

Analyse sonore et multimodale dans le domaine de l'assistance à domicile

Michel VACHER

Soutenance de DHDR

UNIVERSITÉ DE GRENOBLE



Contexte général de mon travail

Analyse sonore et multimodale

Michel Vacher

Introduction

I - Habitat Intelligent

Contexte de l'Habitat Intelligent
Verrous scientifiques

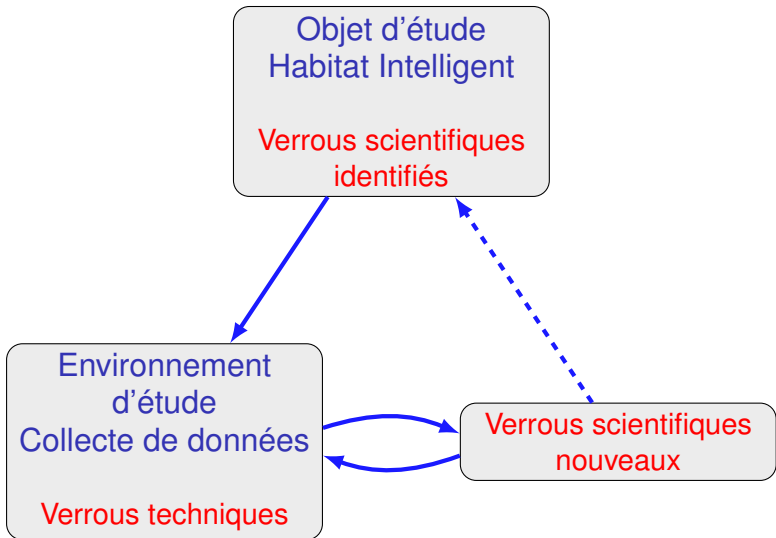
II - Construction de plateformes d'expérimentation

HIS
DOMUS
Expérience pilote

III - Trois des thèmes abordés

Analyse des sons
Reconnaissance de la parole distante
Suivi d'activité

Conclusion et perspectives



I - Contexte, contraintes et attentes dans un habitat intelligent

Contexte de l'Habitat Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote : mise en évidence des questions éthiques

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la parole distante

Suivi d'activité

Conclusion et perspectives

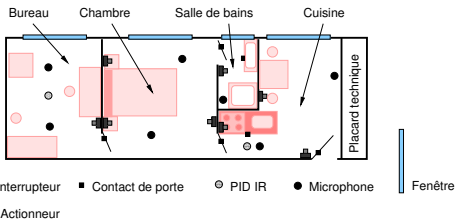
Vieillesse de la population

- ▶ Constat : espérance de vie en hausse
($\approx 22\%$ de plus de 65 ans en 2050)
 - ⇒ proportion de personnes âgées augmentée :
surcharge des institutions spécialisées
- mais ⇒ augmentation de la qualité de vie

- ▶ Conséquences du vieillissement
 - isolement grandissant
 - maladies chroniques et dégénératives
(Alzheimer)
 - diminution de l'autonomie

**maisons intelligentes : un enjeu sociétal
- maintien à domicile -**

Une vision à long terme : un habitat avec des capteurs permettant le maintien à domicile



Une vision à long terme : un habitat avec des capteurs permettant le maintien à domicile



**Adaptabilité en fonction
de l'évolution et des besoins**

Introduction

I - Habitat
Intelligent

Contexte de l'Habitat
Intelligent

Verrous scientifiques

II - Construction de
plateformes
d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des
thèmes abordés

Analyse des sons

Reconnaissance de la
parole distante

Suivi d'activité

Conclusion et
perspectives

Un cahier des charges ambitieux

- ▶ **santé**
capteurs physiologiques, capteurs portés par la personne ...
- ▶ **sécurité**
capteurs portés par la personne, capteurs de chute ...
- ▶ **autonomie/confort**
compensation du handicap en facilitant l'accès aux appareils domestiques
- ▶ *communication avec l'entourage et l'extérieur*
(souvent oubliée)

questionnement éthique

La grande variabilité des désirs, des besoins et des contraintes des personnes conditionne leur acceptation de cet environnement « intelligent ».

Domaine émergent, déjà de nombreux projets ...

- ▶ Europe (FP7 AAL), ANR (TECSAN, VERSO), mosaïque d'industriels. ...
- ▶ se veulent réalistes, mais rarement confrontés à la réalité
- ▶ prédominance des solutions utilisant des caméras
- ▶ e-lío (TECHNOSENS) : SWEET-HOME, CIRDO

**Informations acoustiques
peu, pas ou mal prises en compte !**

Originalité de l'analyse sonore

Les événements sonores, témoins et acteurs de l'Habitat Intelligent

► La parole

- Alarmes
- Commandes
- Informations (rappels)
- Communication avec l'extérieur

► L'environnement sonore

- Bruits d'eau
- Chocs
- Frottements
- Sonneries
- Moteurs

...

