

# Reconnaissance des sons de l'environnement dans un contexte domotique

Mohamed El Amine SEHILI

ESIGETEL-TSP

Le 05 Juillet 2013

**Directeur de thèse :** Bernadette Dorizzi (Télécom SudParis)

**Co-encadrants :** Dan Istrate (Esigetel), Jérôme Boudy (Télécom SudParis)

# Plan

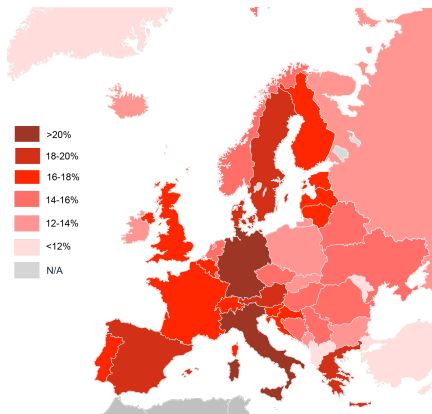
- 1 Contexte : Projet Sweet-Home
- 2 Objectifs de la thèse
- 3 État de l'art en classification des sons de l'environnement
- 4 Première contribution : Mise en œuvre de différents classifieurs issus de la RAL, pour la classification des sons
- 5 Deuxième contribution : Utilisation de représentations adaptées aux différentes classes de sons
- 6 Conclusions et Perspectives

# Plan

- 1 Contexte : Projet Sweet-Home
- 2 Objectifs de la thèse
- 3 État de l'art en classification des sons de l'environnement
- 4 Première contribution : Mise en œuvre de différents classifieurs issus de la RAL, pour la classification des sons
- 5 Deuxième contribution : Utilisation de représentations adaptées aux différentes classes de sons
- 6 Conclusions et Perspectives

# Changements démographiques : un aperçu

Un taux de vieillissement en croissance...



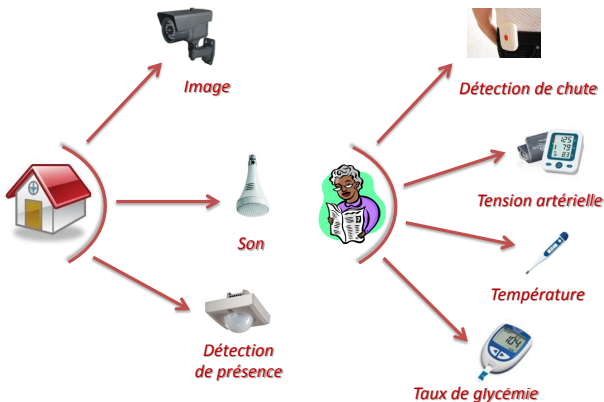
**Figure:** Personnes âgées de 65 ans et plus en Europe en 2010 (Basé sur les données du CIA World Factbook)

# Problématique : maintien des personnes âgées au domicile

# Problématique : maintien des personnes âgées au domicile

## Télé-vigilance médicale :

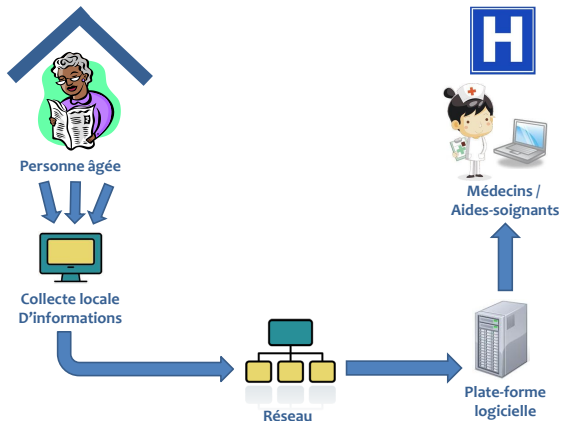
- Utiliser plusieurs types de capteurs (informations sur l'environnement et sur la personne)



# Problématique : maintien des personnes âgées au domicile

## Télé-vigilance médicale :

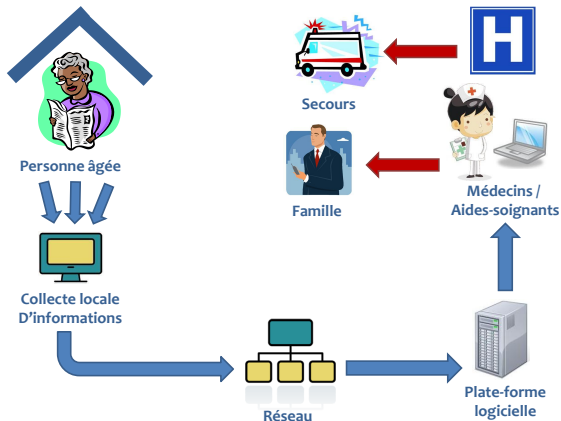
- Utiliser les avancées en télécommunications et en technologies de l'information



# Problématique : maintien des personnes âgées au domicile

## Télé-vigilance médicale :

- Utiliser les avancées en télécommunications et en technologies de l'information





# Projet Sweet-Home

Un projet ANR (*Agence Nationale de la Recherche*) VERSO de recherche industrielle (2010-2013)



# Projet Sweet-Home

Un projet ANR (*Agence Nationale de la Recherche*) VERSO de recherche industrielle (2010-2013)



# Projet Sweet-Home

Un projet ANR (*Agence Nationale de la Recherche*) VERSO de recherche industrielle (2010-2013)



## Détection des situations de détresse :

- Reconnaissance et analyse de la **parole**
- Reconnaissance des **sons** de l'environnement

# Projet Sweet-Home

Un projet ANR (*Agence Nationale de la Recherche*) VERSO de recherche industrielle (2010-2013)



## Détection des situations de détresse :

- Reconnaissance et analyse de la **parole**
- Reconnaissance des **sons** de l'environnement

## Automatisation de la maison :

- Commande **vocale**
- Commande tactile

# Plan

- 1 Contexte : Projet Sweet-Home
- 2 Objectifs de la thèse**
- 3 État de l'art en classification des sons de l'environnement
- 4 Première contribution : Mise en œuvre de différents classifieurs issus de la RAL, pour la classification des sons
- 5 Deuxième contribution : Utilisation de représentations adaptées aux différentes classes de sons
- 6 Conclusions et Perspectives

# Motivations et défis de la thèse

**Objectif :** Reconnaître les sons de l'environnement dans une maison intelligente.

# Motivations et défis de la thèse

**Objectif :** Reconnaître les sons de l'environnement dans une maison intelligente.

**C'est à dire :**

- Séparer la **parole** des **autres sons** → améliorer les performances de la reconnaissance de la parole
- Reconnaître des **sons de la vie courante** → aider le système de décision à détecter des situations anormales

