

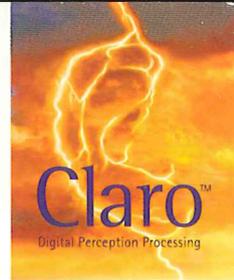
Claro AutoSelect

Classification des sons
pour un système de
gestion automatique,
intelligent des
programmes multiples

Introduction

Il est aujourd'hui largement admis que des environnements auditifs différents requièrent des modes différents de traitement du signal. Les conditions essentielles pour une communication optimale dans le calme sont l'audibilité et la bonne qualité sonore, alors que dans le bruit il faut surtout améliorer le rapport du signal au bruit pour assurer une meilleure intelligibilité vocale. Les aides auditives à programme unique ne fonctionnent bien que dans un nombre limité de situations. Une stratégie plus efficace est l'emploi d'appareils à programmes multiples avec des modes différents de traitement du signal, sélectionnés selon l'environnement acoustique. De tels appareils offrent des programmes dont les réglages sont significativement différents les uns des autres et sont spécialement conçus pour agir de façon optimale dans différents milieux acoustiques.

Plusieurs études ont montré que les utilisateurs préfèrent que leurs aides auditives disposent de courbes de réponse différentes pour les utiliser dans différents environnements acoustiques (Fabry et Stypulkowski, 1992; Keidser, 1996; Keidser, 1998).



Kuk (1994) a trouvé que les utilisateurs préféraient avoir des courbes de réponses adaptées à la situation acoustique plutôt que des réponses fixes. De nombreuses autres études montrent les avantages dans le bruit des multi-microphones utilisés dans des appareils à programmes multiples (Valente et al., 1995; Kuk, 1996a; Kuk, 1996b; Lurquin et Rafhay, 1996; Voss, 1997; Gravel et al., 1999; Kuehnel et al., 1999; May, 1999; Schuchman et al., 1999; Valente et al., 1999; Pumford et al., 2000). Dans une enquête de satisfaction des utilisateurs d'aides auditives «high-tech» (Kochkin, 1996), les programmes multiples se sont révélés être un trait caractéristique significatif qui contribue à la satisfaction des malentendants. Les appareils auditifs à mémoires multiples répondent parfaitement à cet objectif. Cependant, l'inconvénient majeur des appareils à programmes multiples est que c'est l'utilisateur lui-même qui a la charge de sélectionner manuellement le programme auditif approprié, soit à l'aide d'une télécommande, soit avec un commutateur sur l'appareil. Cela suppose qu'il puisse reconnaître à quel moment la sélection est nécessaire.

La toute dernière génération de technologie numérique a rendu possible le développement d'algorithmes intelligents, capables de commuter automatiquement entre des programmes. Les aides auditives à programmes multiples sont ainsi devenues encore plus efficaces et plus simples d'emploi. La procédure de sélection automatique AutoSelect, utilisée dans les ordinateurs auditifs Claro, s'appuie sur une analyse qu'avait proposée Kates (Kates, 1995). Le critère de commutation ne s'appuie pas sur un paramètre unique tel que le niveau global, mais exploite 4 dimensions différentes des caractéristiques du signal. D'autres algorithmes de classification des signaux acoustiques à des fins de réduction du bruit ont été décrits (Ostendorf et al., 1997; Ostendorf et al., 1998; Ludvigsen, 1993) mais n'ont pas encore été utilisés à ce jour pour cette application.

Cet algorithme de classification des sons est utilisé dans Claro pour commuter auto-

matiquement entre deux programmes auditifs, l'un conçu pour le calme et l'autre pour les milieux bruyants. Dès que la commutation a été activée, le mode de traitement du signal est modifié. Par exemple, le réducteur de bruit à haute résolution (Fine-scale Noise Canceler) de même que l'AudioZoom numérique adaptatif (Adaptive digital AudioZoom) sont activés dans le bruit. Ceci garantit l'audibilité, l'intelligibilité, le confort auditif et la qualité sonore optimale, dans le bruit comme dans le calme. AutoSelect ne commutera dans le programme conçu pour le bruit que si de la parole est détectée dans un bruit large bande. En d'autres termes, la commutation n'a pas lieu dans des situations où elle n'est pas souhaitable (bruits de la rue, concert, ...). De plus, il est toujours possible de «court-circuiter» manuellement l'automatisme, soit à l'aide d'un commutateur sophistiqué sur l'appareil lui-même, soit avec la télécommande SoundPilot.

