



## UNE ANTENNE ACOUSTIQUE ADAPTATIVE, ASPECTS MATERIELS ET LOGICIELS.

Y.Grenier, M.Xu, J.Prado, D.Liebenguth.

ENST, Dept. SIGNAL,  
46 rue Barrault, 75634 Paris Cedex 13.

### RESUME.

Cette communication décrit une antenne acoustique adaptative destinée à la prise de sons. Nous traiterons de trois aspects : le contexte des applications visées, à savoir l'audio-conférence et la téléphonie mains libres, la description matérielle du prototype pour l'audio-conférence, et l'élaboration de l'algorithme retenu pour effectuer le traitement adaptatif.

### SUMMARY.

This paper describes an adaptive acoustic antenna designed for sound recording. We shall present three aspects of the design : specification in accordance to the applications which are audio-conferencing and hands-free telephone, description of the hardware prototype for audio-conferencing, and eventually, implementation of the algorithm which adapts the antenna.

### INTRODUCTION.

Les applications de prise de sons dans le contexte de l'audio-conférence sont caractérisées par le besoin d'enregistrer sans distorsion les signaux émis par un ou plusieurs locuteurs, tout en éliminant au maximum un brouilleur de niveau élevé qui est le signal venant de l'autre salle d'audio-conférence et qui est diffusé par haut-parleur. On cherche de plus un effet de "présence" dans l'enregistrement qui s'obtient en réduisant l'importance des trajets sonores réverbérés par rapport au trajet direct.

La réalisation matérielle d'un premier prototype utilise quinze microphones répartis sur un demi-cercle. A chacun d'eux est associée une chaîne d'acquisition comprenant pré-amplificateur, filtre passe-bas et convertisseur. Les voies sont multiplexées par groupe de quatre pour alimenter un Transputer T212. Ceux-ci servent ensuite de point d'entrée pour un réseau de trente deux Transputers T800 qui sont des processeurs flottants en 32 bits. La sortie du réseau alimente un transputer T222 qui pilote le convertisseur numérique analogique.

Les méthodes retenues pour atteindre les buts visés (débruitage, déréverbération, effet de présence) sont des traitements adaptatifs d'antennes. L'un d'eux combine une technique simplifiée de formation de voie adaptative avec une identification, elle aussi adaptative, de la contribution du brouilleur principal. Ce traitement s'implante dans le domaine fréquentiel. Un autre traitement est, dans le domaine temporel, la méthode dite "Generalised Sidelobe

Canceler", proposée par Frost et améliorée par Griffiths et Jim. Un point commun à ces deux approches est leur parallélisme intrinsèque. Nous exploitons ce parallélisme grâce à la structure du réseau configurable de Transputers, et au langage OCCAM. Nous présenterons et comparerons diverses solutions pour la répartition des traitements sur le réseau.

### CONTEXTE DES APPLICATIONS.

Nous visons un certain nombre d'applications du domaine d'intérêt de France Télécom, et dans lesquelles se fait sentir le besoin ou l'utilité d'une prise de sons au moyen d'un réseau de microphones. Ces applications sont de trois types. Elles concernent toutes les trois la téléphonie mains libres en un sens assez général, pour des salles d'audio-conférence, pour des cabines publiques de type Publivoix, pour le radio-téléphone.

Dans l'audio-conférence ou la vidéo-conférence, un groupe de locuteurs situés dans une salle converse avec un second groupe situé dans une salle distante. Le but de la prise de sons est bien entendu de pouvoir transmettre sans distorsion les signaux émis par chacun des locuteurs. Les particularités de cette application sont un niveau de bruit assez faible, mais une réverbération importante lorsque le trajet direct entre locuteur et microphone est de longueur supérieure à un ou deux mètres. Il est souhaitable de réduire cette réverbération afin d'améliorer l'effet de présence de la prise de sons.



A cette réverbération du signal utile s'ajoute celle du signal provenant de l'autre salle, qui risque ainsi d'être réinjecté dans les microphones. Un tel couplage acoustique est extrêmement nuisible, tant par la gêne auditive qu'il introduit, que par les risques d'instabilité ou d'accrochage du système qui s'ensuivent [1]. Le signal provenant de l'autre salle et émis par le haut-parleur constitue donc un brouilleur fort, pour reprendre la terminologie du filtrage d'antennes. On connaît parfaitement le signal émis par le haut-parleur. Par contre la contribution de ce brouilleur au signal capté par chaque microphone est inconnue car la réponse acoustique de la salle mesurée entre le haut-parleur et le microphone est elle aussi inconnue.