# Motivations et défis de la thèse

**Objectif :** Reconnaître les sons de l'environnement dans une maison intelligente.

**C'est à dire :**

- Séparer la **parole** des **autres sons** → améliorer les performances de la reconnaissance de la parole
- Reconnaître des **sons de la vie courante** → aider le système de décision à détecter des situations anormales

**Défis :**



# Motivations et défis de la thèse

**Objectif :** Reconnaître les sons de l'environnement dans une maison intelligente.

**C'est à dire :**

- Séparer la **parole** des **autres sons** → améliorer les performances de la reconnaissance de la parole
- Reconnaître des **sons de la vie courante** → aider le système de décision à détecter des situations anormales

**Défis :**

- Nombre élevé de classes de sons
- Variations entre les sons d'une même classe et similitudes entre les sons de classes différentes
- Bruit de l'environnement (**tout son sans intérêt**)
- Acquisition multi-canaux (**plusieurs micros dans la maison**)

# Classes de sons

## Sons humains



*Respiration*



*Toux*



*Rire*



*Éternuement*



*Pleur*

## Sons non humains



*Sèche Cheveux*



*Bris de glace*

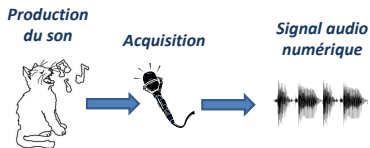


*Claquement de porte*

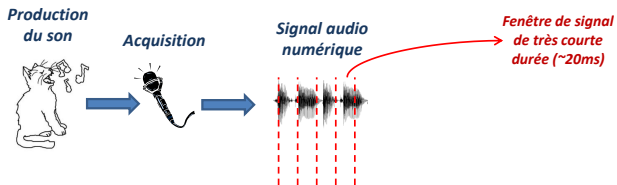


*Clés*

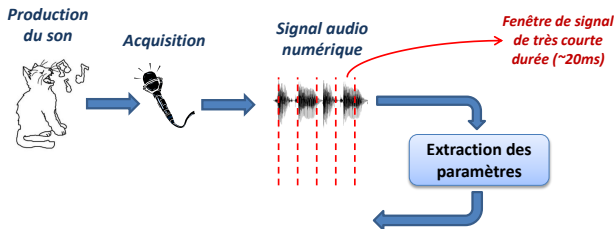
# Architecture d'un système de reconnaissance du son



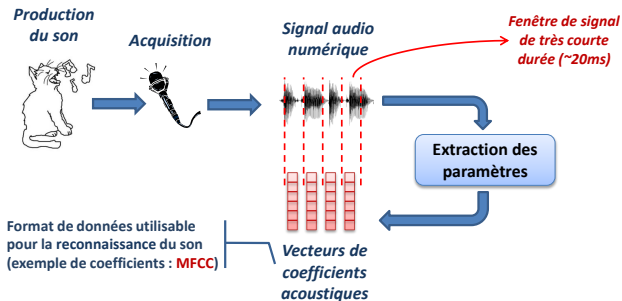
# Architecture d'un système de reconnaissance du son



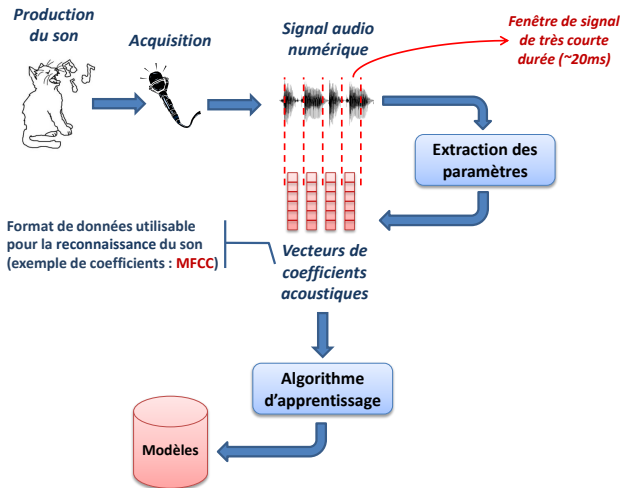
# Architecture d'un système de reconnaissance du son



# Architecture d'un système de reconnaissance du son



# Architecture d'un système de reconnaissance du son



# Voies de recherche

Extraction des  
paramètres

Algorithme  
d'apprentissage



# Voies de recherche

Extraction des  
paramètres

**Première voie de recherche :**  
Exploration et utilisation de plusieurs  
méthodes de reconnaissance

Algorithme  
d'apprentissage

# Voies de recherche

**Seconde voie de recherche :**  
Utilisation de plusieurs familles  
de coefficients acoustiques

Extraction des  
paramètres

**Première voie de recherche :**  
Exploration et utilisation de plusieurs  
méthodes de reconnaissance

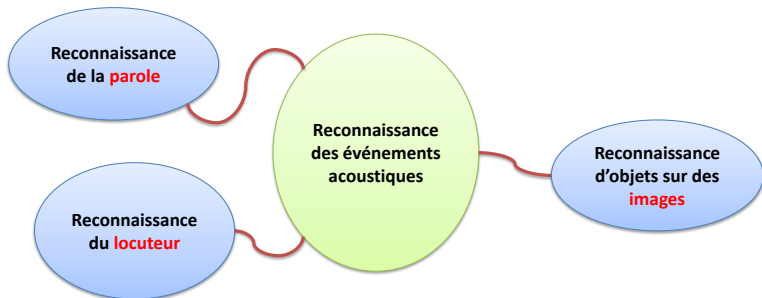
Algorithme  
d'apprentissage

# Plan

- 1 Contexte : Projet Sweet-Home
- 2 Objectifs de la thèse
- 3 État de l'art en classification des sons de l'environnement**
- 4 Première contribution : Mise en œuvre de différents classifieurs issus de la RAL, pour la classification des sons
- 5 Deuxième contribution : Utilisation de représentations adaptées aux différentes classes de sons
- 6 Conclusions et Perspectives

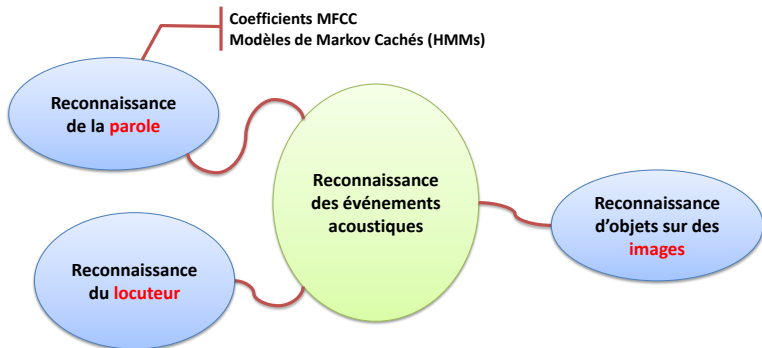
# Domaines les plus influents

Peu de travaux spécifiques...