AutoSelect – classification des sons

Afin de pouvoir commuter automatiquement entre différents programmes auditifs, il faut disposer d'un algorithme puissant de classification des sons. Classifier les sons, dans ce cas, signifie grouper tous les signaux d'entrée qui se présentent dans la vie quotidienne en deux classes distinctes. L'objectif de l'algorithme de classification des sons est de distinguer la parole dans un bruit large bande des autres signaux. Pour atteindre cet objectif, il est nécessaire d'étudier de près les caractéristiques qui rendent la voix humaine tellement unique.

Le tracé supérieur de la *figure 1* représente la structure temporelle de 5 secondes d'un signal vocal continu pur dans le calme et le tracé inférieur représente la même phrase noyée dans un bruit de party avec un rapport du signal au bruit de 0 dB. On voit immédiatement que la parole pure a une structure temporelle claire. Dans le bruit, ses fluctuations sont en partie masquées, car les pauses vocales sont «comblées» par le bruit.

La parole est constituée de phrases, de mots, de syllabes et de phonèmes, les plus petits éléments vocaux. Cette structure unique peut se retrouver dans différentes propriétés physiques utilisées pour décrire un signal acoustique, telles que: la profondeur de modulation, la fréquence de modulation, les fluctuations du niveau du signal ou des propriétés statistiques telles que l'histogramme des amplitudes. A l'inverse, la plupart des signaux bruyants non vocaux n'ont pas la même structure caractéristique que celle de la parole pure.

Critères de classification

Niveau global d'entrée

Le niveau global d'entrée du signal participe à la classification AutoSelect. Il contient des



Figure 1
Cette figure représente, en haut, la structure temporelle d'un signal vocal pur de 5 secondes et, en bas, le même signal vocal, noyé dans du bruit, avec un S/B de 0 dB. Dans ce cas, les pauses vocales sont «comblées» par le bruit.

informations essentielles sur l'environnement acoustique. Les conversations se déroulent en général dans des environnements calmes, où le niveau vocal moyen est d'environ 60-65 dB SPL; nous élevons la voix seulement en présence de bruit perturbant. Un seuil de niveau global peut donc être défini comme critère de commutation pour les situations où le bruit affecte la compréhension de la parole. La figure 2 donne des exemples de niveaux globaux pour de la parole, le bruit de la rue et de la parole dans un brouhaha.

Force de fluctuation du niveau global

La fluctuation du niveau global du signal est la seconde caractéristique utilisée dans le système AutoSelect de Claro. Les fluctuations sont une représentation des variations temporelles à court terme du signal et sont comparables à de la modulation. Pour calculer la force de fluctuation, la moyenne à court terme du signal est comparée à sa moyenne à long terme. Des différences importantes indiquent la présence de fortes fluctuations. Si un signal est pratiquement constant dans le temps, les deux moyennes sont presque identiques ce qui révèle l'absence de fluctuations (comme dans le bruit de la rue).

La parole pure varie fortement à court terme de par sa structure syllabique et fluctue de ce fait beaucoup plus que le bruit, la plupart du temps entretenu, et dont le niveau est pratiquement constant dans le temps. Les

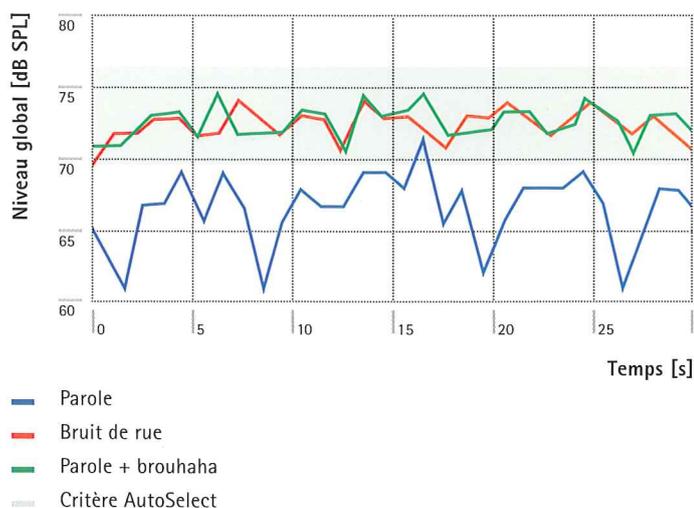


Figure 2
Niveau global de trois signaux différents dans une fenêtre temporelle de 30 secondes. Le niveau global de la parole reste inférieur à celui du bruit de rue ou de la parole dans un brouhaha. La grille utilisée pour distinguer la parole dans le bruit selon ce critère est représentée par la zone ombrée. Pour être qualifié de bruit, le niveau global du signal doit se situer dans cette zone. D'autres critères sont nécessaires pour distinguer la parole dans un bruit large bande des autres environnements sonores.