I - Contexte, contraintes et attentes dans un habitat intelligent

Contexte de l'Habitat Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote : mise en évidence des questions éthiques

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la parole distante

Suivi d'activité

Conclusion et perspectives

Verrou 1 : contraintes nouvelles pour l'analyse sonore

Contexte spécifique de l'Habitat Intelligent

- **signaux sporadiques** : difficulté à déterminer le début et la fin d'un événement sonore au milieu du pseudo-silence
- **signaux hétérogènes** : peu de parole, beaucoup de sons (utiles) et de bruits (gênants)
- **signaux mélangés** : parole en présence de bruit
- **signaux à la limite de la parole** : limite entre parole et sons humains, intérêt de ces informations

**Analyse Sonore
en parallèle de la
Reconnaissance de la Parole**

Verrou 2 : reconnaissance de la parole en conditions distantes

Introduction

I - Habitat Intelligent

Contexte de l'Habitat Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des thèmes abordés

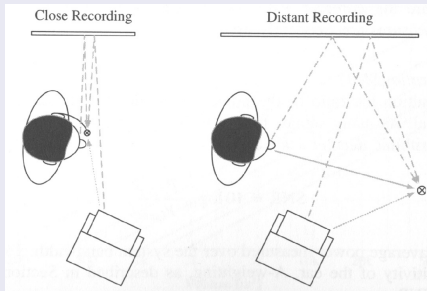
Analyse des sons

Reconnaissance de la parole distante

Suivi d'activité

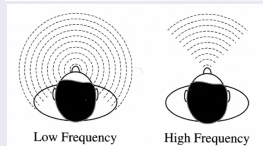
Conclusion et perspectives

Domaine de recherche récent et très vaste : WÖLFEL, McDONOUGH (2009)



Condition proche ou distante

pour des capteur similaires, omnidirectionnels



Cône d'émission basses ou hautes fréquences

Verrou 2 : reconnaissance de la parole en conditions distantes

Applications courantes

- ▶ Bureau ou salle de réunion (réseaux de microphones ou *Microphone Array*)
- ▶ Robots humanoïdes
- ▶ Commandes vocales dans l'habitacle automobile
- ▶ Traduction *Speech to speech*

Contraintes particulières

- ▶ bruit de fond important (RSB - Rapport signal sur bruit)
 - atténuation du signal utile avec la distance
 - microphones omnidirectionnels
 - effet Lombard
- ▶ distorsion due à l'écho et à la réverbération
- ▶ orientation du locuteur par rapport au microphone

Verrou 3 : spécificité de la parole des personnes âgées

Voix âgée

- ▶ Thème encore peu étudié
- ▶ Étude de **VIPPERLA (2009, 2010)** pour la langue anglaise
 - modification de la fréquence fondamentale
 - manque d'énergie
 - instabilité de production de certaines consonnes
 - augmentation du bruit

Langage des personnes âgées

- ▶ Devient un langage spécifique (changement d'usages)
- ▶ Substrats langue étrangère/retour à la langue de l'enfance

Verrou 4 : localisation et suivi des activités

Intérêt

- ▶ localisation = information de contexte pour la domotique
- ▶ activité
 - information de contexte pour la domotique
 - indice du niveau d'autonomie

Contraintes dues aux capteurs

- ▶ capteurs portés par la personne \implies contraintes
- ▶ caméra vidéo \implies intrusive et bruitée
- ▶ microphone \implies importance du bruit, ambiguïté
- ▶ capteurs domotiques classiques \implies informations indirectes

Utiliser la complémentarité des différentes modalités

- ▶ par exemple : fusion de données pour inférer la localisation

I - Contexte, contraintes et attentes dans un habitat intelligent

Contexte de l'Habitat Intelligent
Verrous scientifiques

II - Construction de plateformes d'expérimentation

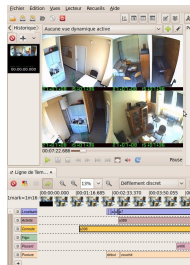
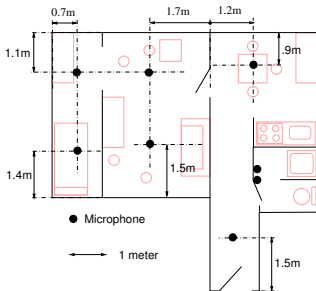
HIS
DOMUS
Expérience pilote : mise en évidence des questions éthiques

III - Trois des thèmes abordés

Analyse des sons
Reconnaissance de la parole distante
Suivi d'activité

Conclusion et perspectives

Habitat Intelligent pour la Santé (HIS)

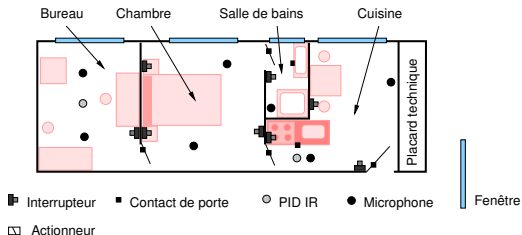


- ▶ **HIS** : appartement expérimental de 47m²
 - Laboratoire TIMC-IMAG (2000)
 - Bâtiment Jean Roget - Faculté de Médecine de Grenoble
- ▶ Projets ★HIS
- ▶ Actimètre, PIR, microphones
- ▶ **Limitations : absence de domotique ?**

Équipement sonore (RESIDE-HIS, DESDHIS, thèse A. Fleury)