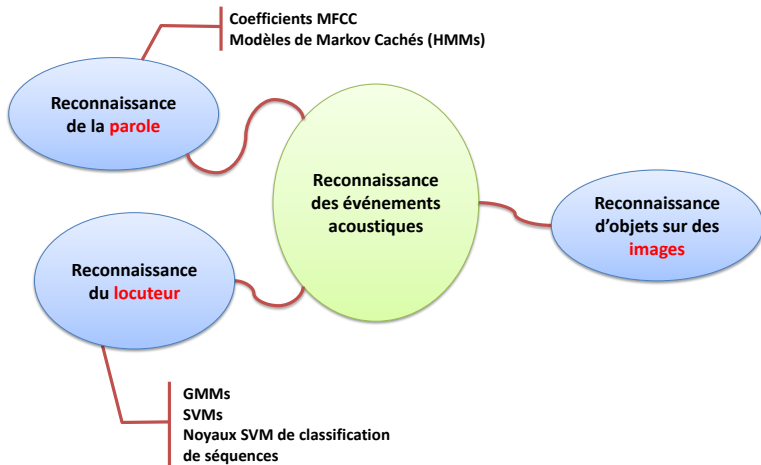
# Domaines les plus influents

Peu de travaux spécifiques...



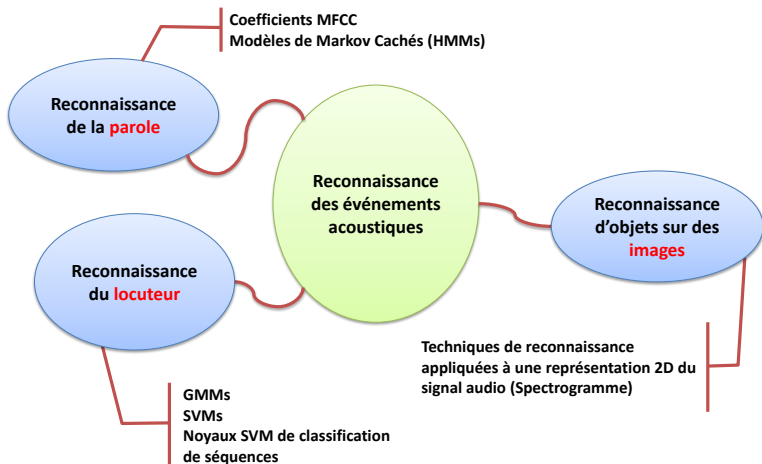
# Domaines les plus influents

Peu de travaux spécifiques...



# Domaines les plus influents

Peu de travaux spécifiques...



# Plan

- 1 Contexte : Projet Sweet-Home
- 2 Objectifs de la thèse
- 3 État de l'art en classification des sons de l'environnement
- 4 Première contribution : Mise en œuvre de différents classifieurs issus de la RAL, pour la classification des sons**
- 5 Deuxième contribution : Utilisation de représentations adaptées aux différentes classes de sons
- 6 Conclusions et Perspectives



# Motivations

## **Pourquoi s'inspirer de la Reconnaissance Automatique du Locuteur (RAL) ?**

# Motivations

## Pourquoi s'inspirer de la Reconnaissance Automatique du Locuteur (RAL) ?

- Domaine plus ou moins proche (un locuteur  $\approx$  une classe de sons)

# Motivations

## Pourquoi s'inspirer de la Reconnaissance Automatique du Locuteur (RAL) ?

- Domaine plus ou moins proche (un locuteur  $\approx$  une classe de sons)
- État de l'art assez riche

# Motivations

## Pourquoi s'inspirer de la Reconnaissance Automatique du Locuteur (RAL) ?

- Domaine plus ou moins proche (un locuteur  $\approx$  une classe de sons)
- État de l'art assez riche

## Méthodes retenues

# Motivations

## Pourquoi s'inspirer de la Reconnaissance Automatique du Locuteur (RAL) ?

- Domaine plus ou moins proche (un locuteur  $\approx$  une classe de sons)
- État de l'art assez riche

## Méthodes retenues

- Modèles de Mélanges Gaussiens (**GMM** pour *Gaussian Mixture Models*)
- Machines à Vecteurs de Support (**SVM** pour *Support Vector Machines*). Deux variantes utilisées :

# Motivations

## Pourquoi s'inspirer de la Reconnaissance Automatique du Locuteur (RAL) ?

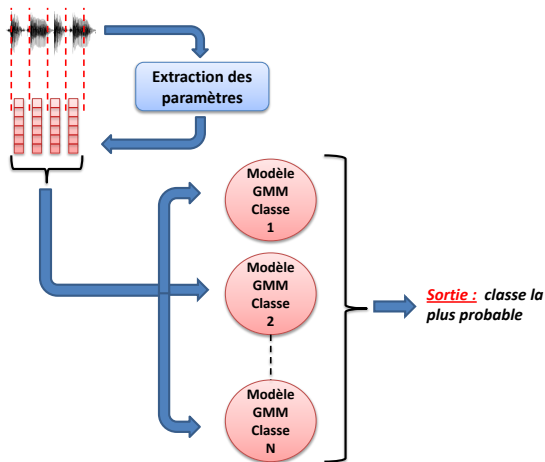
- Domaine plus ou moins proche (un locuteur  $\approx$  une classe de sons)
- État de l'art assez riche

## Méthodes retenues

- Modèles de Mélanges Gaussiens (**GMM** pour *Gaussian Mixture Models*)
- Machines à Vecteurs de Support (**SVM** pour *Support Vector Machines*). Deux variantes utilisées :
  - Utilisation directe des **vecteurs acoustiques** en entrée
  - Utilisation des vecteurs acoustiques **après transformation**

# GMM pour la reconnaissance du son

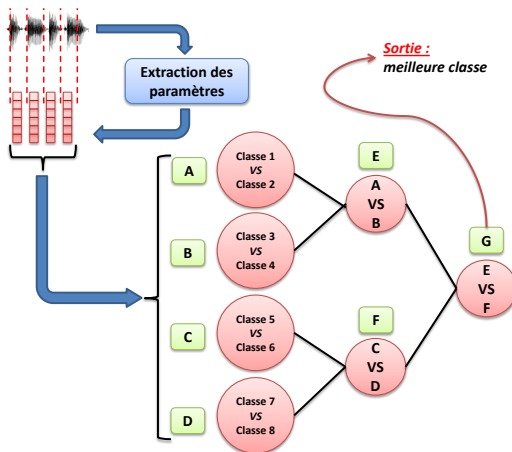
**GMM** : un modèle probabiliste représentant un mélange de plusieurs distributions.



