Capteur sonore intelligent pour la réduction de la consommation d'énergie

Dan Istrate, Vincent Zalc, Theophile Cocquerez, Sami Boutamine

▶ **To cite this version:**

Dan Istrate, Vincent Zalc, Theophile Cocquerez, Sami Boutamine. Capteur sonore intelligent pour la réduction de la consommation d'énergie. Colloque InterUT Systèmes sûrs et durables, Université de Technologie de Compiègne [UTC], Feb 2023, Paris, France. 10.34746/ax2q-b976 . hal-04011806

HAL Id: hal-04011806
https://hal.science/hal-04011806

Submitted on 2 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Capteur sonore intelligent pour la réduction de la consommation d'énergie

Dan Istrate[1], Vincent Zalc[1], Theophile Cocquerez[1], Sami Boutamine[1]
[1]Université de technologie de Compiègne, CNRS, Biomécanique et Bioingénierie, Centre de recherche Royallieu - CS 60 319 - 60 203 Compiègne Cedex
dan.istrate@utc.fr

Résumé - Les nouvelles technologies au service des espaces intelligents ne cessent de se développer et sont en interaction avec les objets de la vie courante, leur émergence s'articule dans plusieurs domaines d'application tels que la réduction énergétique, la santé connectée et la télémédecine. Le développement d'un capteur de présence capable de détecter le nombre de personnes qui parlent dans une pièce est présenté. Le but est de contribuer à la réduction et l'automatisation de la consommation énergétique (chauffage, climatisation, éclairage, autres appareils ménagers…) en l'adaptant au nombre de personnes présentes dans une salle pour offrir plus de confort. Le capteur a été développé sur une carte embarquée (Raspberry PI) et fonctionne en temps réel. Ces travaux se sont déroulés dans le cadre du projet FUI CoCAPs.

Mots-clés : segmentation en locuteurs, detection de presence, reconnaissance des sons, temps réel

I. INTRODUCTION

Les nouvelles technologies ne cessent de se développer de façon exponentielle, mises au service du bien-être des personnes et plus particulièrement les plus fragilisés, elles trouvent alors pleinement leur sens.

Les capteurs connectés (ou intelligents) font partie de ces nouvelles technologies, leur utilisation a permis l'émergence de la santé connectée et de la télémédecine, un domaine où les capteurs sont devenus une aide indispensable à la vie de tous les jours. En effet, la santé connectée est en voie de devenir une préoccupation majeure dans nos sociétés technicisées et un champ de recherche exploré par de nombreuses disciplines en raison de la complexité évidente des dispositifs mobilisés et des enjeux suscités (technologiques, sociaux, économiques, éthiques et politiques, etc.).

La facture énergétique augmente en France pour atteindre 45,1 milliards d'euros en 2018, après quatre années consécutives de baisse continue entre 2012 et 2016 (-56%), pénalisée par la remontée des cours, elle connait une hausse en 2017 (+23%). Face à la hausse des prix des énergies fossiles, la facture énergétique s'accroit de 16% en 2018, malgré la baisse des quantités achetées.

Afin de montrer l'intérêt de l'utilisation d'un capteur de présence pour réduire la consommation énergétique, on s'est basé sur une approche comparative entre deux bâtiments de l'Université de Technologie de Compiègne. Les deux sites ont approximativement la même surface et la même occupation, l'un étant équipé par des capteurs infrarouge pour piloter le système de chauffage et d'éclairage, tandis que l'autre ne bénéficie pas de cette automatisation.

Dans cette approche comparative, deux types de consommation ont été choisis, la consommation de chauffage et la consommation électrique. Le tableau 1.1 présente la consommation de chauffage entre les années 2014 et 2019 inclus. Les valeurs représentent la consommation par m2 sur l'année en Mégawatt.heure (MWh).

	2014	**2015**	**2016**	**2017**	**2018**	**2019**
Bâtiment 1 (sans capteurs)	0,08	0,09	0,08	0,09	0,09	0,09
Bâtiment 2 (avec capteurs)	0,05	0,05	0,04	0,05	0,05	0,04

Tableau 1.1. Consommation de chauffage par m2 (MWh).