IEEE HEALTHCOM, M. VACHER, F. PORTET, A. FLEURY, N. NOURY (2010)

Appartement DOMUS



- ▶ **DOMUS** : appartement expérimental de 30m²
 - Laboratoire LIG/MULTICOM (2009)
 - CTL - Institut Carnot LSI
- ▶ Capteurs et actionneurs domotiques, microphones
- ▶ **Les + :**
 - domotique
 - caméras synchronisées

SWEET-HOME :

- Cahier des charges
- Équipement sonore
- Caméras pour identification des événements sonores

I - Contexte, contraintes et attentes dans un habitat intelligent

Contexte de l'Habitat Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote : mise en évidence des questions éthiques

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la parole distante

Suivi d'activité

Conclusion et perspectives

Introduction

I - Habitat
Intelligent

Contexte de l'Habitat
Intelligent

Verrous scientifiques

II - Construction de
plateformes
d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des
thèmes abordés

Analyse des sons

Reconnaissance de la
parole distante

Suivi d'activité

Conclusion et
perspectives

Expérience pilote dans DOMUS

- ▶ **Domaine peu exploré**
 - peu d'études existantes concernant le domaine du son
 - technologie nouvelle

⇒ | inclure les utilisateurs
dans le processus de conception
- ▶ **Surveillance sonore continue**

⇒ système indiscret ?
- ▶ **Commande vocale**

⇒ utilisabilité ?

Étude d'usage et d'acceptabilité
(SWEET-HOME ANR VERSO-2009)

Expérience pilote dans DOMUS



- ▶ Participants :
 - 8 personnes âgées
 - 7 proches (famille ou amis)
 - 3 professionnels de santé
- ▶ Magicien d'Oz (pilotage par un opérateur en régie)
- ▶ Entretien en codécouverte :
 - visite accompagnée
 - démonstration
 - questionnaire

Scénarios

- ▶ Commande par la voix : des rideaux ou de la lumière ; le magicien d'Oz simule mal comprendre
- ▶ Pendant une émission TV : interruption par une téléconférence avec un proche
- ▶ Message d'alerte : fermer la porte d'entrée ou éteindre les plaques électriques
- ▶ Utilisation d'un agenda partagé

F. PORTET, M. VACHER, C. GOLANSKI, C. ROUX, B. MEILLON (PUC, in press)

Bilan de l'expérience pilote

Acceptabilité : aspects interrogés

- ergonomie générale
- utilisabilité de la commande vocale
- sécurité
- prise en charge par le système d'une activité
- intervention du système pendant une activité
- maintien du lien social

Analyse des réponses et conséquences

- ▶ La commande vocale par mot-clef est jugée utile pour le confort et la sécurité, mais réticence vis à vis de la caméra
⇒ *surdité du système hors mots-clef*
- ▶ Souhait de conserver l'autonomie des activités quotidiennes
- ▶ Crainte de devenir trop dépendant du système
- ▶ Crainte que la visioconférence se substitue aux visites
- ▶ Crainte de prise en mains du système par un intrus
⇒ *identification de présence de plus d'une personne*

I - Contexte, contraintes et attentes dans un habitat intelligent

Contexte de l'Habitat Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote : mise en évidence des questions éthiques

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la parole distante

Suivi d'activité

Conclusion et perspectives

Pré-étude *in vitro*

- ▶ Détection des évènements sonores dans le bruit (bruitage artificiel avec bruit HIS)

claquement de portes, vaisselle, sonneries, pas, serrures, bris de verre, cris, parole

RSB	0dB	+10dB	+20dB	+40dB
Taux d'égale erreur (%)	6,5	0	0	0

Excellents résultats si $RSB \geq 10dB$

- ▶ Classification des évènements sonores (bruitage artificiel avec bruit HIS)

Taux d'erreur de classification
LFCC

claquement de portes, vaisselle, sonneries, pas, serrures, bris de verre, cris, parole, chute d'objet

	RSB	0dB	+10dB	+20dB	+40dB	$\geq +50dB$
GMM	TEC(%)	23,6	15,4	16,5	10,2	3,2
HMM	TEC(%)	29,7	12,6	10,8	9	2

Résultats insuffisants si $RSB \leq 15dB$

Analyse sonore : mise à l'épreuve du réel en HIS

Introduction

I - Habitat Intelligent

Contexte de l'Habitat

Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la
parole distante

Suivi d'activité

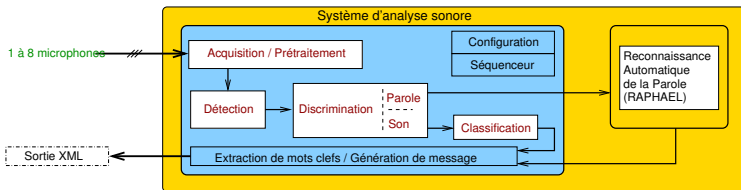
Conclusion et perspectives

Deux expériences :

1. Reconnaissance vocale des appels de détresse
2. Analyse sonore dans la vie quotidienne

M. VACHER, F. PORTET, A. FLEURY, N. NOURY (IEEE HEALTHCOM, 2010)

Les logiciels pour les 2 expériences



AuditHIS (Analyse sonore temps réel)