# SVM pour la reconnaissance du son

**SVM** : une méthode classification supervisée permettant de discriminer deux classes

**Première variante** : classification directe des vecteurs acoustiques

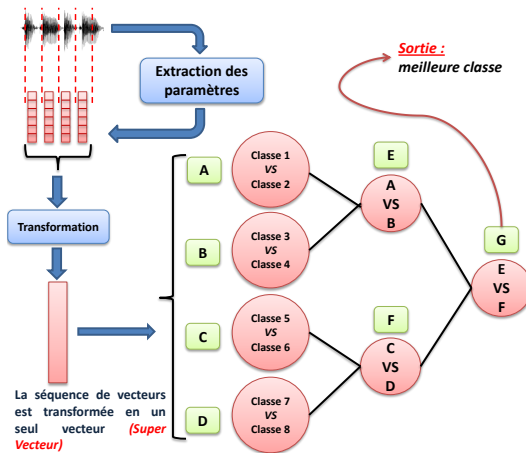


Approche **peu efficace** et **coûteuse** en temps de calcul



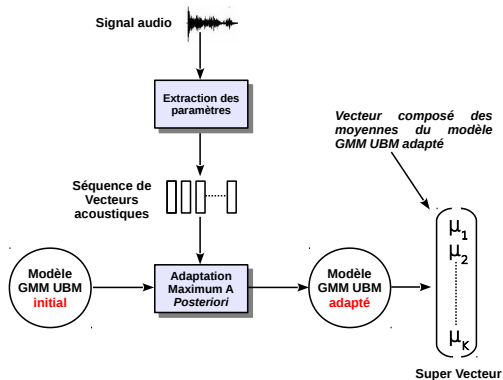
# SVM pour la reconnaissance du son

**Seconde variante** : avant la classification, transformer la séquence de vecteurs acoustiques en un seul grand vecteur, appelé **Super Vecteur**

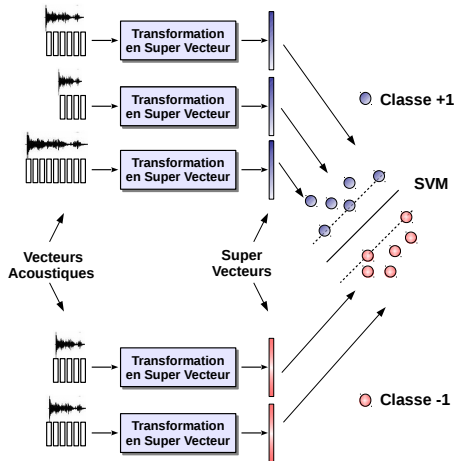


# Méthode de transformation retenue

Noyau **SVM-GSL** (SVM GMM Super Vector Linear Kernel) [Campbell et al, 2006]



# SVM : utilisation d'un noyau de classification de séquences



# Tests et résultats

## Base de données de l'Esigetel :

Classe de sons	Nombre de fichiers	Durée totale (sec).	Conditions d'enreg. variables
Breathing	50	106.44	Oui
Cough	62	181.69	Oui
Dishes	98	303.77	Oui
DoorClapping	114	62.70	Oui
DoorOpening	21	138.94	<b>Non</b>
ElectricalShaver	62	420.33	<b>Non</b>
FemaleCry	36	268.19	Oui
FemaleScream	70	216.83	Oui
GlassBreaking	101	99.52	Oui
HairDryer	40	224.86	<b>Non</b>
HandsClapping	54	218.65	Oui
Keys	36	166.34	Oui
Laugh	49	272.65	Oui
MaleScream	87	202.11	Oui
Paper	63	330.66	Oui
Sneeze	32	51.67	Oui
Water	54	484.72	Oui
Yawn	20	95.87	Oui
<b>Total</b>	<b>1049</b>	<b>3845.94</b>	-

## Tests et résultats

- Fréquence d'échantillonnage des fichiers audio : 16KHz
- Coefficients acoustiques utilisés : MFCC
- Taille de fenêtre : 16ms
- La base de données est divisée en trois parties :
  - Une partie pour créer les modèles GMM et le modèle UBM
  - Une partie pour créer les Super Vecteurs
  - Une partie pour les tests

# Tests et résultats

- Fréquence d'échantillonnage des fichiers audio : 16KHz
- Coefficients acoustiques utilisés : MFCC
- Taille de fenêtre : 16ms
- La base de données est divisée en trois parties :
  - Une partie pour créer les modèles GMM et le modèle UBM
  - Une partie pour créer les Super Vecteurs
  - Une partie pour les tests

## Résultats (comparaison entre GMM et SVM-GSL) :

Méthode	Taux de Reconnaissance (%)
GMM	71.1
SVM-GSL (taille UBM = 256)	74.0
SVM-GSL (taille UBM = 512)	<b>75.4</b>
SVM-GSL (taille UBM = 1024)	75.1

# Tests et résultats

## Résultats par classe de sons :

Classe de sons	Taux de bonnes reconnaissances
Breathing	9/17
Cough	15/21
Dishes	14/33
DoorClapping	36/38
<b>DoorOpening</b>	<b>7/7</b>
<b>ElectricalShaver</b>	<b>20/21</b>
FemaleCry	8/12
FemaleScream	19/24
GlassBreaking	25/34
<b>HairDryer</b>	<b>14/14</b>
HandsClapping	17/18
Keys	10/12
Laugh	2/17
MaleScream	27/29
Paper	19/21
Sneeze	10/11
Water	12/18
Yawn	3/7

Les classes présentant peu de variabilité sont très bien reconnues

# Tests et résultats

## Résultats par classe de sons :

Classe de sons	Taux de bonnes reconnaissances
Breathing	9/17
Cough	15/21
Dishes	14/33
<b>DoorClapping</b>	<b>36/38</b>
DoorOpening	7/7
ElectricalShaver	20/21
FemaleCry	8/12
FemaleScream	19/24
GlassBreaking	25/34
HairDryer	14/14
<b>HandsClapping</b>	<b>17/18</b>
Keys	<b>10/12</b>
Laugh	2/17
<b>MaleScream</b>	<b>27/29</b>
<b>Paper</b>	<b>19/21</b>
<b>Sneeze</b>	<b>10/11</b>
Water	12/18
Yawn	3/7

Certaines classes présentant des variabilités sont également très bien reconnues



# Tests et résultats

## Résultats par classe de sons :

Classe de sons	Taux de bonnes reconnaissances
Breathing	9/17
Cough	15/21
Dishes	14/33
DoorClapping	36/38
DoorOpening	7/7
ElectricalShaver	20/21
FemaleCry	8/12
FemaleScream	19/24
GlassBreaking	25/34
HairDryer	14/14
HandsClapping	17/18
Keys	10/12
Laugh	2/17
MaleScream	27/29
Paper	19/21
Sneeze	10/11
Water	12/18
Yawn	3/7

Certaines classes restent difficiles à reconnaître

## Tests et résultats

Il existe des classes sons très mal séparées par les coefficients MFCC quel que soit le classifieur.