Une seconde application possible du réseau de microphones concerne les cabines téléphoniques publiques du type des cabines Publivox qui sont en cours d'expérimentation. Ces cabines suppriment le combiné, qui est remplacé par un couple microphone et haut-parleur incorporé dans la paroi. De plus la numérotation du correspondant se fait vocalement, grâce à un système de reconnaissance automatique de la parole [2]. La difficulté de la prise de sons dans ce contexte vient de la sensibilité au bruit du système de reconnaissance de parole. Il serait souhaitable de réduire le bruit capté par le microphone, afin d'éliminer les erreurs de numérotation. Les parois en verre de la cabine lui confèrent une assez forte réverbération qui peut entraîner une perte d'intelligibilité du signal vocal et un couplage acoustique entre haut-parleur et microphone.

Une troisième application est pour le radio-téléphone. Il est souhaitable de s'affranchir du combiné, afin de ne pas perturber la conduite automobile. Mais on introduit alors les mêmes ennuis que dans les précédentes applications, la situation étant somme toute assez semblable à celle de la cabine Publivox, par le taux de réverbération, par le rôle du haut-parleur comme brouilleur, par la numérotation vocale... Le niveau de bruit est cependant largement supérieur à l'intérieur d'un habitacle de voiture.

Ce que le réseau de microphones peut apporter par rapport à un microphone unique, dans ces trois applications tient aux propriétés connues de la formation de voie [3]. L'addition des bruits captés se faisant de manière incohérente, multiplier par deux le nombre des capteurs permet de gagner 3 dB sur le rapport signal à bruit, ainsi 8 microphones permettent de gagner une dizaine de dB. Un brouilleur fort reçu de manière cohérente sur le réseau sera éliminé dans un rapport encore plus grand. La partie réverbérée du signal utile et du brouilleur est elle aussi réduite [4].

### METHODES DE FORMATION DE VOIE.

Les traitements d'antenne se font habituellement par filtrage adaptatif [3]. La structure du système se compose, derrière chaque capteur, d'un filtre fixe (souvent à réponse impulsionnelle finie RIF) destiné à remettre en phase le signal utile sur tous les capteurs, suivi par un second filtre RIF, adaptatif celui-ci, les sorties de ces filtres étant sommées afin de constituer le signal en sortie de l'antenne. L'optimisation

des filtres adaptatifs se fait en maximisant un rapport signal à bruit, sous une contrainte. Nous avons considéré trois implantations de ce type de formation de voie.

Dans la première implantation, due à Frost [5], on minimise la variance du signal en sortie de l'antenne, en contraignant la réponse dans la direction visée à respecter un gabarit, par exemple en imposant un gain unitaire, à toutes les fréquences dans la bande, pour le signal provenant de cette direction. L'optimisation est réalisée par un algorithme de gradient sous contrainte. Mais cet algorithme est assez lourd, à cause de la contrainte qui doit être imposée à chaque pas et qui requiert un produit matriciel.

La seconde implantation est une variante de la structure précédente qui sépare l'application des contraintes et la minimisation du critère [6]. L'antenne comporte la formation d'un certain nombre de voies fixes, l'une satisfait la contrainte de gain, et les autres constituent autant de références du bruit permettant une soustraction adaptative du bruit capté dans la première voie. Cette structure a plusieurs intérêts: elle permet d'utiliser tout algorithme adaptatif non contraint, d'introduire des contraintes multiples, de réduire le nombre des voies en faisant une adaptation partielle, et enfin d'exploiter le parallélisme des diverses voies en employant des calculateurs parallèles. Cependant, la parallélisation de ce filtrage adaptatif soulève quelques difficultés que nous discuterons dans la dernière partie.