fluctuations de la parole noyée dans un bruit de fond, quant à elles, s'estompent par la présence du bruit qui «comble» les pauses et les lacunes du signal vocal, comme le montre la *figure 1*. Le niveau moyen de la force de fluctuation (défini comme étant l'écart type du niveau moyen divisé par le niveau moyen) peut donc être un critère qui permet de distinguer la parole dans un bruit large bande des autres signaux. Les fluctuations sont moindres que pour la parole pure et plus importantes que pour un bruit entretenu tel que le bruit de la rue.

Centre spectral

Le centre spectral est la «fréquence moyenne» du signal ou son centre de gravité spectral. Pour la parole, il se situe typiquement autour de 400 à 500 Hz, et il est presque identique pour des locuteurs masculins et féminins. Pour le bruit ou la musique, il peut varier dans une large gamme de fréquences (par exemple, bruit de voiture, aérateur, bruit de cafétéria, bruit de machines ou de perceuses etc.). Cela fournit d'autres informations sur le type de signal et contribue à une classification correcte.

Force de fluctuation du centre spectral

La force de fluctuation du centre spectral quantifie la fluctuation spectrale du signal. Elle est grossièrement définie comme étant le quotient de l'écart type du centre spectral et de la moyenne des centres spectraux. La *figure 3* donne un exemple des fluctuations du centre spectral pour de la parole, de la musique et de la parole dans un brouhaha. De même que pour les fluctuations du niveau du signal, le centre spectral fluctue plus pour la parole pure que pour la musique ou que pour la parole mêlée à un brouhaha large bande.

Matrice de classification des sons AutoSelect – une combinaison des quatre dimensions / caractéristiques du signal

Pour classifier de façon claire et stable la parole dans un bruit large bande, les quatre caractéristiques de signal sont analysées et exploitées comme critères de commutation.

Les quatre grandeurs doivent répondre ensemble à un critère prédéfini pour éviter les commutations intempestives et inutiles. La *figure 4* représente le mode de fonctionnement de la matrice de classification des sons AutoSelect. Pour des raisons pratiques, cet exemple ne montre que deux des quatre dimensions, à savoir la force de fluctuation du niveau du signal par rapport à la fluctuation du centre spectral, pour différents

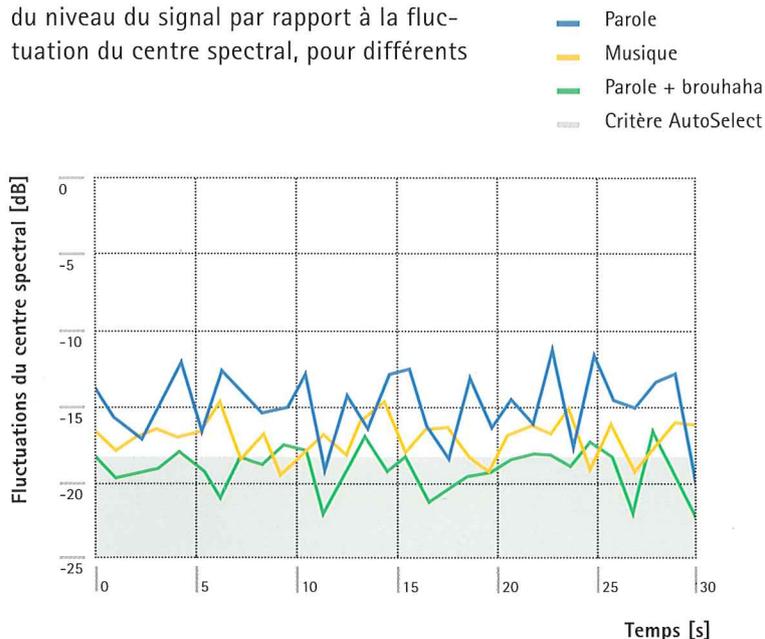


Figure 3

Fluctuations du centre spectral d'un échantillon de 30 secondes de différents signaux sonores. La fluctuation est définie comme étant le quotient de l'écart type du centre spectral et du centre spectral, exprimé en dB. La parole diffère nettement de la musique ou de la parole dans un brouhaha, car elle présente de plus fortes fluctuations de son centre spectral. La zone ombrée indique la grille qui peut être utilisée pour classer le signal comme «de la parole dans un bruit large bande».