### **Notre contribution originale :**

Définir des familles de coefficients variées pour la classification du son.

# Plan

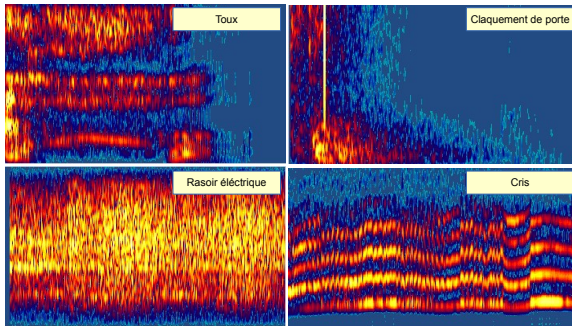
- 1 Contexte : Projet Sweet-Home
- 2 Objectifs de la thèse
- 3 État de l'art en classification des sons de l'environnement
- 4 Première contribution : Mise en œuvre de différents classifieurs issus de la RAL, pour la classification des sons
- 5 Deuxième contribution : Utilisation de représentations adaptées aux différentes classes de sons**
- 6 Conclusions et Perspectives

# Différences entres classes de sons : un aperçu

A l'écoute, certains types de sons sont très différents

# Différences entre classes de sons : un aperçu

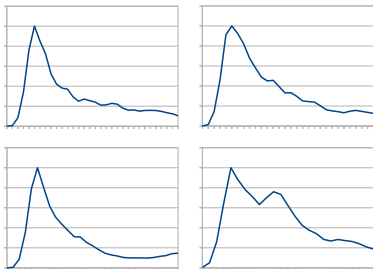
A l'écoute, certains types de sons sont très différents  
Les différences au niveau du signal peuvent aussi être importantes



**Figure:** Spectrogrammes de 4 enregistrements de 4 classes différentes

# Différences entre classes de sons : un aperçu

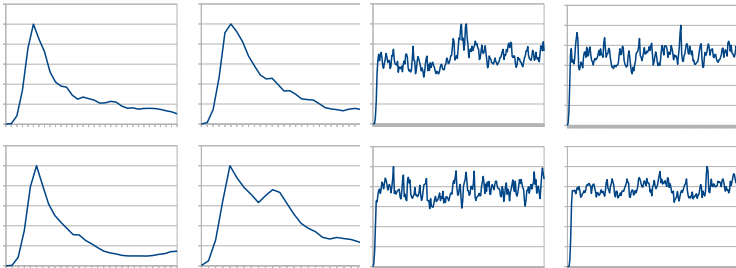
Enveloppe d'amplitude pour des enregistrements de **bris de glace**



# Différences entres classes de sons : un aperçu

Enveloppe d'amplitude pour des enregistrements de **bris de glace**

Enveloppe d'amplitude pour des enregistrements d'un **sèche cheveux**



# Aller au-delà des *MFCC* : utiliser d'autres familles de coefficients acoustiques

## Motivations :

- Améliorer les performances : extraire des caractéristiques du signal non représentées par les *MFCC*
- Opter pour des familles moins complexes si elles permettent d'obtenir les mêmes performances



# Aller au-delà des MFCC : utiliser d'autres familles de coefficients acoustiques

## Motivations :

- Améliorer les performances : extraire des caractéristiques du signal non représentées par les MFCC
- Opter pour des familles moins complexes si elles permettent d'obtenir les mêmes performances

## Approche :

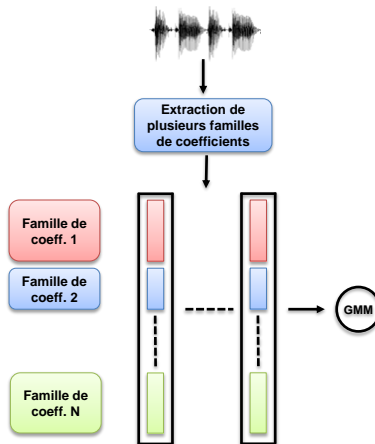
- Utiliser des familles de coefficients acoustiques différentes
- Utiliser le savoir faire acquis en matière de méthodes de reconnaissance (GMM, SVM)

# Familles de coefficients retenues

<b>Famille de coefficients</b>	<b>Nombre de caractéristiques</b>
<i>MFCC</i>	16
<i>Loudness</i>	24
<i>Spectral Crest Factor Per Band</i>	19
<i>Spectral Flatness Per Band</i>	19
<i>Complex Domain Onset Detection</i>	1
<i>Perceptual Sharpness</i>	1
<i>Perceptual Spread</i>	1
<i>Spectral Roll-Off</i>	1
<i>Spectral Decrease</i>	1
<i>Spectral Flatness</i>	1
<i>Spectral Variation</i>	1
<i>Spectral Slope</i>	1
<i>Spectral Shape Statistics</i>	4
<i>Temporal Shape Statistics</i>	4
Total	94

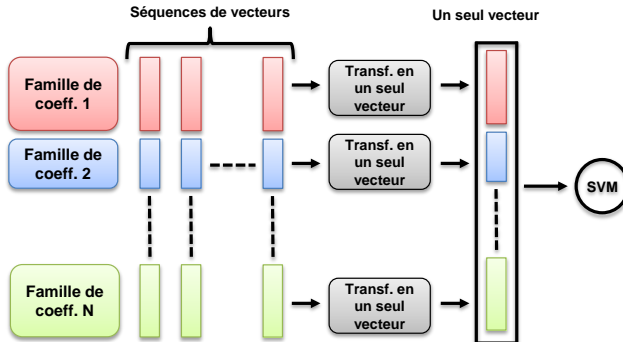
# Utilisation de plusieurs familles avec des GMMs

Utiliser des vecteurs contenant toutes les familles de coefficients :



# Utilisation de plusieurs familles avec des SVMs

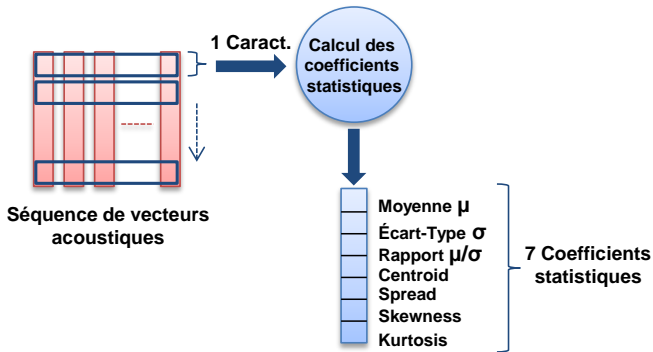
Une transformation des données est nécessaire...