L'implantation de cette structure adaptative s'effectue également dans le domaine fréquentiel [4]. On emploie un prétraitement sous forme de transformée de Fourier par blocs, suivi d'un filtrage spatial à chaque fréquence, et d'une synthèse du signal par transformée de Fourier inverse. L'intérêt est de pouvoir régler séparément à chaque fréquence le diagramme de directivité de l'antenne. Le parallélisme intrinsèque de cette structure est très important, avec dans chaque bloc temporel, les FFT en parallèle sur chaque capteur, puis après un multiplexage entre la fréquence et l'espace, autant de formations de voie en parallèle qu'il y a de fréquences distinctes dans la bande analysée. A ceci s'ajoute la possibilité d'identifier un brouilleur connu affecté d'un filtrage dont la réponse impulsionnelle est très longue, situation fréquente dans les applications citées du fait de la réverbération. Là encore, l'identification se fait intrinséquement en parallèle pour chacune des fréquences de la bande.

### MATERIEL.

Nous avons choisi d'implanter les calculs de la formation de voie adaptative sur un réseau de Transputers INMOS. Le modèle T800 est un microprocesseur 32 bits qui incorpore une unité de calcul en réels flottants, 4 Koctets de mémoire rapide SRAM à 50 ns, et quatre liaisons bit-série, chacune ayant un débit de 20 Mbits/s. Les performances en calcul flottant de ce processeur sont bonnes (de l'ordre de 1 à 1.5 Mflops), mais restent bien en deçà de celles des circuits spécialisés pour le traitement du signal (NEC 77230, TMS C30...). Ce manque de performances est cependant compensé par le parallélisme interne des unités arithmétiques et

flottantes et des liaisons qui se font par DMA. Et surtout, les quatre liaisons permettent de construire par simple connexion des réseaux de Transputers qui fonctionneront en parallèle.

Nous avons retenu une machine disponible commercialement (société Archipel) pour la partie réseau de calcul dans notre prototype. Cette machine comporte 32 T800, avec pour chacun une mémoire de 1Mo (60 ns). La configuration du réseau peut se faire par câblage, ce qui est simple, peu onéreux, mais manque de souplesse, aussi avons-nous retenu une solution plus efficace: une carte de configuration utilisant le circuit INMOS C004 qui permet de connecter, par un commutateur de type "cross-bar", 32 liens en entrée vers 32 liens en sortie.

La structure du traitement implique de pouvoir injecter les signaux issus des microphones dans le réseau, et de pouvoir récupérer en sortie le signal issu du traitement adaptatif. Il nous fallait donc connecter des convertisseurs A/D et D/A à notre réseau. Nous avons pour cela développé une carte d'interface entre des convertisseurs et un Transputer. La conversion se faisant sur 16 bits, le transputer retenu est un T222 (16 bits). Les quatre liens de ce Transputer restent libres, le convertisseur étant vu par l'intermédiaire d'un registre FIFO, comme une partie de l'espace mémoire du T222. Ceci a aussi l'avantage de supprimer la gestion d'interruption quand on acquiert le signal par blocs (dans l'implantation fréquentielle de la formation de voie). Le convertisseur est du type PCM78 Burr-Brown, capable d'échantillonner sur 16 bits à la cadence de 200 kHz. Les considérations de temps d'acquisition ainsi que d'encombrement sur une carte de type PC nous ont conduits à multiplexer 4 voies sur chaque convertisseur.

Une contrainte importante de notre application est la nécessité d'échantillonner en synchronisme les 15 signaux microphones ainsi que la voie de référence du brouilleur connu. Ceci s'obtient sur une carte au moyen d'un bloqueur, et entre cartes en privilégiant une des cartes qui jouera le rôle de maître et pilotera l'horloge d'échantillonnage, les autres cartes recevant leur signal d'horloge de la première. Les cartes esclaves et la carte maître échangeront également plusieurs signaux gérés par des registres d'état, pour l'initialisation, le démarrage et l'arrêt synchrone.

Dans le prototype d'antenne acoustique pour l'audio conférence, ces cartes fonctionneront sur 4 voies avec une cadence d'échantillonnage de 16 kHz, ou de 8kHz, mais les cartes sont également capables, par commutation de cavaliers, de supporter des cadences de 2, 4 et 32 kHz, sur 1 à 4 voies.