Dans le bâtiment équipé par des capteurs de mouvement, le chauffage s'allume automatiquement dès qu'un mouvement est détecté dans une salle, et il s'arrête si aucune activité n'est détectée pendant 30 minutes, alors que le chauffage fonctionne de façon continue dans le bâtiment non équipé de capteurs. Les chiffres présentés montrent la différence de consommation de chauffage entre les deux bâtiments. Celui équipé de capteurs consomme moitié moins par rapport au bâtiment avec une consommation continue.

La consommation d'électricité pour la même période (2014 – 2019), est illustrée par les chiffres présentés dans le tableau 1.2, les valeurs représentent la consommation par m2 pour l'année en Kilowatt.heure (kWh).

	2014	2015	2016	2017	2018	2019
Bâtiment 1 (sans capteurs)	32,83	58,02	63,93	55,76	58,16	60,45
Bâtiment 2 (avec capteurs)	28,79	45,19	51,01	45,73	54,39	54,03

Tableau 1.2. Consommation électrique par m2 (kWh)

Dans le bâtiment doté de capteurs, l'automatisation est utilisée uniquement pour la lumière, ce qui explique que la différence entre les deux soit la plus faible. La lumière s'allume quand le capteur détecte un mouvement et s'éteint lorsqu'aucune activité n'est détectée pendant 15 minutes.

Pour une surface moyenne de 4000 m2, on peut déduire que la consommation de chauffage pour un système piloté par capteurs pour une période d'une année, est de 184 MWh alors que celle du système non piloté est de 348 MWh ; le gain est de 47,1%. Tandis que la consommation électrique est respectivement de 186,08 MWh et de 219,36 MWh ; le gain est dans ce cas de 15,2%.

En effet, dans une salle de réunions, concernant le chauffage, le capteur de présence permettra de gérer au mieux la consommation. Il pourra, par exemple, réduire la consigne de température si le capteur ne détecte pas de parole, au lieu de conserver la température car le capteur de mouvement implique une temporisation. Dans une salle de réunions les participants sont moins amenés à bouger qu'à dialoguer. De plus, l'utilisation du capteur de présence permettra d'estimer le nombre de personnes présentes dans une salle, notamment lors de successions de réunions avec un nombre d'occupants variable, et occupés à chaque fois d'un nombre différents de personnes, ce qui revient à faire un apprentissage de ce nombre de personnes et d'adapter la température adéquate à leur confort et de réduire ainsi le coût d'énergie.

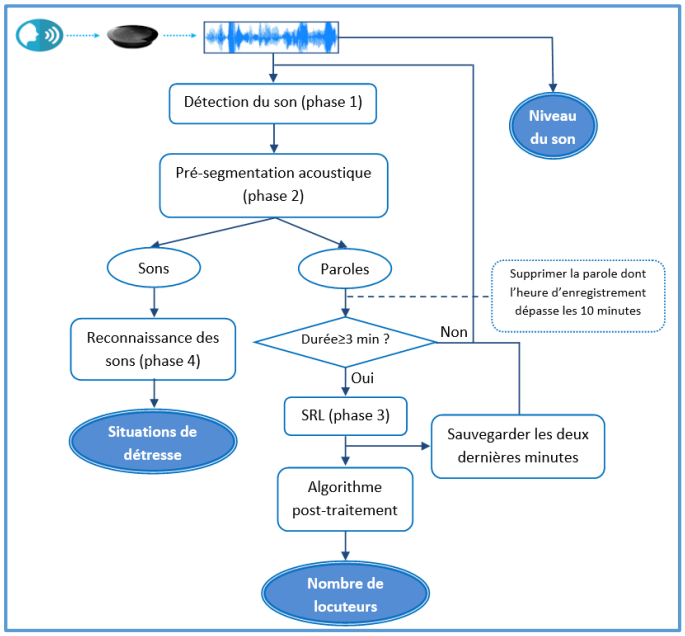
II. DETECTEUR DU NOMBRE DE PERSONNES QUI PARLENT

A. Système proposé

Le système propose d'analyse de l'environnement sonore a le schéma présenté dans la figure 1.