Deux méthodes de transformations proposées : **SVM-StatVect** et **SVM-ProbVect**

# SVM-StatVect

Transformer une séquence de vecteurs acoustiques en un un seul vecteur de coefficients statistiques

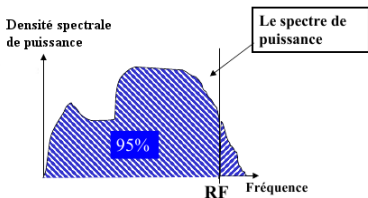


# SVM-ProbVect

Comment exploiter les familles de coefficients à une seule caractéristique ?

## Spectral Roll-Off :

- La fréquence au-dessous de laquelle se trouve 95% du signal
- Une seule valeur par fenêtre de signal



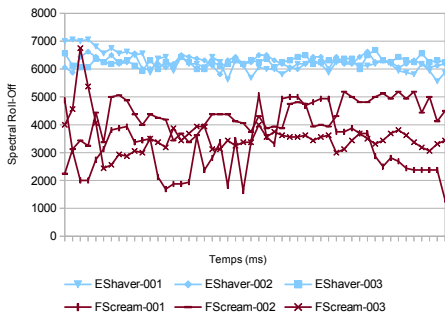
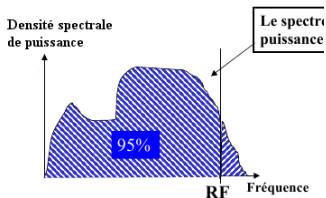
# SVM-ProbVect

Comment exploiter les familles de coefficients à une seule caractéristique ?

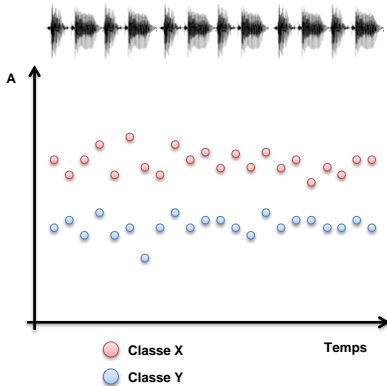
## Spectral Roll-Off :

- La fréquence au-dessous de laquelle se trouve 95% du signal
- Une seule valeur par fenêtre de signal

**Exemple :** variation dans le temps du Spectral Roll-Off pour deux classes de sons



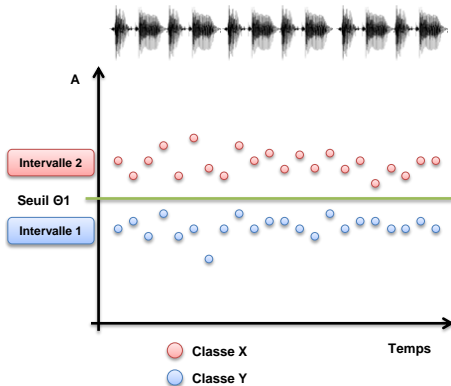
# SVM-ProbVect



Deux classes peuvent être aisément séparables

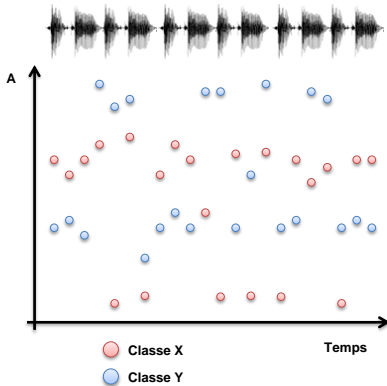


# SVM-ProbVect



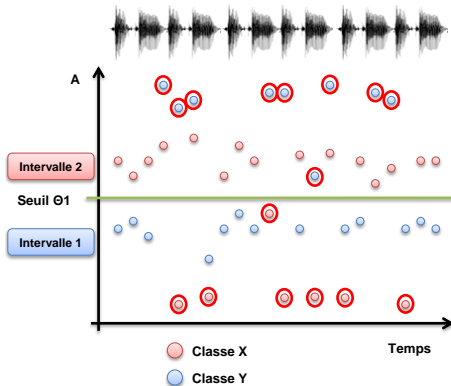
Méthode simple de séparation : trouver un **seuil** et définir deux **intervalles**

# SVM-ProbVect



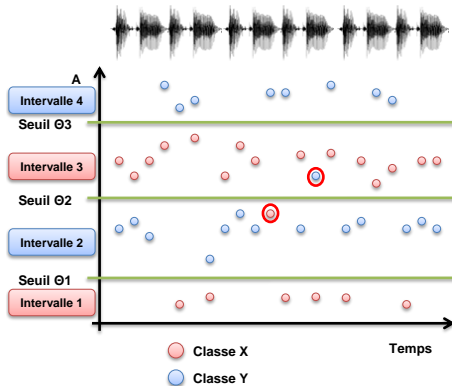
La séparation entre les deux classes peut être plus difficile

# SVM-ProbVect



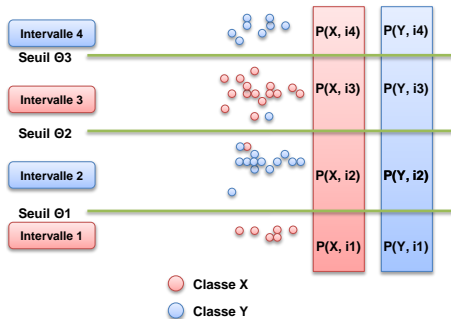
Un seul seuil peut générer beaucoup d'erreurs de classification

# SVM-ProbVect



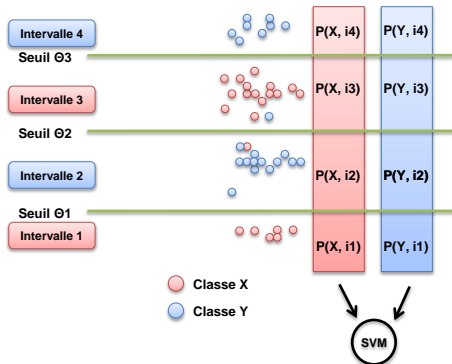
**Une solution** : trouver **plusieurs** seuils pour **minimiser** les erreurs de classification

# SVM-ProbVect



**SVM-ProbVect** : une classe est représentée par un vecteur contenant la **probabilité** de la classe dans **chaque intervalle**

# SVM-ProbVect



**SVM-ProbVect** : les vecteurs de probabilités sont utilisés en entrée d'un SVM

# Comment trouver les intervalles ?

En utilisant des méthodes de **discrétisation** supervisées :