- ▶ Analyse simultanée sur 8 canaux
- ▶ Fréquence d'acquisition : 16 kHz
- ▶ Estimation du rapport signal sur bruit (RSB)
- ▶ Plusieurs modules s'exécutant en temps réel

M. VACHER, A. FLEURY, J.-F. SERIGNAT, N. NOURY, H. GLASSON
(INTERSPEECH, 2008)

Raphael (Reconnaissance Automatique de la Parole)

- ▶ Modèles acoustiques multilocuteurs

Expérience 1 : reconnaissance des appels de détresse

Mise en place des modèles de langage

- ▶ **Modèle de langage spécialisé à petit vocabulaire : 299 unigrammes**
 - 238 énoncés courts de conversation courante
"j'ai bu ma tisane", "il fait beau", "à demain" ...
 - 39 ordres domotiques
 - **93 énoncés courts de détresse**
"au secours", "appelez le SAMU", "j'ai mal" ...

- ▶ **Corpus à analyser : 10 locuteurs français non âgés**
 - enregistrements dans le HIS
 - 20 phrases de détresse et 10 phrases usuelles dans la chambre
 - 3 conversations téléphoniques de 5 phrases dans le séjour
 - ⇒ - 197 énoncés courts de détresse
 - ⇒ - 232 énoncés courts de conversation courantes :

Expérience 1 : reconnaissance des appels de détresse

Aucune autre étude publiée, 1^{ers} résultats

$$RSB_{experimental} = 14,4 \pm 6,5dB$$

► Discrimination son / parole

Taux de bonne discrimination	15dB	20dB	Voix avec meilleur RSB	Voix \geq 80% du meilleur RSB
<i>in vitro</i>	3,8%	5,1%		
<i>in vivo</i>			8,3%	6%

⇒ Modèles 7 classes de son et parole inadaptés (sons de vaisselle !!)

► Reconnaissance des appels de détresse

Taux d'erreur (alarme et normal)	Voix avec le meilleur RSB	3 voix ayant le meilleur RSB
<i>in vivo</i>	33,4%	34,5%

⇒ Modèles acoustiques classiques inadaptés à la parole distante

M. VACHER, A. FLEURY, J.-F. SERIGNAT, N. NOURY, H. GLASSON
(INTERSPEECH, 2008)

Expérience 2 : analyse sonore dans la vie quotidienne

Conditions

- ▶ Expérimentation dans le HIS
- ▶ 15 participants jeunes - expérience 1h
- ▶ Ambiance non contrôlée : bruit à l'extérieur de l'appartement
- ▶ Faibles contraintes
- ▶ 7 Activités de la Vie Quotidienne (AVQ) :
 - Dormir ;
 - Se détendre ;
 - Se nourrir ;
 - S'habiller et se déshabiller ;
 - Éliminer ;
 - Faire sa toilette ;
 - Communiquer.

Expérience 2 : corpus d'événements sonores HIS

Introduction

I - Habitat Intelligent

Contexte de l'Habitat Intelligent
Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS
DOMUS
Expérience pilote

III - Trois des thèmes abordés

Analyse des sons
Reconnaissance de la parole distante
Suivi d'activité

Conclusion et perspectives

TABLE: Sons de la vie courante

Catégorie	Nb. obs.	RSB moyen (dB)	Durée moyenne (ms)	Durée totale (s)
Manipulation d'objets	1302 (69%)	11,9	58,6	76,3 (37,5%)
Appareils	72 (3,8%)	8,0	208,5	15,1 (7,5%)
Sons extérieurs	45 (2,4%)	9,0	174,4	7,9 (4%)
Sons humains (sauf parole)	36 (2%)	12,0	100,8	3,4 (1,7%)
Eau courante	36 (2%)	10,1	1756,1	63,2 (31,1%)
Autres sons	395 (21%)	9,5	93,9	37,1 (18,2%)
Nombre total des observations sonores (sauf parole)	1886 (100%)	11,2	107,8	203,3 (100%)

20% de la durée du corpus

Autres sons = sons non identifiables/sons mélangés

Grande difficulté des tâches d'annotation et de classification

Expérience 2 : nouveau problème posé

- ▶ Attributs inadaptés : classification des sons à partir de leur production
- ⇒ Caractérisation des sons à partir de leurs caractéristiques acoustiques
 - périodicité/non périodicité
 - durée
 - forme de l'enveloppe
 - répétitivité de motifs
 - ...
- ▶ Modèles statistiques inadaptés
 - Faible taille du corpus enregistré :
3mn 23s \iff 15h
(pour 300mn \iff 1500h)
 - Sons imprévisibles
- ▶ Méthode de classification hiérarchique ?
- ▶ Amélioration de la classification en utilisant le contexte

I - Contexte, contraintes et attentes dans un habitat intelligent

Contexte de l'Habitat Intelligent
Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS
DOMUS
Expérience pilote : mise en évidence des questions éthiques

III - Trois des thèmes abordés

Analyse des sons
Reconnaissance de la parole distante
Suivi d'activité

Conclusion et perspectives

Parole distante : analyse multicanal

Introduction

I - Habitat Intelligent

Contexte de l'Habitat
Intelligent
Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS
DOMUS
Expérience pilote