La carte Raspberry Pi 4 Model B a été choisie comme plateforme de développement. C'est une carte électronique qui est équipée d'un processeur composé de 4 cœurs de 64 bits, de type BCM2711 ARM Cortex-A72, à une cadence de 1,5 GHz avec une mémoire de 8 GB. Le système d'exploitation Raspbian (Linux – Debian) est utilisé pour le développement et l'implémentation des algorithmes.

Figure 1. Architecture du système proposé « eBioMed_SpkCOUNT ».



Le choix pour le microphone s'est porté sur le Jabra Speak 510 car il a donné les meilleurs résultats dans les tests effectués et avec une performance stable pour les différentes distances. Sa sensibilité variable automatiquement est un atout pour ce microphone.

eBioMed_SpkCount se décompose en quatre principales phases, la première phase consiste à acquérir le signal audio et enregistrer les sons quand un événement sonore est détecté, la deuxième consiste à séparer la parole de tout autres types de sons, cette dernière est utilisée pour la segmentation et le regroupement en locuteurs (phase 3). Les autres sons (différents de la parole) seront exploités par la reconnaissance des sons (phase 4). Le système fonctionne de façon continue, les quatre phases s'exécutent en parallèle, en se basant sur les avantages de la programmation parallèle et concurrentielle.

Afin d'assurer le traitement en ligne de l'algorithme de détection du nombre de locuteurs (SRL), une fenêtre glissante a été utilisée dont la taille est de 180 secondes et le pas est de 60 secondes (recouvrement des deux dernières minutes "120 secondes"), l'algorithme d'identification du nombre de locuteur s'exécute quand un signal audio d'une durée de 3 minutes

contenant uniquement de la parole est disponible. L'algorithme implémenté (Figure 2) fonctionne selon les étapes suivantes :

- t0 : attente de l'enregistrement de 3 minutes de la parole.
- t1 : 3 minutes de la parole sont enregistrées, l'algorithme de détection du nombre de locuteurs est lancé.
- t2 : une nouvelle minute de la parole est enregistrée, en lui rajoutant aussi les 2 dernières minutes enregistrées en t1 (recouvrement des deux dernières minutes) pour atteindre une durée de 3 minutes et lancer l'algorithme d'identification du nombre de locuteurs une nouvelle fois.
- La parole enregistrée a une durée de validité de 10 minutes ; on supprime tous les segments de parole plus anciens de 10 minutes.

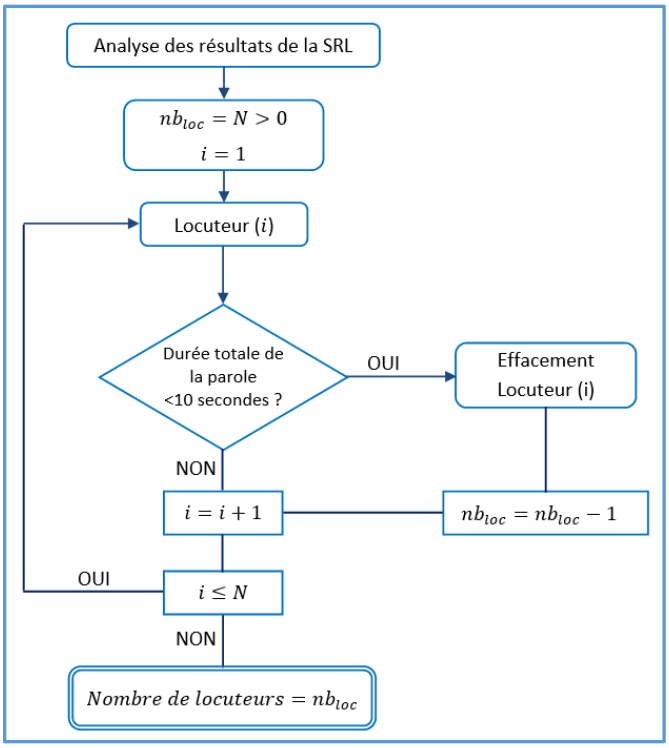


Figure 2- Algorithme de post-traitement du nombre de locuteurs

Après plusieurs tests du système, on a constaté que les locuteurs détectés dont la durée maximale de parole est inférieure à 10 secondes, sont en grande majorité (81,6%) de faux locuteurs. Les segments de ces derniers correspondent essentiellement à la parole superposée, de la parole non superposée mais accompagnée de sonnerie de téléphone et du changement d'intonation et/ou du niveau d'un locuteur. Pour cela, un algorithme de post-traitement a été appliqué sur les résultats obtenus de la SRL, en supprimant toute personne détectée dont la durée de parole est inférieure à 10 secondes.

