la signature devient multi-échelle : on corrèle l'ensemble de la signature, les objets la composant, les sous-objets composant chaque objet...

Toute l'astuce réside dans la façon d'aligner les résultats de cross-corrélation pour effectuer la sommation. Chaque résultat est disposé de sorte que le maximum de corrélation survient au milieu de la matrice, selon une ligne verticale ici, si nous avons deux signaux identiques à comparer. Les espaces non comblés sont remplis avec des zéros. Le résultat de la sommation est donc une matrice comportant des éléments créés par la somme des cross-corrélations effectuées sur différentes échelles. Cette matrice sera appelée la matrice warp. Notons que cette matrice a une dimension de $N(2N - 1)$ pour deux signatures de N éléments. Bien que la corrélation multi-échelle fonctionne avec des longueurs de signature qui diffèrent, dans ce qui suit nous prendrons un nombre identique d'éléments pour les deux signatures afin de simplifier les expressions. De cette matrice, nous extrayons la trace warp qui correspond, à une pente près, à la fonction de recalage $w(n)$. La trace warp nous renseigne sur les

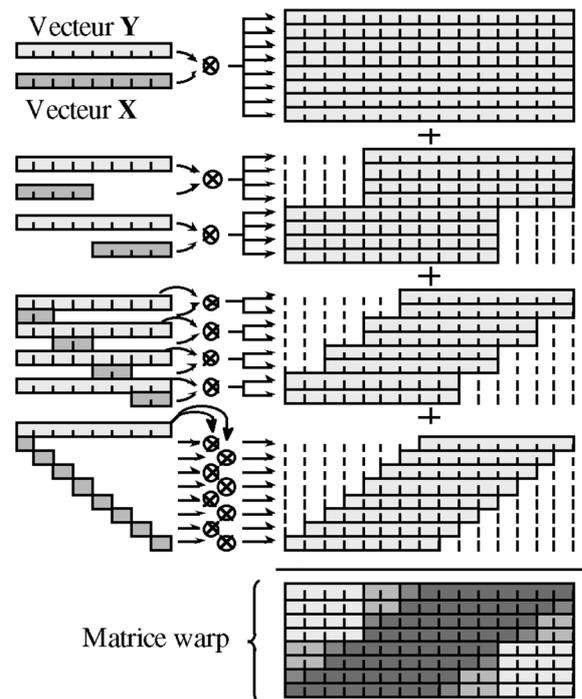


Figure 4. – Illustration de la corrélation multi-échelle dyadique CMED où l'on détaille la méthode de sommation des résultats de cross-corrélations menant à la création de la matrice warp.

écarts temporels distinguant les signatures et permet d'ajuster temporellement celles-ci pour fin de comparaison.

La corrélation est multi-échelle car une même portion de signature est soumise à des corrélations successives sur des échelles de plus en plus réduites. Nous coupons un des vecteurs à des endroits systématiques et répartis pour optimiser le temps de calcul, en espérant que le nombre de coupes et que la résolution obtenue avec le plus petit pas de coupe contrebalanceront le choix d'une coupe non fondée sur la forme de la signature. Comme la dimension de la portion de signature à corréliser est scindée en deux entre chaque échelle, nous ajoutons le terme « dyadique » à la corrélation multi-échelle pour finalement la dénommer « corrélation multi-échelle dyadique » ou CMED.

Le choix d'une progression dyadique fait en sorte que la dimension des colonnes de la matrice warp devient une puissance de deux du nombre d'échelles. La durée de la signature divisée par la dimension des colonnes détermine la résolution temporelle du processus et équivaut au délai séparant deux tranches temporelles contiguës. Dans le contexte de la surveillance de changeurs de prises, nous effectuons des corrélations multi-échelles sur les signatures temps-fréquence avec de 64 à 256 tranches temporelles. Nous obtenons alors des dimensions matricielles allant de 64×127 à 256×511 : l'espace mémoire requis pour manipuler la matrice warp n'est donc pas négligeable.

Un élément x_n du vecteur $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$ affiché à la figure 4 peut être un échantillon temporel de l'enveloppe ou un vecteur contenant une portion d'enveloppe de signature. Si le

signal brut est scindé en un signal multibande, un élément \mathbf{x}_n peut être aussi une sous-matrice correspondant à un ou des échantillons temporels prélevés sur les enveloppes de différentes bandes de fréquences. Enfin, un élément \mathbf{x}_n peut être un vecteur correspondant à une tranche temporelle de la distribution temps-fréquence de la signature, identique au vecteur \mathbf{a}_n de la signature soumise à l'algorithme DTW. Un spectrogramme ou une transformée en ondelettes peuvent être employés pour générer cette distribution temps-fréquence de la signature. Notons que les sous-éléments de \mathbf{x}_n appartiennent à l'ensemble des nombres réels positifs. Enfin, l'algorithme est conçu pour des vecteurs \mathbf{X} et \mathbf{Y} de mêmes dimensions.

3.2. Formulation axiomatique de la corrélation multiéchelle dyadique (CMED)

L'algorithme CMED commande la définition d'un outil de manipulation matricielle peu courant, soit le produit scalaire

$$\mathbf{x}_n \otimes \mathbf{y}_m = \sum_i \sum_j x_{n,i,j} \cdot y_{m,i,j} \quad (8)$$

entre deux matrices de sorte que

$$\|\mathbf{x}_n\|^2 = \mathbf{x}_n \otimes \mathbf{x}_n \quad (9)$$

$$\text{corr}_{n,m}\{\mathbf{X}, \mathbf{Y}\} = \frac{\sum_{k=0}^{\log_2\{N/I\}} \left(p^k \text{corr}_m\{\mathbf{X}, \mathbf{Y}\} \Big|_{\mathbf{x} \begin{cases} \mathbf{x}_i | i \in \text{int}(I(n-1) \cdot 2^k/N) \cdot N/2^k, \text{int}(1 + I(n-1) \cdot 2^k/N) \cdot N/2^k \\ \mathbf{x}_i = 0 \text{ ailleurs} \end{cases}} \right)}{\sum_{k=0}^{\log_2\{N/I\}} P^k} \quad (11)$$

pour $n = 1, 2 \dots \text{int}(N/I)$ et $m = 1, 2 \dots 2N - 1$, où P est un paramètre de pondération d'échelle d'une valeur proche de l'unité. Précisons que l'indice n représente le temps alors que la déviation entre les signatures est donnée selon m . Le résultat final est donc une matrice de dimension $(2N - 1) \times \text{int}(N/I)$. Le paramètre de pondération P a pour fonction d'amplifier ou de diminuer le poids de la corrélation avec le changement d'échelle de l'intervalle corrélé. La variable I limite le nombre de coupes de sorte que la plus petite longueur de coupe donne une suite de I éléments \mathbf{y}_n . Nous avons $I = 1$ pour des signatures temps-fréquence.

Nous constatons que la figure 4 est davantage à la portée de notre compréhension que l'équation 11 : l'idée de la corrélation multiéchelle provient d'ailleurs d'une image mentale et non d'une équation abstraite. Si cette équation exprime précisément le mécanisme de la CMED, elle est par contre difficile à mani-

donner le carré de la norme de la matrice \mathbf{x}_n . On utilise la fonction de corrélation normée

$$\text{corr}_m\{\mathbf{X}, \mathbf{Y}\} = \frac{\sum_{n=1}^N \mathbf{x}_n \otimes \mathbf{y}_{n+m}}{\sqrt{\sum_{n=1}^N \|\mathbf{x}_n\|^2 \Psi_{\mathbf{y}}\{n+m\} \cdot \sum_{n=1}^N \|\mathbf{y}_{n+m}\|^2 \Psi_{\mathbf{x}}\{n\}}} \quad (10)$$

de façon à ne pas biaiser la position maximale de corrélation avec $\Psi_{\mathbf{x}}\{i\}$ comme fonction d'existence où $\Psi_{\mathbf{x}}\{i\}|_{\mathbf{x}_i \text{ existe}} = 1$ et $\Psi_{\mathbf{x}}\{i\}|_{\mathbf{x}_i \text{ n'existe pas}} = 0$. En effet, sans cette normalisation, l'élément \mathbf{y}_{m+j} de plus grande amplitude aura toujours une prédominance et attirera à lui le sommet de corrélation. On remarquera aussi que la corrélation est « non centrée », c'est-à-dire qu'on ne lui a pas retiré la moyenne avant calcul. Dans notre contexte, la valeur moyenne de la portion de signature doit être aussi corrélée puisque celle-ci correspond à la puissance instantanée contenue dans cette portion et que cette puissance contient une information pertinente au recalage.