III - Trois des thèmes abordés

Analyse des sons
Reconnaissance de la
parole distante
Suivi d'activité

Conclusion et perspectives

- ▶ Solution usuelle : RAP sur le canal avec le meilleur RSB
- ▶ Ma proposition : de nouvelles possibilités en utilisant les autres canaux
 1. Fusion de données pour améliorer la reconnaissance
 - Alignement de faisceaux - *Beam Forming*
 - Vote, par exemple méthode ROVER
 - Décodage guidé par les données
 2. Suppression d'une source
 - Annulation d'écho
 - Séparation de sources aveugles - *Blind source Separation*

Données expérimentales de la méthode 1

- ▶ corpus multimodal SWEET-HOME enregistré dans DOMUS
- ▶ distance minimale locuteur/microphone = 2 à 3 m
- ▶ 21 locuteurs (7 femmes)
- ▶ 7 voies simultanées
- ▶ phrases prédéfinies
- ▶ estimation du **RSB** sur chaque voie
- ▶ enregistrement :
 - ▶ **Phase 1** : Activités de la Vie Quotidienne sur 2h
⇒ 2 conversations téléphoniques (**40 phrases**)
 - ▶ **Phase 2** : lecture de 44 phrases dans 3 pièces et dans 3 conditions (aspirateur, radio, appareils éteints)
utilisation des **44 phrases** lues dans la chambre sans radio ni aspirateur

Outils de la méthode 1 : Reconnaissance Automatique de la Parole (RAP)

- ▶ Système Speeral
 - Laboratoire LIA - Avignon (2002)
 - algorithme de recherche A* sur un treillis de phonèmes
 - modèles acoustiques de phonèmes en contexte, HMM 3 états (100h)
 - PLP (Prédiction Linéaire Perceptuelle ou *Perceptual Linear Predictive*)
 - configuration *1xReal Time*
- ▶ Apprentissage
 - phrases de la **Phase 1** annotées manuellement (2min par locuteur)
 - adaptation MLLR pour chaque locuteur
- ▶ Jeu de test
 - phrases de la **Phase 2** (annotation manuelle)

B. LECOUTEUX, M. VACHER, F. PORTET (INTERSPEECH, 2011)

Outils de la méthode 1 : modèles de langage

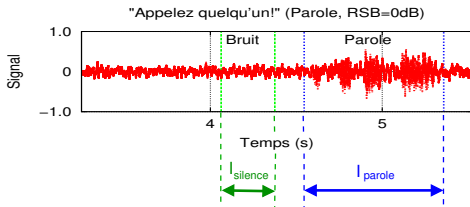
- ▶ Reconnaissance de phrases-clefs
 - Ordres domotiques :
"Allumez la lumière", "Montez la température"
 - Situation de détresse :
"Appelez le docteur", "Aidez moi"

- ▶ Système Speeral
 - Modèle de langage générique (MG)
 - Modèle spécialisé (MS) entraîné sur les phrases à lire
 - meilleure configuration : MG 10%, MS 90 %

Indice de confiance à partir du RSB

Pondération du vote pour chaque canal $j, j \in [1; 7]$:

$$RSB(e) = 10 * \log \left(\frac{\sum_{t \in [0; \Delta T]} s^2(t)}{\sum_{t \in [0; \Delta T]} b^2(t)} \right) \quad (1)$$



$$RSB(P_{k,j}) = 10 * \log \left(\frac{\frac{\sum_{n \in I_{parole}} s(n)^2}{|I_{parole}|}}{\frac{\sum_{n \in I_{silence}} s(n)^2}{|I_{silence}|}} \right) \quad (2)$$

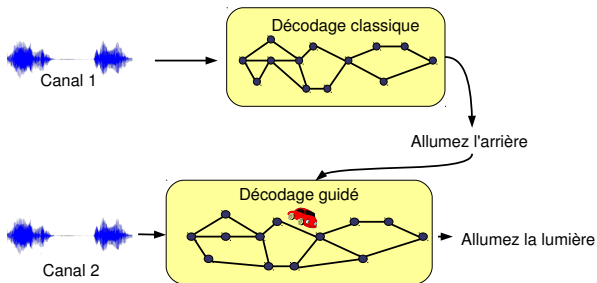
Indice de confiance proposé pour la méthode ROVER : $\phi_{k,j} = \frac{2^{RSB(P_{k,j})}}{\sum_{i=1}^7 2^{RSB(P_{k,i})}}$ (3)

Recognizer Output Voting Error Reduction

FISCUS (1997)

Solutions proposées : décodage guidé DDA

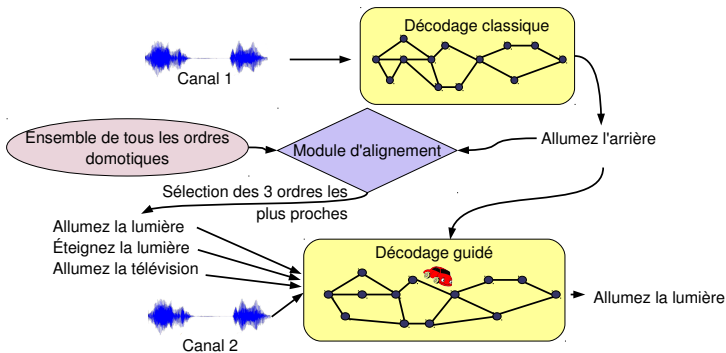
Driven Decoding Algorithm



- ▶ Canal de meilleur RSB sur la 1^e voie
- ▶ Décodage du canal de 2^e meilleur RSB guidé par le 1^{er} décodage
- ▶ Avantage par rapport à ROVER
 - Seconde passe très rapide
 - Au moins 3 voies pour ROVER, seulement 2 pour DDA