Le nombre de locuteurs sert à automatiser et à réduire la consommation de l'éclairage dans un appartement, mais sert aussi au maintien des personnes âgées qui vivent seules dans une pièce. En analysant le nombre de personnes détectées nous pouvons en déduire si la personne est sortie de son appartement, si elle est seule ou accompagnée (passage des différents services ou de la famille).

L'algorithme de reconnaissance des sons s'exécute en parallèle avec les autres algorithmes, chaque son différent de la parole sera traité par le système de reconnaissance des sons de la vie courante. Une alerte pourra être générée si une classe de son de détresse est détectée telles que les cris, les pleurs, le bris de verre, la chute d'un objet… La détection des situations de détresse se base sur la reconnaissance de ce type des sons, mais aussi sur une estimation des activités de la personne qui amène à l'apprentissage des habitudes ; un changement d'habitude pourrait indiquer l'apparition d'un évènement.

eBiomed_SpkCOUNT calcule aussi le niveau sonore en fournissant la valeur moyenne et maximale pour chaque minute, cette information pouvant servir dans certains cas à détecter une activité sonore dans un espace.

III. EVALUATION DU SYSTEME

L'objectif du système *eBioMed_SpkCOUNT* est la détection du nombre de locuteurs à travers la méthode de segmentation et regroupement en locuteurs, et la détection des situations de détresse et d'activités en se basant sur la méthode de reconnaissance des sons.

Plusieurs tests ont été effectués au cours du développement sur différents enregistrements en conditions contrôlées et réelles. Cette évaluation est basée sur le taux d'identifications correctes (TIC), pour le système de comptage du nombre de locuteurs et de la reconnaissance des sons.

Le Tableau 3 présente les résultats du système de segmentation en locuteurs en conditions contrôlés sur des enregistrements.

	3 minutes	**5 minutes**	**10 minutes**
1 locuteur	100%	100%	100%
2 locuteurs	80%	80%	90%
3 locuteurs	90%	80%	90%
Total	**90%**	86,7%	**93,3%**
Performance	**89,99%**		

Tableau 3. Taux des tests correctement identifiés « nombre de locuteurs ».(Signal filtré contenant uniquement de la parole).(20MFCC+ énergie, off-line)

Après avoir obtenus des résultats encourageants lors de la première phase de test en mode off-line, nous nous sommes proposés d'évaluer le système complet en intégrant l'algorithme d'identification du nombre de locuteurs adapté à notre cas d'utilisation. Cette évaluation a été réalisée en mode en ligne et en conditions réelles (en présence de bruit et des personnes qui se coupent la parole). Le système global a été implémenté sur la carte Raspberry Pi 4 Model B, et les résultats sont fournis en temps réel.

La première évaluation en mode on-line a été réalisée pour des durées de 3 minutes. Les résultats sont consignés dans le tableau 4 en termes de taux d'identifications correctes (TIC) et de taux d'identifications correctes amélioré (TICa).

Parmi les objectifs du système est d'automatiser et réduire la consommation de l'énergie tels que l'éclairage, le chauffage et la climatisation, mais aussi assurer un confort pour les utilisateurs. En conséquence une erreur d'un (-/+ 1) locuteur ou même de (-/+ 2) locuteurs est insignifiante et peut être négligée.

20MFCC+Energie	**Sans erreur**	**-/+ [0,1] Locuteur**	**-/+ [0, 1, 2] Locuteurs**
TIC	51,4%	77,1%	97,1%
TIC$_a$	**71,4%**	**85,7%**	**100%**

Tableau 4. TIC et TICa (nombre de locuteurs) on-line

Une deuxième évaluation on-line *d'eBioMed_SpkCOUNT* a été réalisée dans un des bureaux du partenaire industriel du projet CoCAPs. Le test s'est déroulé pendant les heures de travail et pendant les réunions, mais aussi pendant la nuit pour tester si le système détecte la présence des personnes quand le bureau est inoccupé.