# Comment trouver les intervalles ?

En utilisant des méthodes de **discrétisation** supervisées :

- Méthodes utilisées pour transformer des valeurs continues en valeurs discrètes
- Peuvent être utilisées pour créer des arbres de décision
- Permettent de trouver les seuils qui donnent une bonne corrélation entre un intervalle donné et une classe particulière



# Comment trouver les intervalles ?

En utilisant des méthodes de **discrétisation** supervisées :

- Méthodes utilisées pour transformer des valeurs continues en valeurs discrètes
- Peuvent être utilisées pour créer des arbres de décision
- Permettent de trouver les seuils qui donnent une bonne corrélation entre un intervalle donné et une classe particulière

Méthode retenue

- Méthode de Fayyad et Irani (1993)
  - Méthode basée sur la minimisation de l'entropie
  - Bonnes performances pour plusieurs problèmes de classification [Dougherty 95, Liu Huan 2002]

# Résultats : comparaison entre GMM et SVM-StatVect

Famille de Coefficients	GMM	SVM-StatVect
<i>MFCC</i>	<b>75.9</b>	63.2
<i>Loudness</i>	<b>70.3</b>	66.9
<i>Spectral Crest F. P. B.</i>	<b>67.7</b>	66.6
<i>Spectral Flatness P. B.</i>	64.1	<b>65.5</b>
<i>Spectral Shape Stat.</i>	46.8	<b>55.3</b>
Toutes les Familles	72.3	<b>76.5</b>

- **GMM** : une baisse en performances est observée si toutes les familles de coefficients sont fusionnées
- **SVM-StatVect** : Les meilleures performances sont utilisées lorsque toutes les familles sont fusionnées

# Comparaison de toutes les méthodes

## En utilisant toutes les familles de coefficients ensemble :

- GMM : **72.3%**
- SVM-StatVect : **76.5%**
- SVM-ProbVect : **76.2%**
- Fusion entre SVM-StatVect et SVM-ProbVect : **79.0%**

# Comparaison de toutes les méthodes

## En utilisant toutes les familles de coefficients ensemble :

- GMM : **72.3%**
- SVM-StatVect : **76.5%**
- SVM-ProbVect : **76.2%**
- Fusion entre SVM-StatVect et SVM-ProbVect : **79.0%**

## Conclusions :

- Les méthodes SVM-StatVect et SVM-ProbVect sont mieux adaptées à l'utilisation de plusieurs familles de coefficients ensemble
- La fusion des deux méthodes permet d'améliorer les performances

# Résultats expérimentaux : SVM-StatVect

## Performances par famille de coefficients

<b>Famille de Coefficients</b>	<b>Dimension</b>	<b>Taux de Reco. (%)</b>
<i>Loudness</i>	168	66.9
<i>Spectral Crest F. P. B.</i>	133	66.6
<i>Spectral Flatness P. B.</i>	133	65.5
<i>MFCC</i>	112	63.27
<i>Spectral Shape Stat.</i>	12	55.3
<i>Spectral Roll-Off</i>	7	48.0
<i>Spectral Flatness</i>	7	48.0
<i>Spectral Slope</i>	7	45.4
<i>Complex D. O. D.</i>	7	44.6
<i>Spectral Variation</i>	7	43.7
<i>Perceptual Sharpness</i>	7	43.5
<i>Perceptual Spread</i>	7	41.8
<i>Envelope Shape Stat.</i>	7	38.1
<i>Spectral Decrease</i>	7	36.7
<i>Temporal Shape Stat.</i>	4	33.3
<b>Tous les Coefficients</b>	<b>625</b>	<b>76.5</b>

## Performances des familles de coefficients pour 3 classes

Certaines classes de sons sont mieux reconnues avec des familles de coefficients spécifiques

Familles de coefficients	Classes de sons		
	Door Clap.	Keys	Water
MFCC	0.9	0.8	0.4
Loudness	0.9	<b>1</b>	0.6
Spectral Crest Factor Per Band	0.9	0.9	0.5
Spectral Flatness Per Band	0.9	0.7	0.6
Complex Domain Onset Detection	0.7	0	0.7
Perceptual Sharpness	0.9	0.08	0.6
Perceptual Spread	0.8	0	0.9
Spectral Roll-Off	0.7	0	0.5
Spectral Decrease	0.7	0	<b>0.9</b>
Spectral Flatness	0.7	0.7	0.4
Spectral Variation	0.7	0	0.3
Spectral Slope	<b>1</b>	0	<b>0.9</b>
Spectral Shape Statistics	0.9	0.8	0.2
Temporal Shape Statistics	0.7	0	<b>0.9</b>
Envelope Shape Stat.	0.6	0	0.7
Tous les coefficients	<b>1</b>	0.9	0.6

# Sélection de caractéristiques

## Méthode de sélection : SVM Wrapper [Guyon et al, 2002]

- Sélection itérative des caractéristiques les moins significatives d'un modèle SVM
- Un apprentissage et une évaluation sont nécessaires à chaque itération
- Peut être coûteuse en temps de calcul

# Sélection de caractéristiques

## Méthode de sélection : SVM Wrapper [Guyon et al, 2002]

- Sélection itérative des caractéristiques les moins significatives d'un modèle SVM
- Un apprentissage et une évaluation sont nécessaires à chaque itération
- Peut être coûteuse en temps de calcul

## Variante utilisée :

- Le taux de caractéristiques à éliminer est figé
- Une seule sélection est faite



# Résultats de la sélection de caractéristiques

Taux de sél. (%)	Provenance des coefficients		
	SVM-StatVect	SVM-ProbVect	SVM-StatVect + SVM-ProbVect
100	76.5	76.2	<b>79.0</b>
50	76.5	<b>76.5</b>	77.6
20	77.9	75.1	<b>79.0</b>
15	<b>78.5</b>	72.5	78.5
10	75.9	73.4	77.6

Un nombre plus réduit de caractéristiques permet de maintenir ou d'améliorer les performances

# Tests sur le corpus du projet Sweet-Home :

## Description du corpus :

- Enregistrement des activités de la vie quotidienne dans une maison intelligente (appartement Domus, Grenoble)

# Tests sur le corpus du projet Sweet-Home :

## Description du corpus :

- Enregistrement des activités de la vie quotidienne dans une maison intelligente (appartement Domus, Grenoble)
- Sept microphones installés dans l'appartement
- durée d'un enregistrement : une heure environ
- Nombre de sujets : 21

# Tests sur le corpus du projet Sweet-Home :

## Description du corpus :

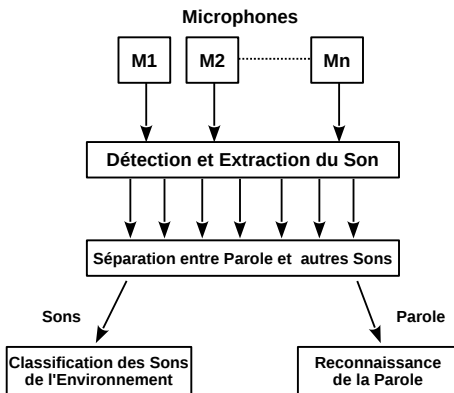
- Enregistrement des activités de la vie quotidienne dans une maison intelligente (appartement Domus, Grenoble)
- Sept microphones installés dans l'appartement
- durée d'un enregistrement : une heure environ
- Nombre de sujets : 21
- Exemples d'activités :
  - Parler au téléphone
  - Préparer une boisson chaude
  - Utiliser l'aspirateur
  - Se brosser les dents
  - Etc.

