The use of kernel methods for audio events detection

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

Alissar Nasser, Denis Hamad, Jean-Luc Rouas, Sébastien Ambellouis

INRETS

The French national institute for transport and safety research

Outline

1 Introduction

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

INRETS

The French national institute for transport and safety research

Outline

1 Introduction

2 Experimental Data

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

INRETS

The French national institute for transport and safety research

Outline

1 Introduction

2 Experimental Data

3 Processing
- Visualisation
- Feature reduction
- Classification step of reduced data

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

Outline

1 Introduction

2 Experimental Data

3 Processing
- Visualisation
- Feature reduction
- Classification step of reduced data

4 Conclusion & Perspectives

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

Outline

1 Introduction

2 Experimental Data

3 Processing
- Visualisation
- Feature reduction
- Classification step of reduced data

4 Conclusion & Perspectives

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

Introduction

Audio Events Classification

- Usually $\Rightarrow$ Supervised techniques
- But: requires hand labelling...

$\Rightarrow$ Use unsupervised Kernel methods:

- dimensionality reduction + visualisation
- unsupervised data clustering

Introduction
Experimental Data
Processing
Visualisation
Feature reduction
Classification step of reduced data
Conclusion & Perspectives

Outline

1 Introduction

2 Experimental Data

3 Processing
- Visualisation
- Feature reduction
- Classification step of reduced data

4 Conclusion & Perspectives

Experiments

Context (SAMSIT project)

- Public transport application

$\Rightarrow$ Detecting abnormal audio events

Three sound classes

- Speech
- Spray Bomb
- Ambient noise

Introduction
Experimental Data
Processing
Visualisation
Feature reduction
Classification step of reduced data
Conclusion & Perspectives

Test Signal used

Total: 2338 samples

- 605 speech samples,
- 1273 noise samples,
- 460 spray bomb samples.

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

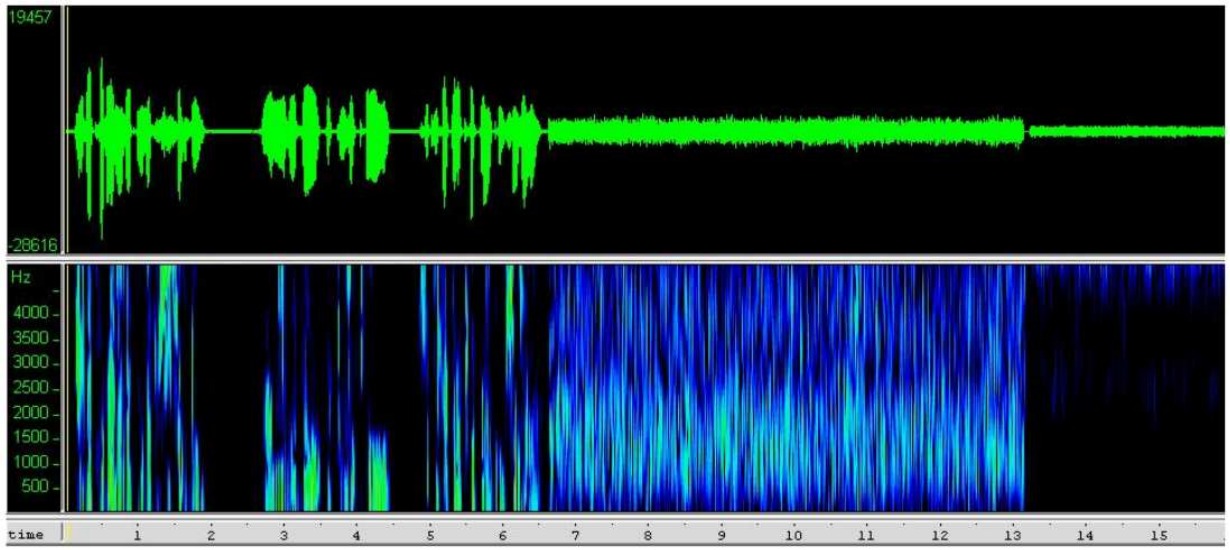

Figure: Spectrum of the signal used

Outline

1 Introduction

2 Experimental Data

3 Processing
- Visualisation
- Feature reduction
- Classification step of reduced data

4 Conclusion & Perspectives

Approach

Introduction
Experimental Data
Processing
Visualisation
Feature reduction
Classification step of reduced data
Conclusion & Perspectives