L'évaluation a été réalisée en conditions réelles telles que la présence du bruit, des personnes se coupant la parole et la visite des collègues dans le bureau. Les résultats sont présentés dans le tableau 5, ils représentent le taux d'identifications correctes amélioré (TICa).

20MFCC+Energie	**Sans erreur**	**-/+ [0,1] Locuteur**	**-/+ [0, 1, 2] Locuteurs**
TIC$_a$	58,3%	75%	83,3%

Tableau 5 TICa (nombre de locuteurs) on-line

Les résultats obtenus montrent une diminution du TICa par rapport à celui du premier test, cela étant dû aux conditions d'enregistrement. Le positionnement en hauteur du microphone Jabra 510 Speak a amené une réduction de la qualité du signal. On peut signaler aussi la présence du bruit et de la parole superposée plus souvent par rapport au premier test (sonnerie de téléphone est superposée à la voix d'un locuteur pendant 30s).

IV. Conclusion and Perspectives

Consommer de l'énergie est une nécessité, la technologie permet d'en produire en grande quantité, utilisant toutes les ressources possibles (fossiles, eau, vent, soleil). Cependant, les coûts de l'énergie (carburant, électricité, gaz) augmentent inexorablement dans tous les pays y compris en France dont la facture énergétique a augmenté pour atteindre 45,1 milliards d'euros en 2018. La consommation d'énergie a donc des frais et des impacts notamment au niveau mondial, la réduire est un enjeu financier et environnemental. Plusieurs mesures de maitrise de l'énergie sont possibles pour y parvenir, telles que l'adaptation, l'automatisation et la réduction de la consommation de l'éclairage, de la ventilation, et du chauffage en fonction de nos besoins.

La prise en charge des personnes âgées est un enjeu sociétal majeur, surtout quand selon une étude IFOP (Institut Français d'Opinion Publique) 83% de ces personnes souhaitent vieillir à domicile plutôt que d'intégrer une maison de retraite. Il est donc important de leur garantir le confort et notamment la sécurité pour éviter les accidents domestiques fréquents tels que les chutes, car des milliers de décès de séniors par an à cause d'une chute grave sont comptés. Il est indispensable de mettre en place des solutions technologiques (alarmes, détecteurs, caméras, capteurs) afin de favoriser le maintien à domicile des séniors pour permettre de gérer au mieux leur perte d'autonomie, et détecter leur situation de détresse.

Le capteur proposé s'inscrit dans ces deux contextes. Le but est d'une part de contribuer à la réduction et l'automatisation de la consommation énergétique (chauffage, climatisation, éclairage…) en l'adaptant au nombre de personnes présentes dans une salle pour offrir plus de confort, et d'autre part pour offrir une assistance aux personnes âgées à travers l'analyse automatique des sons de la vie courante, à vieillir à leur domicile, en bonne santé en confort et sécurité. Pour ce faire, deux tâches dans le traitement de la parole sont utilisées à savoir la segmentation et le regroupement en locuteurs (SRL) et la reconnaissance des sons.

Il existe à l'heure actuelle, des capteurs de mouvement qui font partie intégrante du monde qui nous entoure, mais il n'existe pas sur le marché de capteurs de présence qui permettent d'estimer le nombre de personnes présentes dans une pièce et de détecter leur situation de détresse.

L'utilisation de capteurs de présence augmentera non seulement la sécurité et le maintien des personnes âgées à domicile mais aussi leur confort, la réduction et l'automatisation de la consommation énergétique à travers le contrôle de l'éclairage, de la climatisation et de nombreux autres appareils ménagers.

Les travaux présentés ont plusieurs contributions : développement d'un capteur de présence sur un système embarqué avec une puissance de calcul limité, choix du

microphone et son positionnement, l'adaptation et la combinaison des algorithmes existants, assurer le fonctionnement continu, en ligne et en temps réel du système. Ainsi que fournir des réponses pour des courtes durées.