La corrélation normée définie à l'équation précédente permet d'énoncer plus succinctement la CMED. Ainsi, cette dernière s'énonce

puler mathématiquement. Il y aurait sans doute un gain à trouver une expression plus aisée à manipuler tout en présentant un résultat similaire. Nous pourrions ainsi déduire analytiquement les propriétés mathématiques plutôt que de procéder par simulations numériques. De même, des développements mathématiques permettraient de déboucher sur d'autres applications.

3.3. Extraction, interpolation et propriétés de la trace warp de recalage temporel

On peut visualiser la matrice warp ou sa transposée par différents moyens, soit une présentation pseudo-3D avec des bâtonnets, avec des surfaces en relief, ou encore avec une présentation 2D avec variation sur l'intensité lumineuse (figure 5). Pour des

signatures similaires, on constate la présence d'un sommet de corrélation au voisinage d'une ligne passant par le centre de la matrice. Nous nommerons « trace warp » le tracé parcourant ce sommet d'une extrémité à l'autre de la matrice warp. La trace warp est habituellement discernable à l'œil et sa position $T(n)$ expose la distorsion temporelle distançant les deux signatures. Une façon simple d'extraire la trace warp est de ranger dans $T(n)$ les indices correspondant aux valeurs maximales de corrélation, et décalés de N de sorte qu'une position nulle correspond à une trame temporelle identique chez les deux signatures. Cependant, le saut d'une valeur d'indice matriciel correspond à un saut de plusieurs échantillons temporels sur la signature d'origine. Par exemple, pour un spectrogramme, on a habituellement plusieurs dizaines d'échantillons temporels entre chaque transformée rapide de Fourier (TRF), chaque TRF correspondant à un incrément de l'indice n . C'est pourquoi nous utilisons une méthode d'interpolation de la trace warp permettant de préciser davantage la position du sommet de corrélation. Cette interpolation accroît substantiellement la résolution temporelle

de l'extraction de la trace warp. Pratiquement, nous fixons la résolution d'interpolation de façon à ce qu'elle corresponde au pas d'échantillonnage du signal brut.

L'algorithme d'interpolation retenu est de type 1D et est appliqué sur chaque colonne de la matrice warp transposée. L'interpolation est réalisée par la convolution d'un filtre passe-bas de type FIR [19] constitué d'un « sinc » multiplié par une fenêtre spectrale de type Blackman-Harris [20]. Rappelons que le « sinc », ou $\sin(t)/t$, correspond à la fonction de reconstruction de Shannon et est l'interpolateur exact d'un signal à bande limitée [21]. Le produit par une fenêtre spectrale permet de limiter la durée du sinc tout en minimisant l'oscillation de gain dans la bande spectrale. L'erreur d'interpolation devient alors fonction des propriétés de la fenêtre appliquée. Cependant, la principale cause d'inexactitude dans l'interpolation provient la présence de discontinuités dans la matrice warp : le signal à interpoler n'est pas une fonction à bande limitée. Par exemple, le début et la fin des valeurs non nulles de chaque colonne génèrent des oscillations locales à ces frontières qui peuvent

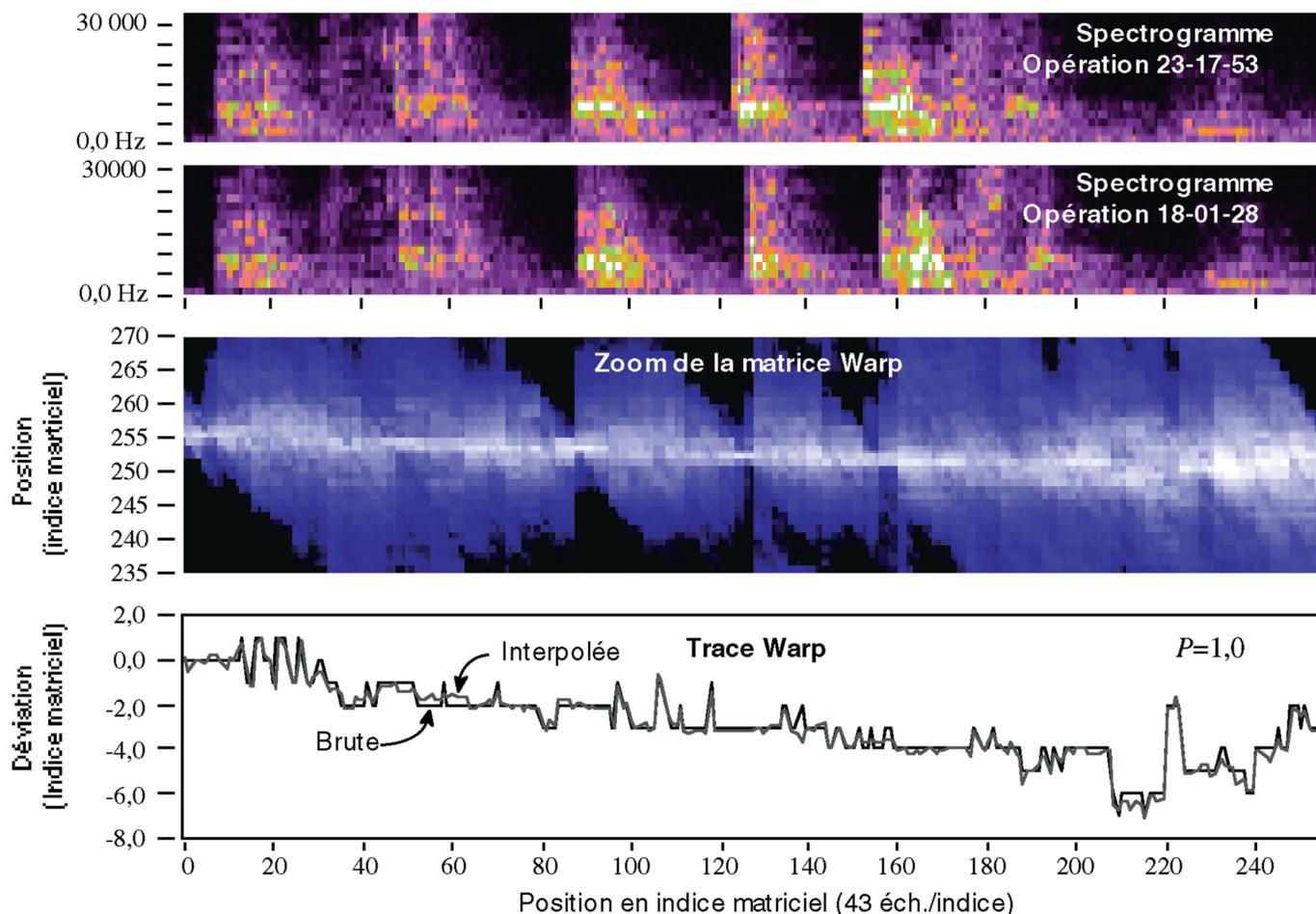


Figure 5. – Spectrogrammes extraits sur 16 bandes avec fenêtre de Blackman-Harris, zoom de la matrice warp (256 × 511) et trace warp transposées, réalisées à partir des mêmes signatures qui ont leurs enveloppes illustrées à la figure 1.

dépasser le maximum de corrélation plausible, soit l'unité. Nous avons donc introduit une fonction d'échelle telle que

$$\text{corr}^*_{n,j}\{x, y\} = \frac{1}{1,0000001 - \text{corr}^2_{n,j}\{x, j\}} \quad (12)$$

devient la fonction à être interpolée. Cette fonction s'accroît lorsque la corrélation est proche de l'unité. De plus, si nous représentons $\text{corr}^*_{n,j}\{x, y\}$ par le cosinus $\cos(\theta)$ entre deux vecteurs soumis à un produit scalaire, nous pouvons alors écrire

$$\text{corr}^*_{n,j}\{x, y\} = \frac{1}{1,0000001 - \cos(\theta)^2} \approx \frac{1}{\sin(\theta)^2} \quad (13)$$

où la valeur $\sin(\theta)$ donne la proportion non colinéaire des vecteurs. Une variante de cette fonction consiste à soustraire la corrélation $\text{corr}^*_{n,j}\{x, y\}$ plutôt que son carré. Évidemment, l'interpolation d'une fonction arbitraire ne donne pas nécessairement la bonne valeur de la position du maximum de corrélation. Néanmoins, notre expérience démontre que cette interpolation donne une valeur satisfaisante, plus vraisemblable que l'interpolation de $\text{corr}^*_{n,j}\{x, y\}$. Éventuellement, il faudra trouver la fonction d'interpolation, 1D ou 2D, et la fonction d'échelle optimales.