types de signaux, y compris du bruit et de la parole dans le bruit à différents S/B. Avec deux critères, le signal vocal pur se trouve dans une région qui est clairement distincte de tous les autres signaux. La zone ombrée indique la grille de sélection utilisée pour caractériser la parole pure selon ces deux critères. Quatre critères sont nécessaires pour identifier de la parole dans un bruit spécifique.

Si une situation bruyante qui risque d'affecter l'intelligibilité vocale est détectée, le programme est commuté sur NoiseAdapt et le traitement du signal est modifié pour répondre spécifiquement à cette situation (DPP NoiseAdapt). De plus, le réducteur de bruit à haute résolution et l'AudioZoom numérique adaptatif sont activés. Cela améliore le rapport du signal au bruit et donc le confort et l'intelligibilité pour l'utilisateur. Comme des temps d'adaptation assez lents sont préférables pour modifier les caractéristiques de l'aide auditive (Köbler et Leijon, 1999), AutoSelect a un temps de réaction de l'ordre de 10 secondes. Ceci évite toute commutation intempestive si l'utilisateur se déplace rapidement d'un environnement à l'autre.

Résumé

Le système de classification des sons AutoSelect de Claro distingue clairement la parole dans un bruit large bande des autres signaux, autorisant une commutation automatique et intelligente des programmes. Ceci garantit l'audibilité, l'intelligibilité, le confort auditif et la qualité sonore optimale dans le bruit comme dans le calme et réduit le besoin de commutation manuelle entre les programmes.

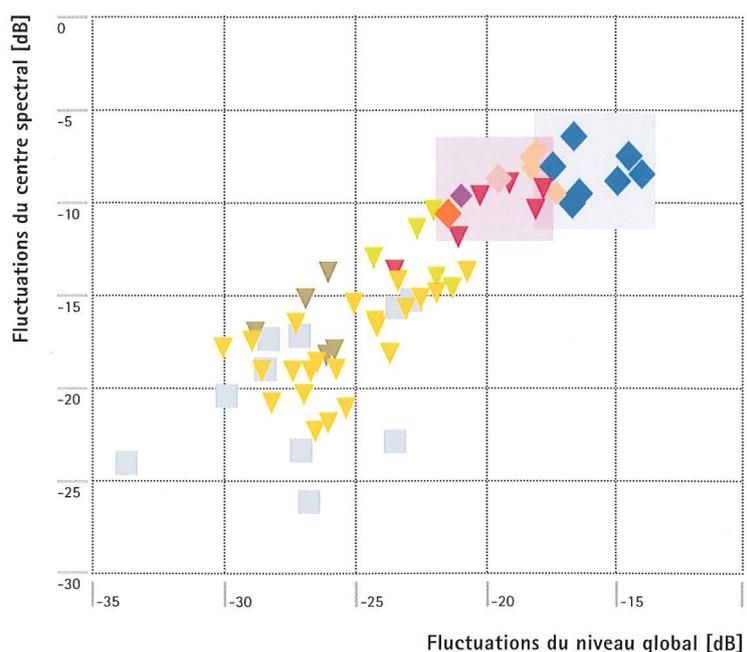


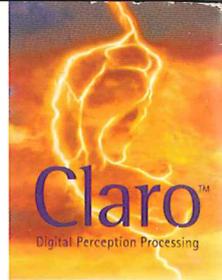
Figure 4

Cette figure représente deux des quatre dimensions utilisées pour classifier la parole dans un bruit large bande par rapport autres signaux. Différentes valeurs du S/B ont été choisies dans le cas de la parole dans le bruit. La zone ombrée montre que deux dimensions peuvent être utilisées pour distinguer la parole pure des autres signaux. L'ensemble des quatre dimensions permet une classification puissante de la parole dans un bruit large bande.