1. Pre-processing: MFCC feature extraction
2. Feature reduction & visualisation: PCA & Kernel PCA
3. Classification: Kernel K-means

Visualisation

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

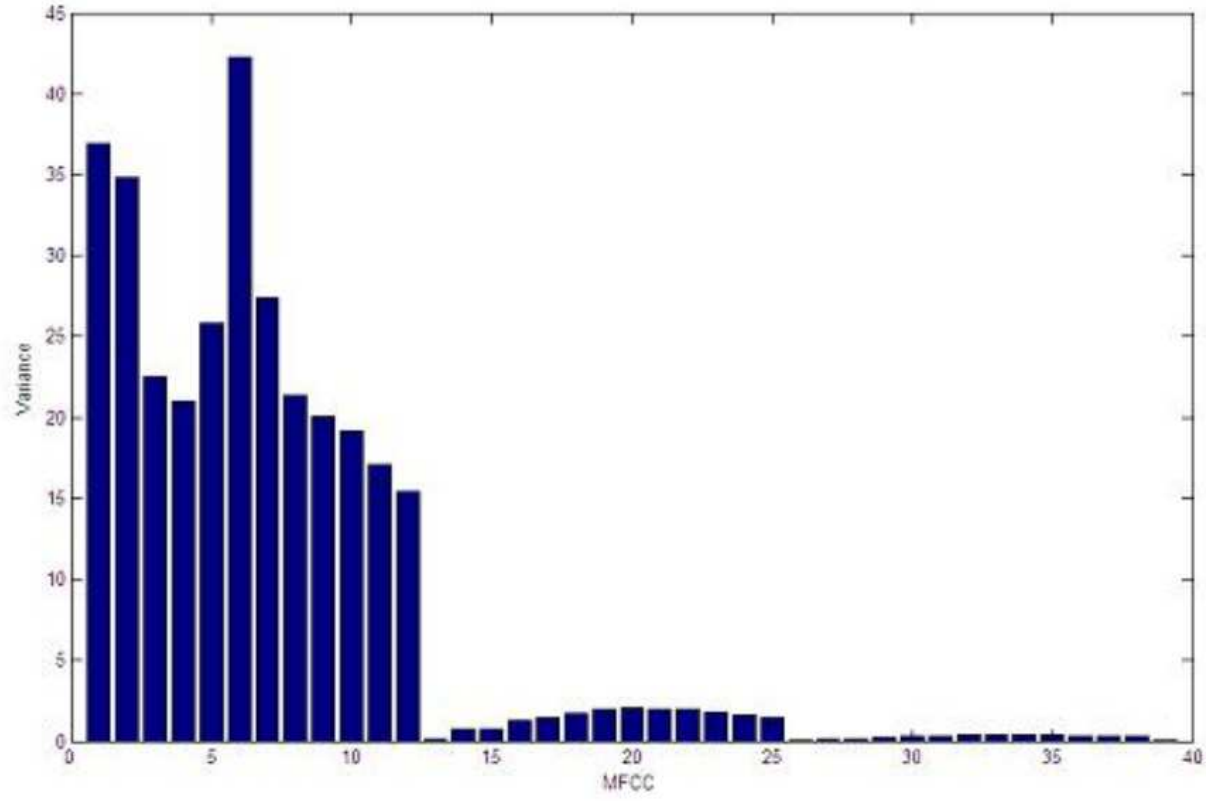