Nous définirons la trace autowarp

$$T_0(n) = \text{indice}(\max_l\{\text{coorm}_{n,l}\{\mathbf{Y}, \mathbf{Y}\}\}) \quad (14)$$

comme la trace issue de la corrélation multiéchelle d'une signature par rapport à elle-même. La trace autowarp est située dans la colonne du centre de la matrice warp et doit avoir une position indépendante de la signature \mathbf{Y} et constante pour tout n . Usuellement, nous avons $T_0(n) = N$ en absence d'interpolation de la trace warp. Cependant, la position de la trace autowarp peut différer très légèrement, de moins d'un élément d'indice de colonne, selon le type d'interpolation numérique réalisée. Par exemple, pour certains algorithmes d'interpolation 1D ou 2D, la trace warp présente des déviations qui peuvent être amplifiées aux abords des frontières de la matrice warp. Ces artefacts de manipulations numériques déplacent la trace warp. Pour éliminer ce déplacement et reporter la valeur de position de la trace warp de N , on définit la valeur de la trace warp

$$T(n) = \text{indice}(\max_l\{\text{coorm}_{n,l}\{\mathbf{X}, \mathbf{Y}\}\}) - T_0(n) \quad (15)$$

comme la valeur d'indice de la position du maximum de corrélation soustrait de la position de la trace autowarp. Cette dernière fonction donne directement la grandeur de déviation temporelle selon l'indice n entre les signatures.

Les signatures \mathbf{X} et \mathbf{Y} sont reliées par la trace warp telle que

$$\mathbf{x}_{n+T(n)} \approx \mathbf{y}_n \quad (16)$$

exprime la relation de recalage temporel. Comme la fonction de recalage de l'algorithme DTW relie d'une façon similaire les

signatures, soit

$$\mathbf{b}_{w(n)} \approx \mathbf{a}_n. \quad (17)$$

la trace warp $T(n)$ et la fonction de recalage $w(n)$ sont deux fonctions réalisant la même transformation et sont reliées par la relation

$$T(n) \approx w(n) - n. \quad (18)$$

Cependant, contrairement à la fonction de recalage $w(n)$ qui ne peut évoluer que de ± 1 élément par pas, la trace warp peut comporter de fortes discontinuités s'étalant sur plusieurs dizaines d'éléments matriciels.

Pour l'implantation réalisée, on constate que l'interpolation de la trace warp présente une fraction importante du temps d'exécution. En effet, si l'on interpole l'ensemble de la matrice warp, l'accroissement du temps de calcul est proportionnel au produit $(2N - 1) \times \text{int}(N/I)$, soit une croissance en N^2 . Une interpolation limitée au voisinage de la position de la trace warp brute amènerait donc une réduction importante de l'effort de calcul avec une croissance en N de cet effort.

Un exemple d'extraction de la trace warp avec interpolation est illustré à la figure 5 où un délai de 0,65 ms (43 échantillons) est présent entre deux éléments contigus de la matrice. On observe un écart de 4 éléments matriciels entre le premier et le cinquième transitoire, soit 2,6 ms, ce qui correspond approximativement au retard observé entre les enveloppes illustrées à la figure 1.

Précisions que la trace warp et sa réciproque doivent être recalées l'une par rapport à l'autre afin de les afficher sur la même référence d'écoulement temporel. En effet, prenons la fonction de transfert de la référence d'écoulement temporel

$$m \approx n + T_{\mathbf{X},\mathbf{Y}}(n) \quad (19)$$

et sa réciproque

$$n \approx m + T_{\mathbf{Y},\mathbf{X}}(m). \quad (20)$$

De celles-ci, nous déduisons la relation

$$T_{\mathbf{X},\mathbf{Y}}(n) \approx -T_{\mathbf{Y},\mathbf{X}}(m). \quad (21)$$

Premièrement, soulignons l'absence d'égalité dans ces expressions puisque l'extraction d'une trace n'est pas un processus exact mais une estimation. Deuxièmement, les références d'écoulement temporel ne sont pas les mêmes de part et d'autre dans l'équation 21. Référencé à un même indice d'écoulement temporel, nous obtenons alors

$$T_{\mathbf{X},\mathbf{Y}}(n) \approx -T_{\mathbf{Y},\mathbf{X}}(n + T_{\mathbf{X},\mathbf{Y}}(n)) \quad (22)$$

où $n + T_{\mathbf{X},\mathbf{Y}}(n)$ correspond au terme de recalage temporel requis pour comparer la trace et sa réciproque. Bien entendu, la

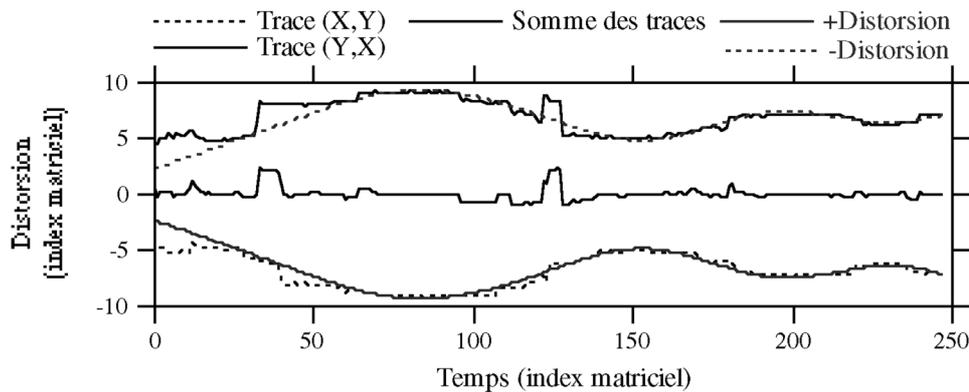


Figure 6. – Écart sur la réciprocity de la trace warp.

réciroque $T_{Y,X}$ de la trace warp étant une fonction discrète de l'indice m , il faut interpoler celle-ci à la position $n + T_{X,Y}(n)$. Enfin, si l'on pose l'hypothèse que les erreurs d'estimation de la trace warp et de sa réciroque ne sont pas corrélées entre elles, on peut combiner ces estimations pour réduire ces erreurs. Il suffit de sommer les termes de droite et de gauche de la précédente équation et de diviser cette somme par deux. La précision est alors accrue au prix d'une augmentation de l'effort de calcul de plus du double.

La figure 6 présente une trace warp reliant la signature X à la signature Y et la trace warp réciroque. Dans cet exemple, la signature Y est issue d'une interpolation temporelle de la signature X , à laquelle nous avons ajouté une distorsion connue sur la position temporelle d'interpolation. Cette distorsion est aussi affichée sur cette figure. Précisons que cette distorsion est induite par un rééchantillonnage du signal temporel avec un pas d'échantillonnage modulé par la fonction de distorsion temporelle. Ce rééchantillonnage modifiant aussi les fréquences propres des transitoires, cette distorsion temporelle dénature partiellement les transitoires présentées sur un spectre de 16 bandes. On constate que la trace warp diffère d'une part de la distorsion insérée, et, d'autre part, il n'y a pas une récirocity parfaite entre les traces puisque la somme des traces ne donne pas zéro. Les écarts observés proviennent des hachures verticales présentes sur la transposée de la matrice warp, telles que celles observées à la figure 5. Ces hachures s'expliquent par la nature transitoire du signal et par les frontières abruptes du découpage réalisé dans la construction de la matrice warp.