B. LECOUTEUX, G. LINARÈS, Y. ESTÈVE, G. GRAVIER (ICASSP, 2008)

Solutions proposées : décodage guidé DDA à 2 niveaux (DDA-2)



- ▶ Canal de meilleur RSB sur la 1^e voie
- ▶ Localisation des ordres domotiques par alignement DTW au niveau phonétique pour guider la 2^e voie

B. LECOUEUX, M. VACHER, F. PORTET (INTERSPEECH, 2011)

Résultats de la méthode 1 (reconnaissance multicanal)

	$TEM_{\pm\sigma}(\%)$
canal de meilleur RSB	$18,3 \pm 12,1$
Beamforming	$16,8 \pm 8,3$
ROVER sur 7 canaux	$12,2 \pm 6,1$
DDA_{+RSB}	$11,4 \pm 5,6$
$DDA_{2-level+RSB}$	$8,8 \pm 3,7$

Conclusion

- ▶ amélioration significative/analyse monovoie mais nécessité de **modèles acoustiques parole distante**
- ▶ ROVER : plus gourmand en temps calcul
- ▶ DDA-2 : bonnes performances pour des phrases prédéfinies
⇒ **intéressant pour acceptabilité/éthique**

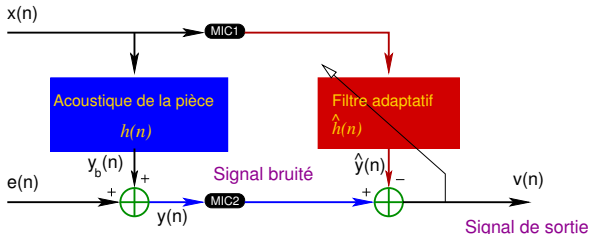
Méthode 2 : reconnaissance en présence de sources de bruit connues

Utilisation de 2 microphones

TV ou radio



Parole



$$y_b(n) = h(n) * x(n)$$

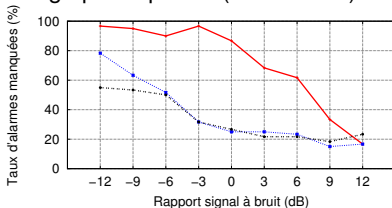
$$y(n) = e(n) + h(n) * x(n)$$

$$v(n) = e(n) + h(n) * x(n) - \hat{h}(n) * x(n)$$

$$v(n) = e(n) + \varepsilon(n)$$

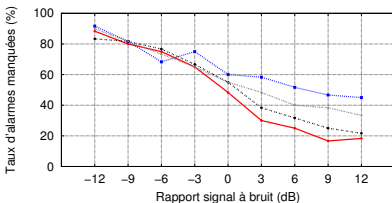
Expérience 2 : résultats

Bruitage par la parole (France Info)



Parole bruitée sans traitement
Après annulation d'écho
Après annulation d'écho et post-filtrage

Bruitage par la musique (classique et moderne)



(C) après annulation d'écho
(C) après annulation d'écho et post-filtrage
(M) après annulation d'écho
(M) après annulation d'écho et post-filtrage

M. VACHER, A. FLEURY, N. GUIRAND, J.-F. SERIGNAT, N. NOURY
(SPED, 2009)

I - Contexte, contraintes et attentes dans un habitat intelligent

Contexte de l'Habitat Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote : mise en évidence des questions éthiques

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la parole distante

Suivi d'activité

Conclusion et perspectives

Activités de la Vie Quotidienne (AVQ)

Évaluation du niveau de dépendance de la personne âgée

- ▶ Établi annuellement et manuellement par le médecin gériatre
- ▶ Se base sur les scores AGGIR ou ADL
- ▶ Détermine le niveau d'allocation (APA...)

Grille AGGIR **A**utonomie **G**érontologique - **G**roupe **I**so-Ressource

- 10 variables discriminantes (toilette, habillage, ...)
- 7 variables illustratives (cuisine, ménage, ...)

(A) Aucune aide, (B) aide partielle, (C) aide totale.