Figure: Variance representation of the 39 MFCC features

Visualisation

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

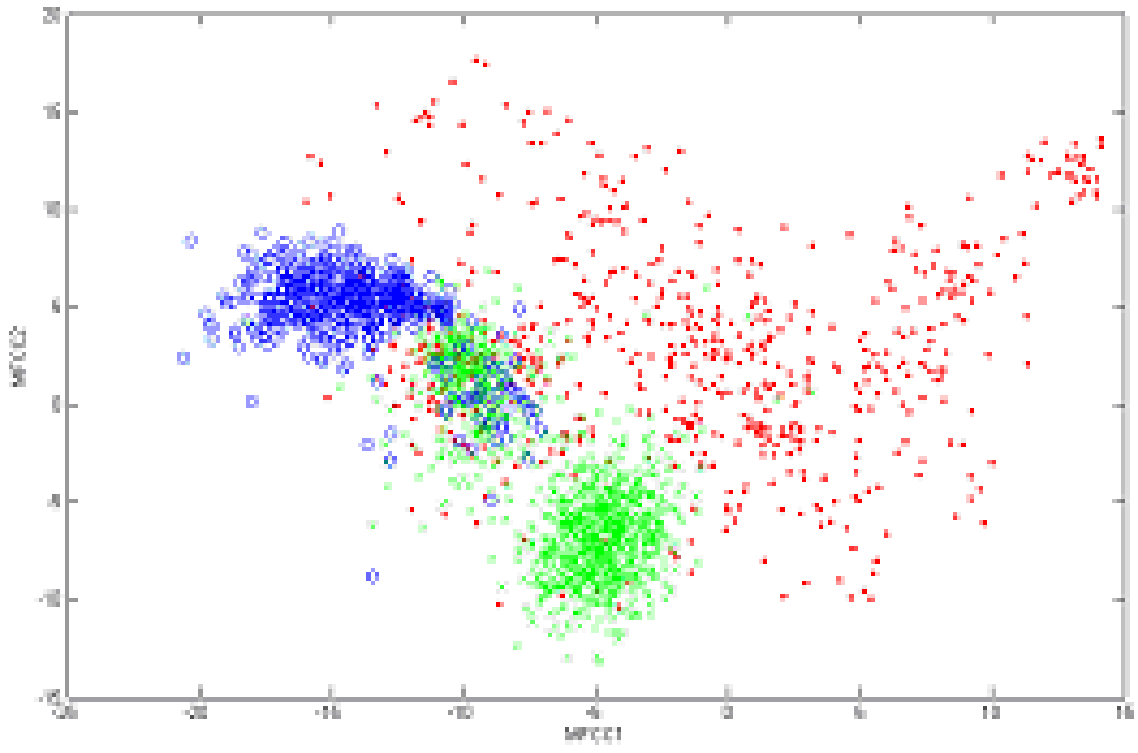

Figure: Scatter plot on the 1st and 6th MFCC.