3.4. Correction de la déviation temporelle d'une signature

Rappelons que la correction de la déviation s'exerce sur une signature et non sur le signal temporel d'origine, telle qu'expliquée au chapitre 1.2. Précisons aussi que l'on ne corrige pas la distorsion temporelle absolue d'une signature mais bien sa

déviation avec une référence. Dans une application de surveillance, cette référence devra donc être représentative du phénomène moyen à baliser. Pour le recalage d'une enveloppe, on se contente d'interpoler celle-ci aux positions $I \cdot (n + T(n))$ où I est le nombre d'échantillons par éléments de la matrice warp. Pour le recalage d'un spectrogramme, on recalcule celui-ci en glissant une fenêtre spectrale sur les données brutes avec un pas de $d \cdot (n + T(n))$ au lieu d'un pas constant d . D'autre part, dans ces deux dernières expressions, il suffit de remplacer $n + T(n)$ par $w(n)$ pour réaliser un recalage selon l'algorithme DTW.

3.5. Application de la CMED sur des enveloppes de signatures vibro-acoustiques

Un calcul d'enveloppe rejette une partie importante de l'information présente sur le signal brut. Par exemple, du signal vibro-acoustique de changement de prises échantillonné à 65 kéch./s, on extrait une enveloppe lissée et décimée ayant un taux d'échantillonnage de 2k à 5 kéch./s. Le support d'information passant de 65 kéch./s à quelques milliers d'échantillons par seconde, il est raisonnable de suggérer qu'il y a une perte d'information du même ordre de grandeur. Il faut donc s'attendre à ce que la CMED réalisée sur une enveloppe livre un moins bon résultat que pour une signature temps-fréquence. En effet, lorsque les transitoires présentes proviennent de modes de vibration de différentes fréquences, la fréquence permet de distinguer les transitoires alors que les profils d'enveloppes de vibration modale sont similaires.

Le traitement présenté à la figure 5 pour des signatures temps-fréquence est repris à la figure 7 pour des enveloppes. Le taux de décimation étant de 16, les 12 000 échantillons de la signature d'origine donnent une enveloppe de 750 échantillons. Pour cette analyse, nous avons fixé $I = 2,9$ de façon à limiter le nombre de coupes à 256, d'où, sur l'échelle horizontale, on retrouve une échelle en indice matriciel allant de 1 à 256. Le

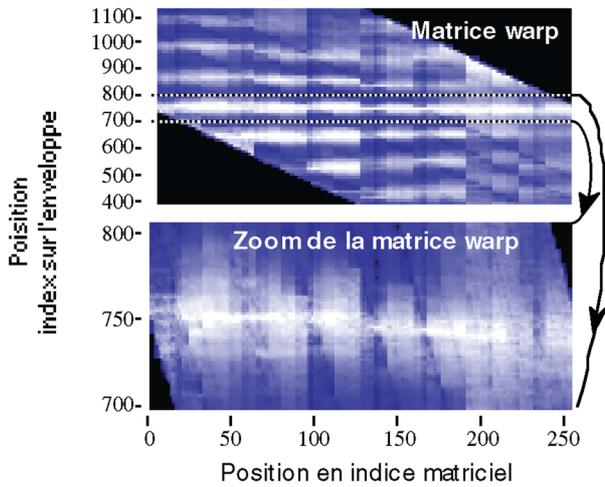


Figure 7. – Matrice warp transposée réalisée à partir des enveloppes de signatures présentées à la figure 1.

résultat de la corrélation multiéchelle sur l’enveloppe de signal diffère légèrement de celle effectuée à partir de spectrogrammes. La comparaison d’enveloppe apparaît moins précise. En effet, avec des enveloppes, la corrélation entre des portions de deux transitoires, pour une même signature, est trop proche de la valeur de corrélation obtenue pour une portion du même transitoire observée sur deux signatures. On remarque aussi la présence de corrélations croisées, entre les transitoires, générant des bandes horizontales plus claires. Ce sont surtout les formes globales des transitoires et l’agencement de ceux-ci qui font émerger la trace warp issue d’enveloppes. On arrive à visualiser cette trace warp malgré la quantité réduite d’information véhiculée par l’enveloppe.

3.6. Application de la CMED sur des signatures temps-fréquences et effet du paramètre de pondération d’échelle P

Comparativement à un calcul d’enveloppe, habituellement le calcul d’une signature temps-fréquence préserve une plus grande part de l’information présente sur le signal brut. Dans les exemples illustrés ici, le signal échantillonné à 65 kéch./s est transposé sur 16 bandes de fréquence échantillonnées à 1500 tranches spectrales/s, donnant ainsi un facteur de réduction de 2,7 dans la dimension du support de l’information. Soulignons aussi que l’on compare toujours des signatures ayant des paramètres de traitement spectral identiques : le type de fenêtre spectrale, la largeur de la fenêtre, le nombre de bandes de fréquence ou de lignes spectrales doivent être identiques. Par contre, la durée temporelle peut varier entre les deux signatures. Dans les corrélations réalisées, c’est le logarithme de la puissance spectrale qui donne le meilleur résultat en regard de l’amplitude ou de la puissance spectrale. Les matrices warp présentées dans cet article sont donc issues de signatures constituées du logarithme de la puissance.

La normalisation de la matrice warp par le dénominateur de l’équation 11 est telle que la valeur est unitaire lorsque la corrélation est parfaite. Les signatures temps-fréquences de commutation de changeur de prises affichent une corrélation sous la trace warp habituellement au-dessus de 0,8, pour des signatures présentant le même phénomène. Nous observons une valeur de corrélation plus élevée, soit plus de 0,95, pour une comparaison de logarithmes de signatures temps-fréquence. Ces valeurs élevées de corrélation apparaissent en blanc sur les illustrations de la matrice warp telle celle illustrée à la figure 5.

Sur la matrice warp, le facteur de pondération d’échelle a pour effet d’amplifier la contribution des détails de petite échelle ($P > 1,0$) ou la contribution des détails de grande échelle

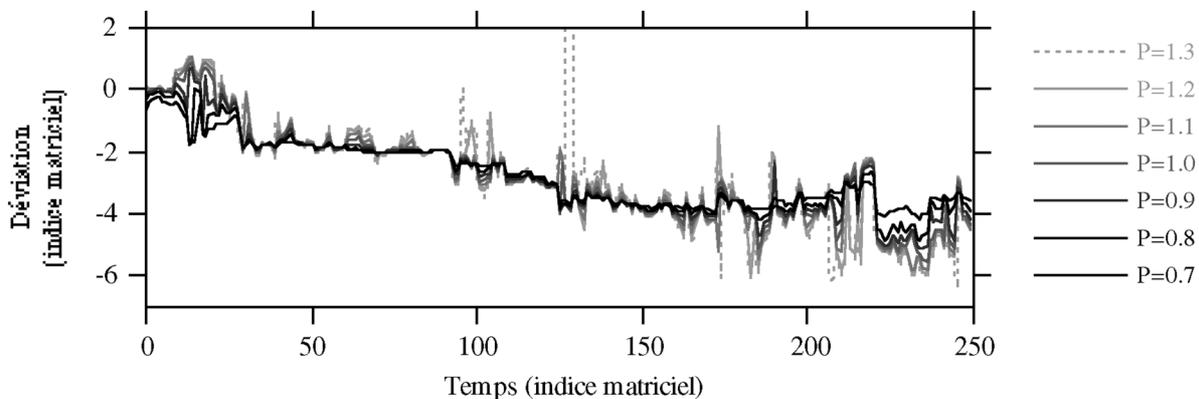


Figure 8. – Traces warp obtenues pour différentes valeurs du paramètre de pondération d’échelle P .