La plupart des systèmes existants sur le marché, en relation avec le traitement de la parole, fonctionnent en mode hors ligne (off-line). Ces systèmes traitent un signal audio préalablement enregistré, alors que celui que nous proposons a la particularité de non seulement fonctionner en ligne (on-line), en temps réel sur un système embarqué tout en traitant les données au fur et à mesure de leur enregistrement, mais aussi de regrouper tous les traitements tels que l'enregistrement du signal, la segmentation et regroupement en locuteurs, la reconnaissance automatique des sons, pour fonctionner ensemble, en parallèle et de façon continue. Un algorithme a été développé pour assurer aussi le fonctionnement en ligne (on-line).

La plupart des algorithmes fonctionnent initialement en mode off-line sur des durées de paroles grandes. Un algorithme a été développé et adapté à notre système pour fournir une réponse chaque minute en temps réel, sur peu de données de parole (fenêtre glissante de 3 minutes) et cela avec un recouvrement de deux dernières minutes, pour une segmentation optimisée en locuteurs en mode continu.

Le capteur intelligent de présence a été testé dans différents environnements réels (salles de réunion, bureaux, visioconférences ...), en temps réel surtout pendant des visioconférences à cause du COVID-19 et sous diverses conditions réelles (paroles, sons de la vie courante) avec l'utilisation de 18 classes de sons de la vie quotidienne où il a obtenu des résultats encourageants. L'évaluation de ce système ayant été présentée en termes de taux d'identifications correctes (TIC).

L'évaluation *d'eBioMed_SpkCOUNT* a été faite en conditions réelles « off-line/on-line ». Les taux d'identifications correctes (TIC) obtenus pour l'évaluation de notre système pour le comptage de locuteurs semblent encourageants, cependant, nous n'avons pas pu faire plusieurs comparaisons en l'absence de systèmes de comptage évalués sur de la parole non superposée.

Les résultats obtenus par le capteur sont très satisfaisants et encourageants pour la suite, néanmoins ils peuvent être améliorés en tenant compte des points suivants :

- Adapter l'emplacement du microphone de façon à obtenir des résultats optimaux pour une meilleure acquisition sonore. Développer une méthodologie pour trouver l'emplacement optimal de microphone.
- Résoudre le problème de la variabilité intra-locuteur (état émotionnel, voix chantée, voix parlée, voix chuchotée…) dans l'identification d'un locuteur donné, l'évolution naturelle ou volontaire de la voix d'une personne, en se basant sur d'autres paramètres acoustiques utilisés seuls ou couplés.
- Intégrer dans notre capteur un algorithme de reconnaissance vocale tels que Google Home ou Amazon Alexa, afin d'améliorer et de renforcer le maintien à domicile des personnes âgées et faciliter leur quotidien. Cet algorithme sera capable de détecter les chutes d'objets, les malaises des personnes âgées et les appels de détresse. Dans ce même contexte, apporter une amélioration des outils et systèmes présents pour les rendre plus adaptés aux personnes âgées, afin de détecter les troubles de paroles des séniors (vieillissement du conduit vocal, bruit de bouche), celle-ci étant moins bien reconnue qu'une voix plus jeune par les systèmes de reconnaissance vocaux en général.
- L'adaptation du système de reconnaissance des sons, dans le cadre de la pandémie COVID-19, pour des fins diverses.

REMERCIEMENTS

Ces travaux ont eu lieu dans le cadre du projet FUI COCAPS; nous remercions BPI France pour le financement et les partenaires du projet pour la collaboration.

BIBLIOGRAPHIE

[1] C. Barras, Xuan Zhu, S. Meignier and J. -. Gauvain. Multistage speaker diarization of broadcast news, in IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, no. 5, pp. 1505–1512, September 2006.

[2] S. Meignier, J.F. Bonastre, C. Fredouille, T. Merlin. Modèle de Markov évolutif pour les tâches de suivi de locuteurs. JEP2000, Aussois, France. pp. 4, 2000.

[3] S. Meignier et T. Merlin. LIUM SpkDiarization: an open-source toolkit for diarization. In CMU SPUD Workshop, Dallas, Texas (USA), pp. 1–6, 2010.