Visualisation

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

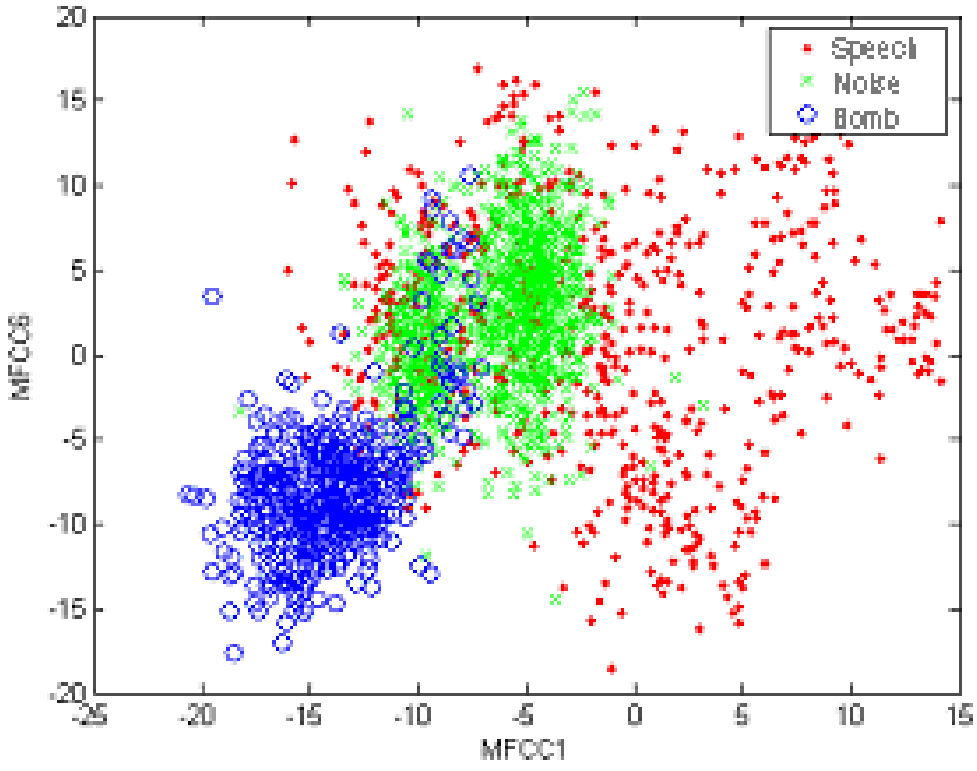

Figure: Data scatter plot on the 1st and 2nd MFCC

MFCC features reduction

- Select only the 12 principal components given from both PCA and kernel PCA

$\Rightarrow$ 93.69% information

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

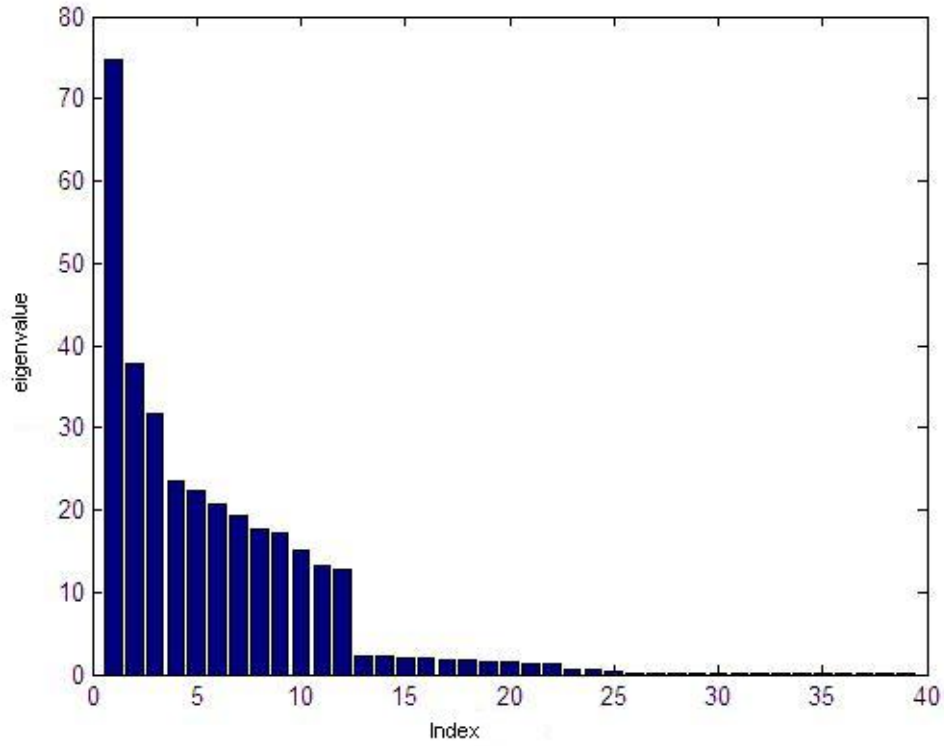

Figure: KL-transform eigenvalues in decreasing order

MFCC features reduction

Introduction

Experimental Data

Processing

Visualisation

Feature reduction

Classification step of reduced data

Conclusion & Perspectives

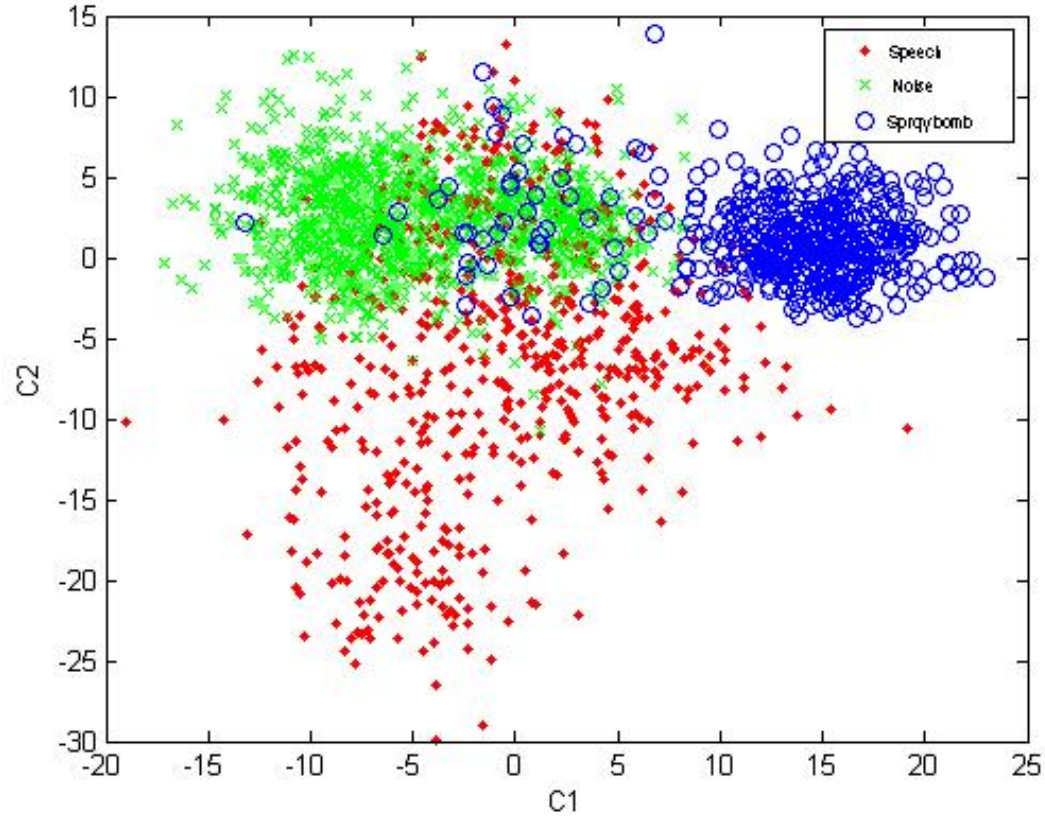

Figure: Real classes on the 1st and 2nd PCA components.

Classification step of reduced data - PCA

Introduction
Experimental Data
Processing
Visualisation
Feature reduction
Classification step of reduced data
Conclusion & Perspectives

Results → ↓ Truth	Speech	Noise	Spray
Speech	69.42	27.44	3.14
Noise	0.55	97.25	2.2
Spray	0.00	8.48	91.52

Table: Confusion matrix of k-means clustering on 12 PCA components

Classification step of reduced data - Kernel PCA

Introduction
Experimental Data
Processing
Visualisation
Feature reduction