($P < 1,0$). Le facteur de pondération affecte significativement la trajectoire de la trace warp. Par exemple, la même analyse que celle illustrée à la figure 5 est reprise à la figure 8 et ce, pour différentes valeurs de pondération d'échelle P allant de 0,7 à 1,3. On constate une progression régulière dans le lissage de la trace warp avec la diminution du paramètre de pondération, une pondération inférieure à l'unité réduisant l'importance des corrélations sur les détails plus fins. On peut ainsi limiter ou encourager les grands déplacements dans le recalage temporel de la signature. Le paramètre « pondération » permet donc d'ajuster l'amplitude des déviations de la trace warp selon le contexte d'utilisation.

4. Comportement des algorithmes DTW et CMED

Comme nous n'avons pas développé un algorithme DTW qui pouvait traiter l'enveloppe d'une signature vibro-acoustique, seules des signatures temps-fréquence seront utilisées pour comparer ces algorithmes. La figure 9 expose les déviations temporelles estimées par les algorithmes DTW et CMED, et ce, toujours pour les mêmes signatures exposées aux figures 1 et 5. Dans cet exemple, comme dans plusieurs autres, l'algorithme DTW oscille dans son estimation de la déviation temporelle. Cette oscillation s'explique par deux contributions. Premièrement, l'algorithme DTW n'effectuant pas d'interpolation, la trajectoire de la fonction de recalage doit passer par une valeur entière d'indice matriciel : la trajectoire oscille alors autour de la valeur vraie, l'erreur crête maximale étant de 1/2 pas temporel. Deuxièmement, les signatures de changement de prises contiennent des discontinuités sur une petite échelle, soit de

l'ordre de la milliseconde ou de quelques pas temporels, en rappelant qu'un pas équivaut à 43 échantillons dans nos exemples. Comme l'algorithme DTW ne dispose pas d'une pondération d'échelle, ces discontinuités ne sont pas lissées et affectent la trajectoire de la fonction de recalage temporel. Par exemple, l'écart atteint 4 pas temporels à l'indice 210 sur la figure 9.

Afin de détailler davantage le comportement de ces algorithmes, nous avons altéré une signature et comparé celle-ci à l'original. Dans un premier temps, nous avons repris les signatures utilisées pour produire la figure 6 où une distorsion temporelle continue est induite numériquement. Cette distorsion étant connue, sur la figure 10 nous affichons la déviation correspondante de pair avec les estimations réalisées par la CMED et l'algorithme DTW. Dans cet exemple, la CMED suit la déviation avec un écart type moyen de 0,38 pas alors que l'algorithme DTW a un écart type de 2,44 pas, soit plus de six fois supérieur. L'écart crête atteint 5 pas pour la DTW : la précision obtenue est du même ordre que la déviation temporelle induite. Dans cet exemple, si l'on lisse les détails des déviations, en fixant adéquatement le paramètre de pondération P , la CMED affiche un résultat plus précis que l'algorithme DTW. Pour les mêmes signatures comparées, mais avec les deux signatures décalées de 10 tranches temporelles, la DTW livre un résultat très proche du précédent alors que la CMED livre un résultat moins précis et qui s'approche de celui de la DTW : avec la CMED le résultat apparaît fonction de la correspondance entre la position de certaines transitoires et la position des coupes dyadiques. Une solution envisagée serait d'appliquer une fenêtre spectrale au lieu du créneau dans la coupe multiéchelle. La coupe serait alors effectuée avec une progression moins rapide que la coupe dyadique. De plus, un choix adéquat de progression permettrait de ne pas aligner les positions des coupes sur les différentes échelles, soit à titre d'exemple, une progression de 5/7 au lieu de 1/2. Ainsi, les hachures de la coupe dyadique s'estomperaient en partie au prix d'un effort accru de calcul.

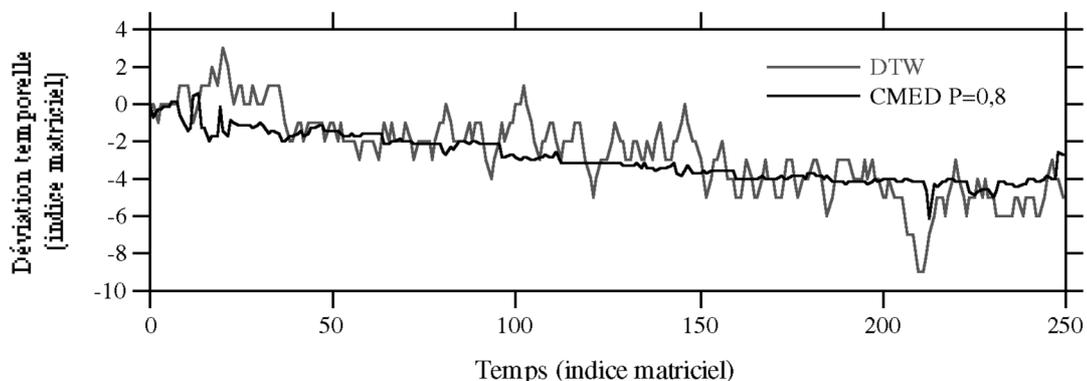


Figure 9. – Déviation temporelle observée par les algorithmes CMED et DTW dans la comparaison des signatures de changement de prises illustrées au haut de la figure 5.

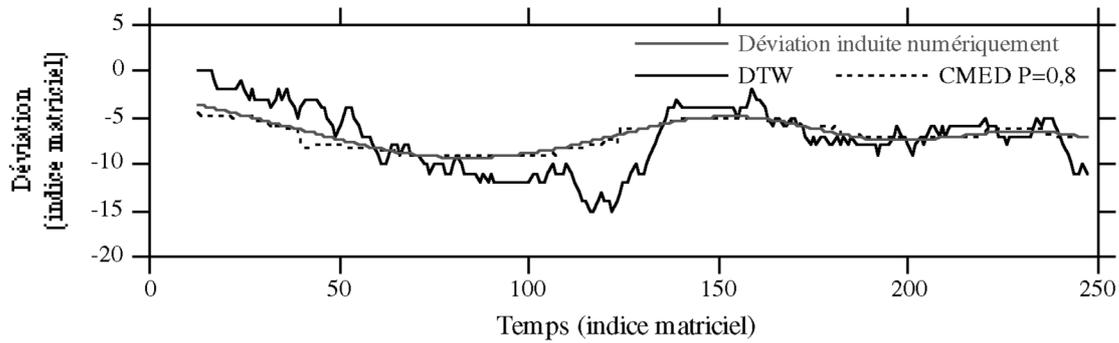


Figure 10 : Déviation temporelle estimée par les algorithmes CMED et DTW dans la comparaison d'une signature où l'on a inséré numériquement une distorsion temporelle.

Nous avons ajouté un bruit gaussien au signal d'accélération à la signature distordue pour juger de la sensibilité au bruit. De plus, nous avons introduit la distorsion temporelle dans le calcul de la signature temps fréquence en remplaçant le pas constant séparant deux tranches spectrales consécutives par un pas modulé par cette distorsion. De cette façon, la distribution spectrale des transitoires reste inchangée permettant ainsi un recalage plus précis. La figure 11 présente les estimations des déviations temporelles obtenues pour différentes grandeurs d'écart type. L'enveloppe du signal donné au haut de la figure 11 illustre la correspondance entre la position des transitoires et les sauts

observés dans la trace warp avec l'ajout du bruit. La CMED livre une trace warp qui se détache par sauts importants là où le signal est de plus faible amplitude. La DTW présente un résultat plus robuste lorsque le signal de faible amplitude est de courte durée. Par contre, à partir de la tranche temporelle 170, la DTW décroche complètement pour cette plage où la durée du signal de faible amplitude est importante.