Grille ADL **A**ctivity of **D**aily **L**iving (Katz)

- Hygiène Corporelle
- Habillage
- Faire sa toilette
- Locomotion
- Elimination
- Prendre un repas

Données expérimentales

- ▶ Expérimentation dans l'appartement HIS de la faculté de médecine
- ▶ Capteurs utilisés :
actimètre, détecteur de présence IR, contact de porte, capteur environnemental, microphone
- ▶ Scénario
 - Une entrée et une sortie de l'appartement
 - Simulation d'une matinée (7 AVQ)
 - Chaque activité est faite au moins une fois sans contrainte de temps ou d'ordre
 - Activité d'hygiène demandée : se laver les mains et les dents
 - Mouvement de synchronisation pour le capteur cinématique
- ▶ 13 participants (6 femmes)
 - 24 à 43 ans
 - 51min 40s (entre 23min 11s et 1h 35min 44s)

Données expérimentales

Introduction

I - Habitat

Intelligent

Contexte de l'Habitat

Intelligent

Verrous scientifiques

II - Construction de
plateformes

d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des

thèmes abordés

Analyse des sons

Reconnaissance de la
parole distante

Suivi d'activité

Conclusion et
perspectives

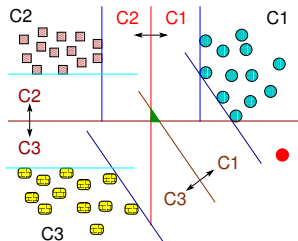
Distribution des trames suivant les AVQ (durée d'une trame = 3min)

Activité	Nombre de trames	Pourcentage (%)
Dormir	49	21,1
Préparer et prendre le petit déjeuner	45	19,4
Se vêtir et se dévêtir	16	6,9
Se reposer	73	31,4
Faire sa toilette	14	6
Élimination	16	6,9
Communication	19	8,1
Total	232	100

Méthode : classification SVM

Séparateur à Vaste Marge

- ▶ SVM (Boser et Vapnik, 1992) : classificateur binaire reposant sur les notions de marge maximale et de noyaux
- ▶ Méthodes existantes : M-SVM, un-contre-tous et un-contre-un
- ▶ Retenue : un-contre-un



- $\frac{N \cdot (N-1)}{2}$ classificateurs (N classes)
- Différentiation des classes C_i et C_j , $0 < i \leq N$ et $0 < j < i$
- Vote majoritaire :

$$C = \max_{k=1..N} \text{Card}(\{y_{i,j}\} \cap \{k\})$$

Résultats

- ▶ protocole « tous sauf un » (*leave one out*)
- ▶ noyau gaussien
- ▶ taux de bonne classification = 86%

Activité	Taux de bonne classification						
	Dormir	Se reposer	Se vêtir/ se dévêtir	Se nourrir	Éliminer	Faire sa toilette	Communiquer
Dormir	93,9%	0%	0%	0%	0%	0%	0%
Se reposer	13,8%	78,1%	1,3%	1,3%	4,2%	1,3%	0%
Se vêtir/se dévêtir	6,25%	12,5%	75%	0%	0%	0%	6,25%
Se nourrir	0%	0%	2,2%	97,8%	0%	0%	0%
Éliminer	0%	0%	0%	6,25%	93,75%	0%	0%
Faire sa toilette	7,1%	0%	0%	7,1%	21,5%	64,3%	0%
Communiquer	0%	5,3%	5,9%	0%	0%	0%	89,5%

- ▶ les 2 classes les moins présentes sont les moins bien classées

IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE,
A. FLEURY, M. VACHER, N. NOURY (2009)

Classification SVM des AVQ : conclusions

Introduction

I - Habitat Intelligent

Contexte de l'Habitat

Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la
parole distante

Suivi d'activité

Conclusion et perspectives

- ▶ classification de 7 AVQ, taux d'erreur 86%
- ▶ comparaison des résultats avec ceux de la littérature :
 - taux d'erreur = 88% Philipose et al. (2004)
mais nombreux capteurs, marquage de tous les objets
- ▶ pistes pour amélioration :
 - sélection d'attributs
 - paramètres incluant le temps de présence dans la pièce
 - classe de rejet

I - Contexte, contraintes et attentes dans un habitat intelligent

Contexte de l'Habitat Intelligent
Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS
DOMUS
Expérience pilote : mise en évidence des questions éthiques

III - Trois des thèmes abordés

Analyse des sons
Reconnaissance de la parole distante
Suivi d'activité

Conclusion et perspectives

Perspectives à court terme

1. Parole en conditions distantes

- ▶ Cas des sources de bruit non identifiées dans l'appartement ou à l'extérieur

 - ▶ *Outils* :
 - Corpus multimodal SWEET-HOME
 - Séparation de sources aveugles : problèmes à résoudre
 - identification de la voie parole
 - cas des mélanges convolutifs
- Collaboration L. Girin & C. Servièrre GIPSA Lab.*
- Analyse Factorielle
- Collaboration prévue avec D. Matrouf LIA*

Introduction

I - Habitat Intelligent

Contexte de l'Habitat
Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la
parole distante

Suivi d'activité

Conclusion et perspectives

2. Caractérisation de la voix âgée

Intérêt pour l'étude des caractéristiques phonétiques et prosodiques de la parole âgée

- ▶ Protocole à construire pour l'acquisition de gros corpus (parole lue ou parole spontanée)
- ▶ Des modèles à établir :
 - jusqu'où adapter les modèles de parole non âgée ?
 - nécessité de construire de nouveaux modèles ?
- ▶ Quelles sont les phrases de détresse ?
Introduire la reconnaissance de l'émotion
Reconnaissance de la parole avec émotion

Perspectives à court terme

3. Reconnaissance des AVQ

- ▶ *Classification des AVQ*
- ▶ Détection du début et de la fin
- ▶ Enrichissement des modèles au fil du temps
- ▶ Reconnaissance en temps réel

Outils :

- ▶ Données appartement DOMUS
- ▶ méthode SVM
- ▶ **méthode MLN** (*Markov Logic Network*)
Réseaux Logiques de Markov

Collaboration A. Fleury, École des Mines de Douai

Perspectives à court terme

4. Interprétation des ordres vocaux en fonction du contexte

« allume la lumière » [?]

laquelle ?

comment ?