Classification step of reduced data
Conclusion & Perspectives

Results → ↓ Truth	Speech	Noise	Spray
Speech	77.85	21.65	0.50
Noise	0.79	97.72	1.49
Spray	0.00	9.13	90.87

Table: Confusion matrix of k-means clustering on 12 Kernel PCA components

Kernel K-means classification (after kernel PCA)

Introduction
Experimental Data
Processing
Visualisation
Feature reduction
Classification step of reduced data
Conclusion & Perspectives

Results → ↓ Truth	Speech	Noise	Spray
Speech	77.02	22.48	0.50
Noise	0.63	98.43	0.94
Spray	0.00	9.35	90.65

Table: Confusion matrix of Kernel k-means clustering

Outline

1 Introduction

2 Experimental Data

3 Processing
- Visualisation
- Feature reduction
- Classification step of reduced data

4 Conclusion & Perspectives

Conclusion & Perspectives

Introduction
Experimental Data
Processing
Visualisation
Feature reduction
Classification step of reduced data
Conclusion & Perspectives

Conclusions

- Automatic and non supervised audio event detection
- Classes are better separated thanks to kernel PCA
- Kernel PCA combined with kernel K-means have showed slightly better performances

Perspectives

- Few annotation data is available

$\Rightarrow$ Use this data in a semi-supervised approach