Il est important de souligner que, dans d'autres essais, ce n'est pas seulement en fin de transitoire, là où l'amplitude est tenue et que le bruit domine, qu'il y a méprise dans la trajectoire de la trace warp. En effet, en présence de transitoires similaires et

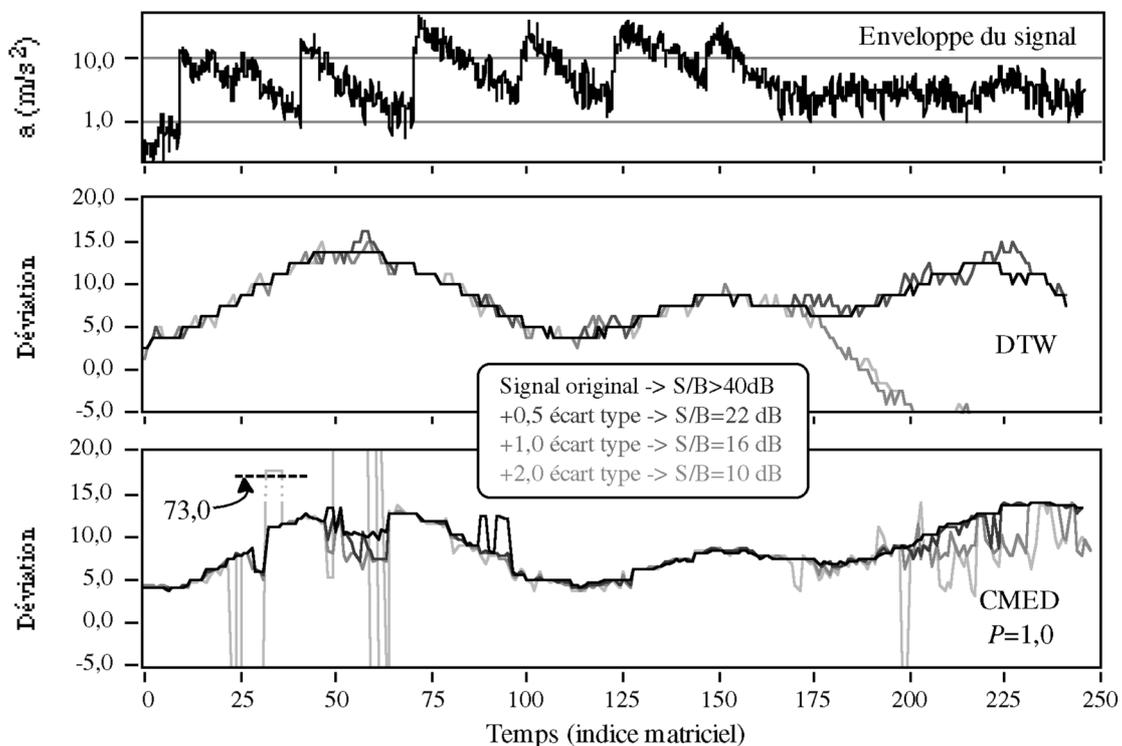


Figure 11. – Sensibilité au bruit des algorithmes CMED et DTW dans la comparaison d'une signature où l'on a inséré numériquement une distorsion temporelle et un bruit blanc.

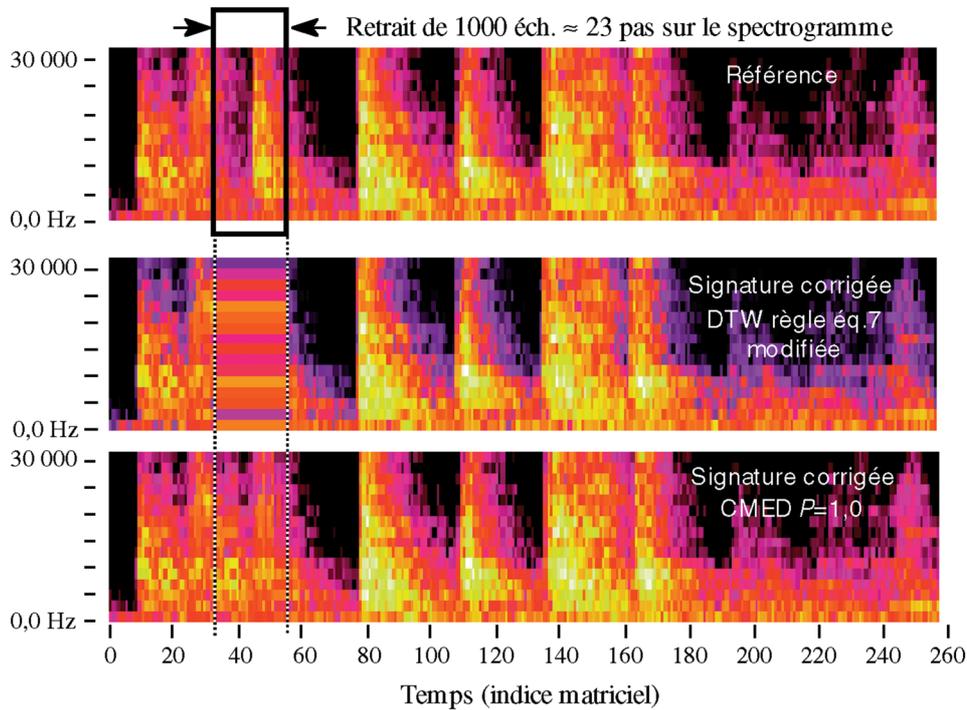


Figure 12. – Spectrogrammes illustrant le comportement des algorithmes CMED et DTW lors d’un découpage de 1000 échantillons temporels sur la signature à corriger.

successifs, il suffit de peu de bruit pour confondre deux transitoires. Il s’agit que le bruit modifie l’un de ceux-ci pour qu’il ressemble à l’autre et l’algorithme d’extraction de la trace warp hachurera la trace pour qu’elle passe par le transitoire contaminé. Encore ici, diminuer le facteur de pondération P réduit considérablement l’occurrence de sauts intempestifs. Cependant, cet exemple soulève une lacune dans l’extraction de la trace quand celle-ci hésite entre deux sommets de corrélation : la grandeur du saut à réaliser n’est pas mise en perspective avec le gain de corrélation réalisé avec ce saut. Précisons que ce n’est pas la CMED qui ici est en défaut mais bien l’algorithme d’extraction de la trace.

Dans un second temps, nous avons introduit une discontinuité temporelle dans une signature. Différents tests [22] démontrent que la signature n’est pas bien traitée par l’algorithme DTW au voisinage de la discontinuité. Ici, nous reproduisons un test qui consiste simplement à couper une portion de la signature brute, soit retirer 1 000 échantillons temporels sur une signature de 11 000 échantillons. Exprimé en indice matriciel, ce retrait revient à créer une discontinuité de 23,25 pas sur la signature temps-fréquence telle qu’illustrée à la figure 12. La signature à corriger n’est pas illustrée sur cette figure : pour la visualiser, il suffit de retirer le contenu du rectangle sur la signature de référence et d’accoler les sections restantes. Sur cette même figure, on observe que l’algorithme CMED cherche à reconstruire le transitoire qui a été effacé alors que l’algorithme DTW ne fait que répéter la même tranche spectrale jusqu’à la correction de la

discontinuité. Dans ce dernier exemple, la discontinuité est d’une grandeur inférieure (23 pas) à la demi-largeur (30 pas) du tampon contenant les trajectoires les plus probables retenues par l’algorithme DTW. Cependant, dans le cas d’une discontinuité commandant une excursion loin à l’extérieur du tampon, l’algorithme DTW décroche, la fonction de recalage temporel $w(n)$ progresse alors aléatoirement et ne peut garder le cap pour rattraper l’écart généré par la discontinuité.