La partie analogique comporte 15 microphones de type électret, avec leurs préamplificateurs, et les filtres passe-bas anti-repliement. La seizième voie est destinée à un signal de référence, qui constitue un brouilleur connu: par exemple, ce sera le signal provenant de l'autre salle dans le contexte de l'audioconférence, ou l'autoradio dans le cas du téléphone de voiture. Cette voie ne comporte que le filtre anti-repliement, sans les préamplificateurs. Du point de vue mécanique, les quinze microphones sont répartis sur un demi-cercle de 1 mètre de diamètre.

## LOGICIEL.

Les développements du logiciel nécessaire pour cette antenne acoustique se font dans le langage OCCAM qui est bien adapté à la programmation des Transputers, le langage et le matériel ayant été conçus l'un en fonction de l'autre. Il s'agit d'un langage structuré, qui ne possède que des types de données scalaires, mais permet cependant des structures complexes, pouvant comprendre des variantes. Ceci se fait par le biais des protocoles associés à chaque canal. Les trois processus élémentaires du langage sont l'affectation, la lecture sur un canal et l'écriture sur un canal. Le canal OCCAM est donc l'élément essentiel du langage, il est le pendant du lien bit-série du Transputer.

L'intérêt du langage tient surtout en la possibilité de mettre en séquence ou en parallèle les processus. Le modèle du parallélisme est simple et efficace. Il repose sur la communication entre processus par des canaux, chaque échange de données donnant lieu à un rendez-vous. Des processus peuvent s'exécuter en parallèle, soit sur un même processeur, soit sur des processeurs différents. Le langage permet de le spécifier en configurant un réseau, et en plaçant les processus devant s'exécuter sur chaque Transputer.

Nous allons donner ici un certain nombre d'indications issues de nos premières expériences autour de la structure temporelle de la formation de voie adaptative dite "Generalised Sidelobe Canceller". Rappelons que la structure comporte une formation de voie "au repos", un certain nombre de formations de voies fixes orthogonales à la voie "au repos", puis une annulation adaptative du bruit présent dans la voie "au repos", en prenant comme références les voies dites "orthogonales". Réaliser cette structure conduit donc à une sorte de pipe-line hétérogène. Chaque étage de ce pipe-line sera implanté sur plusieurs processeurs en parallèle si le nombre d'opérations à effectuer l'exige. Pour simplifier la conception du logiciel, il est souhaitable d'avoir peu de processus différents qui seront dupliqués le plus possible, avec une structure de réseau très régulière.

Avant de parvenir à cette conclusion, nous avons tenté d'optimiser le placement des processus sur les différents processeurs, en confiant à un programme le soin de déterminer la configuration [7]. Les données communiquées à ce programme étaient les temps d'exécution de chaque bloc élémentaire de calcul, déterminés en simulant chaque bloc sur un T800, puis les chemins de données entre ces blocs élémentaires. Le programme cherchait alors de manière combinatoire la configuration qui permettait à chaque Transputer de travailler un temps maximal dans la durée de chaque cycle (62.5  $\mu$ s entre deux échantillons), tout en respectant le nombre total de 4 liens avec ses voisins.

Les résultats de cette optimisation ont été assez décevants. Ceci peut être imputé à un manque de sophistication du programme de placement automatique, mais la raison nous en semble plus profonde. La structure trouvée par le programme n'était pas dépourvue d'attraits, elle correspondait bien à un certain type d'optimum, et elle était relativement économe en Transputers. Son échec est venu de la difficulté à prendre en compte les diverses synchronisations



entre les traitements élémentaires, certains rendez-vous entraînant pour l'un des processus un temps d'attente trop grand. Nous avons pu parvenir à une seconde solution, donnant de meilleures performances en imposant une configuration plus régulière (un tableau de 6 colonnes de 5 Transputers), et en optimisant empiriquement, mais soigneusement les synchronisations entre processus. Nous avons en particulier observé que dans le cas d'une structure de calcul très régulière comme en filtrage adaptatif, où la même séquence d'instructions se répète à chaque cycle, il n'est pas très rentable de rendre les échanges de données asynchrones lorsqu'ils ne concernent que quelques octets.