[chambre]

[bureau]

...

lampe de chevet

lampe bureau

...

[lecture]

[repos]

atténuée

forte

Outils :

- ▶ Corpus multimodal SWEET-HOME
- ▶ arbres de décision ou règles expertes (*Portet, 2005*) (*Callens, 2008*)
- ▶ informations temporelles, imprécises et incertaines (*Dempster-Shafer, 1976*)
- ▶ combinaison de preuves (*Hong, 2009*)

Perspectives à long terme

Introduction

I - Habitat Intelligent

Contexte de l'Habitat
Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la
parole distante

Suivi d'activité

Conclusion et perspectives

1. Étude de l'évolution de la richesse du vocabulaire
 - ▶ Alzheimer : après 85 ans
(25% des femmes et 20% des hommes)
 - ▶ détection de la réduction progressive du vocabulaire
utilisé

Résultats CIRDO sur la voix âgée

2. Cas des personnes âgées qui reviennent à la langue
de leur enfance
 - ▶ RAP multilingue
 - ▶ **Converser for HealthCare, M. SELIGMAN**
personnel de santé anglophones
⇔ patients hispanophones
 - ▶ *Collaboration envisagée : Pr. Kawahara (Kyoto)*



Merci à toutes les personnes avec qui j'ai collaboré ces dernières années :
Jean-François Serignat, Jean Caelen, Eric Castelli, Norbert Noury, Vincent Rialle, Dan Istrate, Anthony Fleury, Hubert Glasson, Laurent Besacier, Dominique Vaufreydaz, Yannick Fouquet, Denis Tuffelli, Solange Rossato, François Portet, Hervé Blanchon, Christian Boitet, Benjamin Lecouteux, Pedro Chahuara, Bernard Cassagne, Jean-Claude Durand, Brigitte Meillon, Nicolas Bonnefond, Francis Jambon, Christian Perrot, Sylvie Pons, Gilles Virone, Frédéric Aman, Stéphane Chaillol, Sylvain Méniard, Viet Bac Le, Nicolas Gac, Noé Guirand, Remus Dugheanu, Caroline Golanski, Camille Roux, Anne Pellegrin, Mathieu Gallissot, Christophe Debeaux, Yohan Payan, Claude Aynaud, Quentin Lefol, Nicolas Vuillerme, Corentin Dijoux, Ayoub Maatallaoui, Didac Cristobal Canals, Mihai Bobu, Josep Torras-Flaquer, Pelayo Menendez Garcia, Takoua Ouled ben Ali, Bruno Grentzinger, Rania Bouaziz, Alexandre Aminot, Yuko Sasa, Juline Legrand, Thierry Chevalier, Thierry Joubert, Saïda Bouakaz, Serge Smidtas. . .

Introduction

I - Habitat Intelligent

Contexte de l'Habitat Intelligent

Verrous scientifiques

II - Construction de plateformes d'expérimentation

HIS

DOMUS

Expérience pilote

III - Trois des thèmes abordés

Analyse des sons

Reconnaissance de la parole distante

Suivi d'activité

Conclusion et perspectives

Mes 6 publications marquantes



M. Vacher, A. Fleury, F. Portet, J.-F. Serignat, and N. Noury

Complete Sound and Speech Recognition System for Health Smart Homes : Application to the Recognition of Activities of Daily Living,

New Developments in Biomedical Engineering, INTECH Book, pp. 645 – 673, Feb. 2010.



Portet, F., Vacher, M., Golanski, C., Roux, C., and Meillon, B.

Design and evaluation of a smart home voice interface for the elderly - acceptability and objection aspects

Personal and Ubiquitous Computing, pages 1-30 (in press)



Vacher, M., Portet, F., Fleury, A., and Noury, N.

Development of audio sensing technology for ambient assisted living : Applications and challenges

International Journal of E-Health and Medical Communications, 2(1) :35-54, 2011.



A. Fleury, M. Vacher and N. Noury

SVM-Based Multi-Modal Classification of Activities of Daily Living in Health Smart Homes : Sensors, Algorithms and First Experimental Results

Information Technology in Biomedicine, IEEE Trans. on, 14(2) :274-283, 2010.



Istrate, D., Castelli, E., Vacher, M., Besacier, L., and Serignat, J.-F. (2006)

Information extraction from sound for medical telemonitoring

Information Technology in Biomedicine, IEEE Transactions on, 10(2) :264-274, 2006.



M. Vacher, D. Istrate and J.-F. Serignat

Sound Detection and Classification through Transient Models using Wavelet Coefficient Trees

12th European Signal Processing Conference EUSIPCO, Vienne (Autriche), pp. 1171-1174, 2004.