Toujours sur la figure 12, il est fort intrigant de voir comment la CMED tente d’assembler le contenu spectral avoisinant la discontinuité afin de recréer la composante effacée par le retrait des échantillons temporels. L’examen des estimations des déviations temporelles présentées à la figure 13 nous montre comment la CMED se débrouille pour combler le « trou », le vide d’information, en modulant déviation temporelle. Cette modulation permet de synthétiser une portion de spectrogramme qui s’apparente à celle détruite à partir des caractéristiques spectrales au voisinage de la discontinuité. Évidemment, inventer quelque chose et la disposer là où il n’y devrait rien avoir ne constitue pas une solution. En fait, il n’y a pas de solution possible pour ce test dans l’intervalle où l’on doit combler une discontinuité. Cet exemple permet surtout de démontrer que la CMED ne présente pas de glissement de la réponse à une discontinuité, contrairement à l’algorithme DTW : la surface séparant ces deux estimations de déviation détermine le gain obtenu avec la CMED. Enfin, notons qu’une fois passée la discontinuité, la CMED se plaque sur la valeur non entière de la discontinuité,

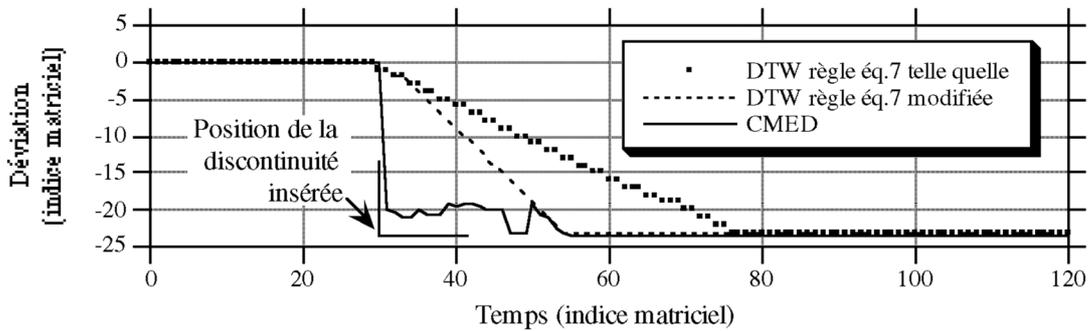


Figure 13 . – Comportement des algorithmes CMED et DTW lors d’un découpage de 1000 échantillons temporels avant calcul de la signature temps-fréquence.

soit à 23,25 pas, alors que l’algorithme DTW reste systématiquement 0,25 pas au côté de la valeur vraie.

La figure 14 présente un exemple d’inversion de l’ordre d’arrivée de deux transitoires. Dans cette simulation numérique, le découpage s’est aussi effectué sur le vecteur d’échantillons temporels, avant le calcul du spectrogramme. Pour un tel découpage, l’algorithme DTW ne peut corriger du repliement temporel où le temps s’écoule en sens inverse par endroits. Par contre, avec un paramètre de poids inter-échelle P ajusté convenablement, l’algorithme CMED passe ce test. Notons que les deux transitoires faisant l’objet de l’inversion apparaissent dissemblables sur le spectrogramme. En fait, des transitoires trop similaires n’auraient pu être remis dans le bon ordre par l’algorithme CMED. Il y a donc la notion de dissemblance entre les objets temps-fréquence qui intervient sur les différentes échelles. La pondération P permet seulement de choisir entre les échelles petites ou grandes pour commander le découpage temporel. L’introduction d’une pondération $P(k)$ fonction de la dimension pourrait s’avérer utile lorsque les objets ont des dissimilarités sur un nombre réduit de dimensions intermédiaires. On soulignerait ainsi les différences entre ces transitoires. Inversement, si la connaissance du phénomène physique commande une non-inversion de ces transitoires, le paramètre de pondération P , ou la fonction de pondération $P(k)$, devra être ajusté pour que cette inversion ne se produise pas.

Ces deux derniers tests constituent des cas difficiles démontrant le bon comportement de l’algorithme proposé. Cependant, l’algorithme conventionnel DTW dispose d’un avantage sur le plan de la simplicité et de l’effort réduit de calcul. Si ce dernier avantage est essentiel pour le traitement en continu de la parole, il n’est pas requis pour notre application où l’on peut se permettre de traiter l’information bien après la mesure. Précisons que le temps de calcul requis avec l’algorithme DTW est de moins de quelques secondes alors que l’algorithme de CMED prend plus de 5 minutes d’exécution (CPU M604 de 200 MHz) pour ce dernier exemple. La croissance du temps de calcul de l’algorithme DTW optimisé est linéaire avec la dimension de la signa-

ture alors que la CMED a un effort de calcul qui croît en $N^3 \ln(N)$ sans optimisation et en N^3 avec une optimisation partielle, où N est le nombre de pas temporels de la signature. L’accroissement de l’espace mémoire requis croît respectivement en N et en N^2 pour ces deux algorithmes. Entre autres, 43 Méga-octets sont requis dans LabVIEW™ pour rouler l’algorithme sur des signatures de 256 pas. Soulignons que LabVIEW™ est lent et inefficace pour ce type de traitement : l’implantation sous Fortran, Pascal ou C++ réduirait notablement la mémoire et le temps de calcul. Ces algorithmes sont publics et disponibles sur les sites ftp.ireq.ca/pub/fleonard/dtw et ftp.ireq.ca/pub/fleonard/cmmed.

5. Conclusion

Dans un système de surveillance d’équipements de commutation électrique de haute puissance, le recalage temporel permet de distinguer les déviations provenant de la distorsion temporelle, des déviations provenant de l’amplitude associée à chaque transitoire. L’acuité de l’algorithme de recalage a un impact direct sur la qualité du diagnostic. Dans ce contexte de comparaison de signatures vibro-acoustiques, l’algorithme proposé affiche un comportement robuste et estime la distorsion temporelle plus précisément que l’algorithme DTW. De plus, cet algorithme peut corriger une discontinuité de la trame temporelle et même une inversion dans une séquence d’apparition de transitoires. Plusieurs limitations présentes chez l’algorithme DTW proviennent d’un ensemble limité de trajectoires permises pour la fonction de recalage temporel. Or, la corrélation multiéchelle donne accès à un plus vaste ensemble de solutions, le prix à payer étant un effort de calcul accru. Enfin, l’algorithme DTW a atteint sa maturité alors que l’algorithme proposé est à ses débuts : il reste à explorer plusieurs raffinements qui amélioreraient la corrélation multiéchelle et le module d’extraction de la trace warp.