# Tests sur le corpus du projet Sweet-Home :

## Objectif des expérimentations :

- Évaluer les différentes composantes du système

## Architecture du système :



# Tests sur le corpus du projet Sweet-Home :

## Protocole de test :

- **Trois** scénarios utilisés en apprentissage
- **Cinq** scénarios utilisés pour évaluer la séparation entre la parole et les autres sons
- Dont **quatre** utilisés pour évaluer la reconnaissance des sons de l'environnement

# Tests sur le corpus du projet Sweet-Home :

## Protocole de test :

- **Trois** scénarios utilisés en apprentissage
- **Cinq** scénarios utilisés pour évaluer la séparation entre la parole et les autres sons
- Dont **quatre** utilisés pour évaluer la reconnaissance des sons de l'environnement
- **18** classes de sons prises en considération :
  - Ouverture/fermeture de portes
  - Aspirateur
  - Écoulement d'eau
  - Etc.

# Tests sur le corpus du projet Sweet-Home :

## Protocole de test :

- **Trois** scénarios utilisés en apprentissage
- **Cinq** scénarios utilisés pour évaluer la séparation entre la parole et les autres sons
- Dont **quatre** utilisés pour évaluer la reconnaissance des sons de l'environnement
- **18** classes de sons prises en considération :
  - Ouverture/fermeture de portes
  - Aspirateur
  - Écoulement d'eau
  - Etc.
- Coefficients acoustiques : **MFCC**
- Méthode de reconnaissance : **SVM-GSL**



# Tests sur le corpus du projet Sweet-Home :

## Performances de la détection du son :

$\tau$ (%)	20	50	80	100
<b>Sensibilité (%)</b>	96.1	93.4	90.3	88.6

Sensibilité de détection pour différentes valeurs du taux de recouvrement ( $\tau$ ) avec les références

# Tests sur le corpus du projet Sweet-Home :

## Performances de la séparation entre la parole et les autres sons :

Sujet	# de phrases	Canal	FN Det.	FN Reco.	TP D+R
S01	44	C6	0	4	40
		C7	1	2	41

# Tests sur le corpus du projet Sweet-Home :

## Performances de la séparation entre la parole et les autres sons :

Sujet	# de phrases	Canal	FN Det.	FN Reco.	TP D+R
S01	44	C6	0	4	40
		C7	1	2	41
S03	41	C6	0	2	39
		C7	1	7	33
S07	45	C6	4	5	36
		C7	6	3	36
S09	40	C6	0	0	40
		C7	0	0	40
S13	42	C6	0	4	38
		C7	1	0	41

# Tests sur le corpus du projet Sweet-Home :

## Reconnaissance des sons de la vie courante par canal :

Canal	C1	C2	C3	C4	C5	C6	C7	Fusion de tous les canaux
<b>Moy. TP D+R(%)</b>	31.3	33.3	14.9	24.1	21.8	13.3	9.6	<b>59.1</b>

# Plan

- 1 Contexte : Projet Sweet-Home
- 2 Objectifs de la thèse
- 3 État de l'art en classification des sons de l'environnement
- 4 Première contribution : Mise en œuvre de différents classifieurs issus de la RAL, pour la classification des sons
- 5 Deuxième contribution : Utilisation de représentations adaptées aux différentes classes de sons
- 6 Conclusions et Perspectives**

# Conclusions :

- La reconnaissance des sons de l'environnement est un domaine de recherche assez récent : il existe très peu de travaux

# Conclusions :

- La reconnaissance des sons de l'environnement est un domaine de recherche assez récent : il existe très peu de travaux
- Les méthodes utilisées en reconnaissance du locuteur fonctionnent visiblement bien

# Conclusions :

- La reconnaissance des sons de l'environnement est un domaine de recherche assez récent : il existe très peu de travaux
- Les méthodes utilisées en reconnaissance du locuteur fonctionnent visiblement bien
- SVM-StatVect et SVM-ProbVect s'avèrent efficaces pour l'utilisation de plusieurs familles de coefficients ensemble en entrée d'un SVM



# Conclusions :

- La reconnaissance des sons de l'environnement est un domaine de recherche assez récent : il existe très peu de travaux
- Les méthodes utilisées en reconnaissance du locuteur fonctionnent visiblement bien
- SVM-StatVect et SVM-ProbVect s'avèrent efficaces pour l'utilisation de plusieurs familles de coefficients ensemble en entrée d'un SVM
- Certaines classes de sons sont reconnues plus efficacement avec des familles de coefficients spécifiques

# Conclusions :

- La reconnaissance des sons de l'environnement est un domaine de recherche assez récent : il existe très peu de travaux
- Les méthodes utilisées en reconnaissance du locuteur fonctionnent visiblement bien
- SVM-StatVect et SVM-ProbVect s'avèrent efficaces pour l'utilisation de plusieurs familles de coefficients ensemble en entrée d'un SVM
- Certaines classes de sons sont reconnues plus efficacement avec des familles de coefficients spécifiques

# Perspectives :

- Enrichir la base de données

# Perspectives :

- Enrichir la base de données
- Explorer les autres avancées en reconnaissance du locuteur

# Perspectives :

- Enrichir la base de données
- Explorer les autres avancées en reconnaissance du locuteur
- Utiliser d'autres familles de coefficients acoustiques

# Perspectives :

- Enrichir la base de données
- Explorer les autres avancées en reconnaissance du locuteur
- Utiliser d'autres familles de coefficients acoustiques
- Inclure d'autres sources d'information (dépendance temporelle)

# Perspectives :

- Enrichir la base de données
- Explorer les autres avancées en reconnaissance du locuteur
- Utiliser d'autres familles de coefficients acoustiques
- Inclure d'autres sources d'information (dépendance temporelle)
- Utiliser d'autres méthodes de fusion et de sélection pour une meilleure exploitation des familles de coefficients