Les nécessités de l'implantation en parallèle peuvent dicter certains choix d'algorithmes, ainsi le filtrage adaptatif présent dans la version temporelle de la formation de voie pose quelques problèmes si on le réalise par l'algorithme le plus simple, de gradient stochastique. Soient  $A_t$  le modèle de filtre à l'instant  $t$ ,  $X_t$  le vecteur des échantillons du passé du signal référence,  $y_t$  le signal à débruiter, l'algorithme s'écrit :

$$\varepsilon_t = y_t - A_{t-1}^T X_t$$

$$A_t = A_{t-1} + \mu X_t \varepsilon_t$$

Dans cet algorithme, il est possible de découper en tranches chacun des vecteurs  $A_t$  et  $X_t$ , pour répartir les tranches sur différents processeurs. Mais la variable  $\varepsilon_t$  doit être commune à tous les processus qui implantent l'algorithme en parallèle. Une telle variable n'est pas permise par OCCAM, aussi chaque processus doit-il calculer la contribution locale à cette variable, recevoir la somme des contributions venant d'un côté (gauche/droite), renvoyer de l'autre côté cette somme augmentée de la contribution locale, et utiliser pour  $\varepsilon_t$  la somme de ces trois contributions, locale, gauche et droite. Aux deux extrémités du pipe-line, le traitement est différent, car l'extrémité gauche doit aussi additionner  $y_t$ , tandis que l'extrémité droite initialise à zéro la contribution. Le squelette d'un processus OCCAM (en milieu de pipe-line) réalisant ceci est présenté ci-dessous.

```

VAL INT n IS 12 :
[n+1]REAL x, a :
REAL s.left, s, s.right, epsilon :
CHAN OF ANY in.left, out.left,
              in.right, out.right :
WHILE running
SEQ
  out.right ! x[n]
  ... shift X.
  in.left ? y, x[0]
PAR
  in.left ? s.left
  ... compute partial sum (s := A.X).
  in.right ? s.right
PAR
  out.left ! s.right + s
SEQ
  epsilon := (s.left + s) + s.right;
  ... adapt A (A := A + mu*epsilon*X)
  out.right ! s.left + sum

```

Cette solution n'est pas la plus efficace qui soit, pour une application avec des contraintes de temps réel, car les processeurs du (pseudo) pipe-line doivent en fait tous

s'attendre après le calcul des sommes partielles. Une solution à ce problème est de passer à un algorithme en treillis qui est plus coûteux, mais permet d'avoir un véritable pipe-line, sans la remontée de droite à gauche que l'on a ci-dessus.

## CONCLUSION.

Nous avons présenté quelques uns des problèmes qui se sont posés à nous dans l'élaboration d'un prototype d'antenne acoustique devant fonctionner en temps réel. Dans le contexte d'une structure parallèle, il nous apparaît que l'étude et le choix de cette structure doit se faire simultanément au niveau de l'algorithme, du matériel et du logiciel, comme en témoignent les remarques encore partielles que nous déployons ci-dessus.

## REMERCIEMENTS.

Nos remerciements vont à plusieurs de nos collègues qui ont contribué avec beaucoup d'efficacité à la réalisation du prototype d'antenne acoustique: M.Chotard pour la partie mécanique, A.Chaigne pour le calibrage des microphones, M.Krueger pour la conception des convertisseurs, P.Le Borgne pour la programmation de l'algorithme adaptatif.

## REFERENCES.

- [1] A.GILLOIRE, J.P.JULLIEN, "L'acoustique des salles dans les télécommunications", *Echo des Recherches*, n°127, pp 43-54, 1987.
- [2] C.GAGNOULET, D.JOUVET, "Développements récents en reconnaissance de la parole", à paraître dans *l'Echo des Recherches*, 1989.
- [3] R.A.MONZINGO, T.W.MILLER, "Introduction to adaptive arrays", J.Wiley & sons, 1980.
- [4] M.XU, Y.GRENIER, "Débruitage par antenne acoustique", *Colloque GRETSI sur le traitement du signal et des images*, 1989.
- [5] O.L.FROST, "An algorithm for linearly constrained adaptive array processing", *Proc. of IEEE*, Vol.60, n°8, pp , 1972.
- [6] L.J.GRIFFITHS, JIM, "An alternative approach to linearly constrained adaptive beamforming", *IEEE Trans. on AP*, Vol.30, n°, 1982.
- [7] P.LE BORGNE, "Formation de voies sur réseau de Trtransputers", stage ENST, 1989.