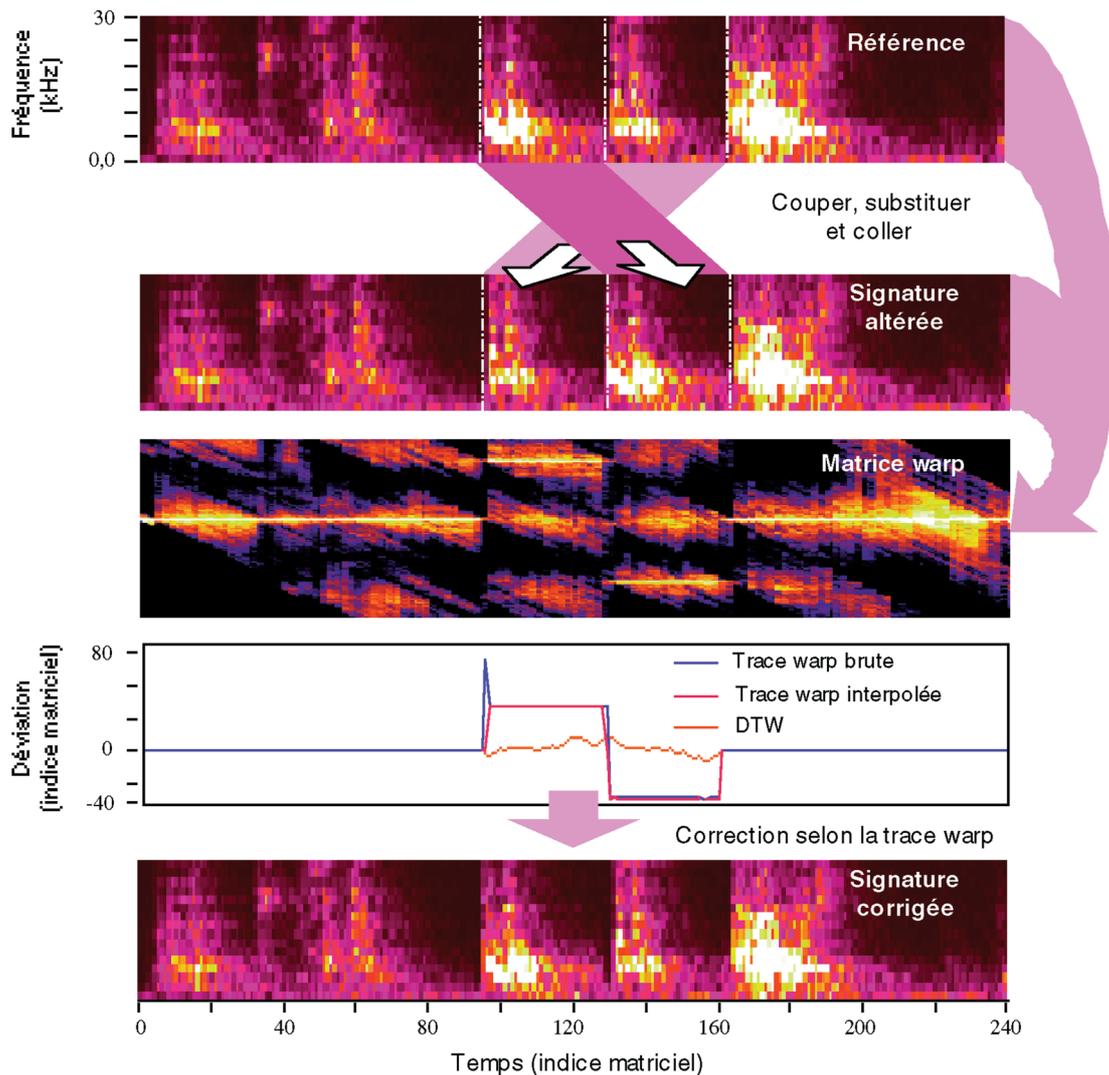


Figure 14. – Comportement des algorithmes CMED et DTW lors d’une inversion dans l’ordre d’arrivée de deux transitoires avant calcul de la signature temps-fréquence.

6. Remerciements

Soulignons la contribution des autres membres de l’équipe travaillant à la mise au point d’un système de surveillance de changeur de prises, soit messieurs Marc Foata, chercheur et responsable de projet, Mohamed Farhat, chercheur, Réal Beauchemin et Jean-Yves Paquin, techniciens. Mentionnons l’existence d’une collaboration de plus de trois années entre ABB Corp. et Hydro-Québec. À cet effet, nous remercions Tord Bengtsson et son équipe de ABB Corp. pour la qualité des échanges techniques entretenus. Enfin, nous exprimons notre gratitude à l’endroit de Claude Rajotte de TransÉnergie, Hydro-Québec, pour avoir supporté le développement de ce système.

BIBLIOGRAPHIE

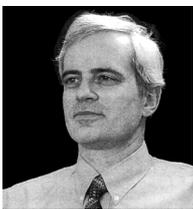
- [1] M. Belanger, « Transformer Diagnosis : Part 1. A statistical Justification For Preventive Maintenance », *Electricity Today*, p. 6-8, juin 1999.
- [2] T. Bengtsson & al., « Acoustic diagnosis of tap changers », Cigré 1996, conférence 12-101, Paris.
- [3] M. Foata, T. Bengtsson, « Vibration monitoring of tap changers and windings », Canadian Electrical Association, report CEA No. 474 T 955, juillet 1997.
- [4] F. Léonard, M. Foata, C. Rajotte, « Procédé de traitement de signatures vibro-acoustiques dans un système de commutation électromécanique haute-tension », Brevet canadien no. 5024.330, 8 février 1999. Brevets en cours de dépôt à l’international.
- [5] F. Léonard, « La transformée en objets et les vers temps-fréquence », Troisième conférence internationale sur les méthodes de surveillance et

- techniques de diagnostic acoustiques et vibratoires, tenue au CETIM de Senlis, France, octobre 1998.
- [6] H. Sakoe and S. Shiba, « A dynamic programming approach to continuous speech recognition », Proc Int. Cong. Acoustics, Budapest, Hungary, Paper 20C-13, 1971.
- [7] F. Itakura, « Minimum prediction residual principle applied to speech recognition », IEEE Transactions on acoustics, speech, and signal processing, Vol. 23, No. 1, pp. 67-72, février 1975.
- [8] H. Sakoe and S. Shiba, « Dynamic programming algorithm optimization for spoken word recognition », IEEE Transactions on acoustics, speech, and signal processing, Vol. 26, No. 1, pp. 43-49, février 1978.
- [9] A.R. Brenner, K. ECK, G. Engelhard and T.G. Noll, « Phase aberration correction using dynamic time warping », IEEE Ultrasonics symposium, Vol. 2, pp. 1361-1365, 1995.
- [10] R. Martens, L. Claesen, « Dynamic Programming Optimisation for On-line Signature Verification », IEEE, Proc. of the Fourth Con. on Document Analysis and Recognition, Vol. 2, pp. 653-656, 1997.
- [11] S. B. Dolins, J. D. Reese, « A Curve Interpretation and Diagnostic Technique for industrial Processes », IEEE Trans. on Industry Applications, Vol. 28, No. 1, pp. 261-267, 1992.
- [12] M. Runde, T. Aurud, L. E. Lundgaard, G. E. Ottesen, K. Faugstad, « Acoustic diagnosis of high voltage circuit-breakers », IEEE Transactions on Power Delivery, Vol. 7, No. 3, juillet 1992.
- [13] M. Runde, T. Aurud, L. E. Lundgaard, G. E. Ottesen, K. Faugstad, « Vibration analysis for diagnostic of circuit-breakers », IEEE/PES Winter meeting, janvier 1996.
- [14] D. Burr, B. Ackland, and N. Weste, « Array configurations for dynamic time warping », IEEE Transactions on acoustics, speech, and signal processing, Vol. 32, pp. 119-127, février 1984.
- [15] M. J. Irwin, « A digit pipelined dynamic time warp processor », IEEE Transactions on acoustics, speech, and signal processing, Vol. 36, pp. 1412-1422, septembre 1988.
- [16] G. N. Stainhaouer et G. Carayannis, « New Parallel Implementations for DTW Algorithms », IEEE Transactions on acoustics, speech, and signal processing, Vol. 38, No. 4, avril 1990.
- [17] L. R. Rabiner et R. W. Schafer, Digital processing of speech signals, Prentice-Hall Inc., p. 478-485, 1990.
- [18] L.R. Rabiner, A.E. Rosenberg, and S.E. Levinson, « Considerations in dynamic time warping for discrete word recognition », IEEE Transactions on acoustics, speech, and signal processing, Vol. 26, No. 6, pp. 575-582, décembre 1978.
- [19] R. W. Schafer et L. R. Rabiner, « A Digital Signal Processing Approach to Interpolation », Proc. of IEEE, Vol. 61, No. 6, p. 692-702, juin 1973.
- [20] F. J. Harris, « On the use of windows for harmonic analysis with the discrete Fourier transform », Proc. of IEEE, Vol. 66, No. 1, p. 51-83, 1978.
- [21] C. E. Shannon, « Communication in Presence of Noise », Proceeding of the Institute of Radio Engineers, Vol. 37, No. 1, pp. 10-21, janvier 1949.
- [22] F. Léonard, « Comparaison de signatures vibro-acoustiques et repliement temporel par corrélation multiple de type radix 2 », Hydro-Québec, rapport IREQ-96-135, mai 1996.

Manuscrit reçu le 5 mars 1999.

LES AUTEURS

F. LÉONARD



François Léonard est né le 9 mars 1956. Gradué de l'École polytechnique de Montréal en génie physique, il y termina sa maîtrise en 1979. Depuis, il occupe un poste de chercheur à l'institut de recherche d'Hydro-Québec (IREQ). Œuvrant principalement dans le domaine du traitement du signal mécanique et du diagnostic, il développe ses propres outils, dont certains pour l'analyse temps-fréquence et d'autres pour la statistique, tel le krigeage.