- ◆ Parole pure
- bruit
- ◆ 2 locuteurs, 0/7 dB S/B
- ◆ 1+2 locuteurs, 7 dB S/B
- ◆ 1+3 locuteurs, 7 dB S/B
- ◆ 1+4 locuteurs, 7 dB S/B
- ▼ Parole + brouhaha, 4/10/15 dB S/B
- ▼ Parole + bruit de party, 4/10/15 dB S/B
- ▼ Parole + bruit de voiture, 10/5/0/-5 dB S/B
- ▼ Parole + bruit, 4 dB S/B

■ 2 critères permettent d'identifier la parole pure

■ 4 critères sont nécessaires pour identifier la parole dans un bruit spécifique



Bibliographie

Fabry D. and Stypulkowski P. Evaluation of fitting procedures for multiple-memory programmable hearing aids. In: Paper presented at the annual meeting of the American Academy of Audiology, 1992.

Gravel J. S., Fausel N., Liskow C. and Chobot J. Children's speech recognition in noise using omnidirectional and dual-microphone hearing aid technology [In Process Citation]. *Ear + Hearing* 1999; 20: 1-11.

Kates J. M. Classification of background noises for hearing aid applications. *Journal of the Acoustic Society of America* 1995; 97: 461-70.

Keidser G. Selecting different amplification for different listening conditions. *Journal of the American Academy of Audiology* 1996; 7: 92-104.

Keidser G. Fitting multiple-memory hearing aids: A practical guideline. *The Hearing Journal* 1998; 51: 36, 38-39, 42-43, 46, 48.

Köbler S. and Leijon A. Noise analysis of real-life listening situations for maximal speech audibility in hearing aid fitting. *Scandinavian Audiology* 1999; 28(3): 179-189.

Kochkin S. Customer satisfaction and subjective benefit with high performance hearing aids. *The Hearing Review* 1996; 3: 16-26.

Kuehnel V., Margolf-Hackl S. and Kiessling J. Multi-microphone technology for severe-to-profound hearing loss. *Scandinavian Audiology Supplement* 2000; (in press).

Kuk F. A screening procedure for modified simplex in frequency-gain response selection. *Ear + Hearing* 1994; 15: 62-70.

Kuk F. Hearing aid survey tests user satisfaction. *Hearing Instruments* 1996a; 24-29.

Kuk F. Subjective preference for microphone types in daily listening environments. *The Hearing Journal* 1996b; Vol. 49.

Ludvigsen C. Schaltungsanordnung für eine automatische Regelung von Hörhilfsgeräten. *Deutsches Patentamt* 1993, DE 43 40 817 A1.

Lurquin P. and Raffhay S. Intelligibility in noise using multi-microphone hearing aids. *Acta oto-rhino-laryngologica belg.* 1996; 50: 103-109.

May A. Case Study on a Child's Use of Multi-Mic & FM Technology. *The Hearing Review* 1999; 6: 36, 38, 47.

Ostendorf M., Hohmann V. and Kollmeier B. Empirische Klassifizierung verschiedener akustischer Signale und Sprache mittels einer Modulationsfrequenzanalyse. In: *Fortschritte der Akustik, DAGA 1997*, Deutsche Gesellschaft für Akustik e.V., 1997: 608-609.

Ostendorf M., Hohmann V. and Kollmeier B. Klassifikation von akustischen Signalen basierend auf der Analyse von Modulationspektren zur Anwendung in digitalen Hörgeräten. In: *Fortschritte der Akustik – DAGA 98*, Deutsche Gesellschaft für Akustik e.V., 1998: 402-403.

Pumford J. M., Seewald R. C., Scollie S. and Jenstad L. Speech recognition with in-the-ear and behind-the-ear dual-microphone hearing instruments. *Journal of the American Academy of Audiology* 2000; 11(1); (in press).

Schuchman G., Valente M., Beck L. and Potts L. User satisfaction with an ITE directional hearing instrument. *The Hearing Review* 1999; 6: 12, 16, 21, 22.

Valente M., Fabry D. A. and Potts L. G. Recognition of speech in with hearing aids using dual microphones. *Journal of the American Academy of Audiology* 1995; 6: 440-449.

Valente M., Sweetow R. and May A. Using Microphone Technology to Improve Speech Recognition. *High Performance Hearing Solutions* 1999; 3: 10-13.

Voss T. Clinical evaluation of multi-microphone hearing instruments. *The Hearing Review* 1997; 4: 36, 45, 74.