Word recognition

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

POLYTECH
NICE-SOPHIA

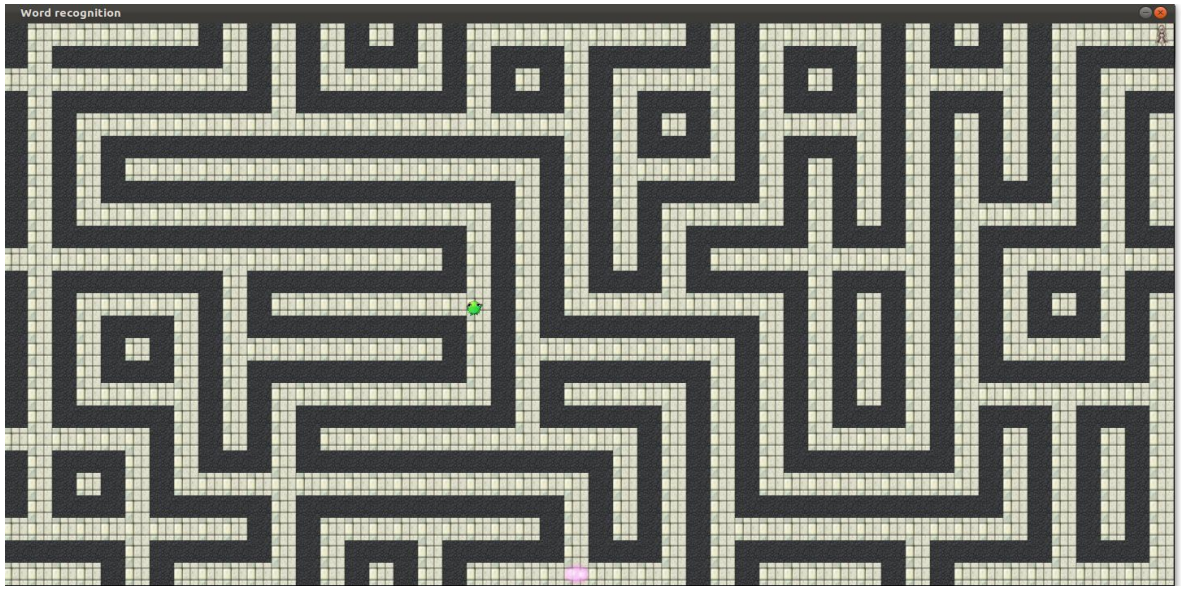
MAM4 G1

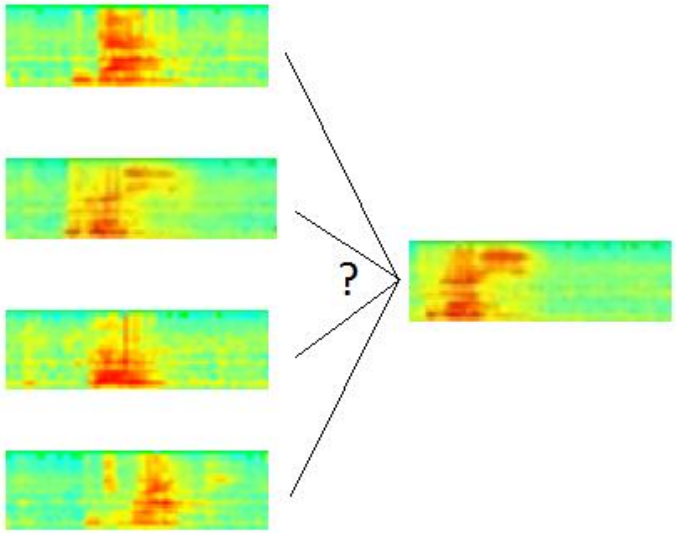
Introduction

- Home automation system
- Disable persons
- Games

Introduction

- Game by voice control



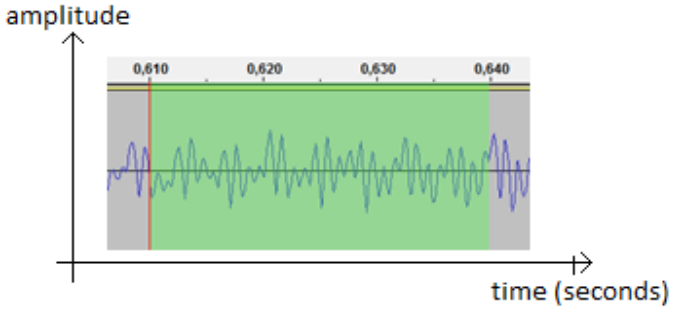


Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Contents

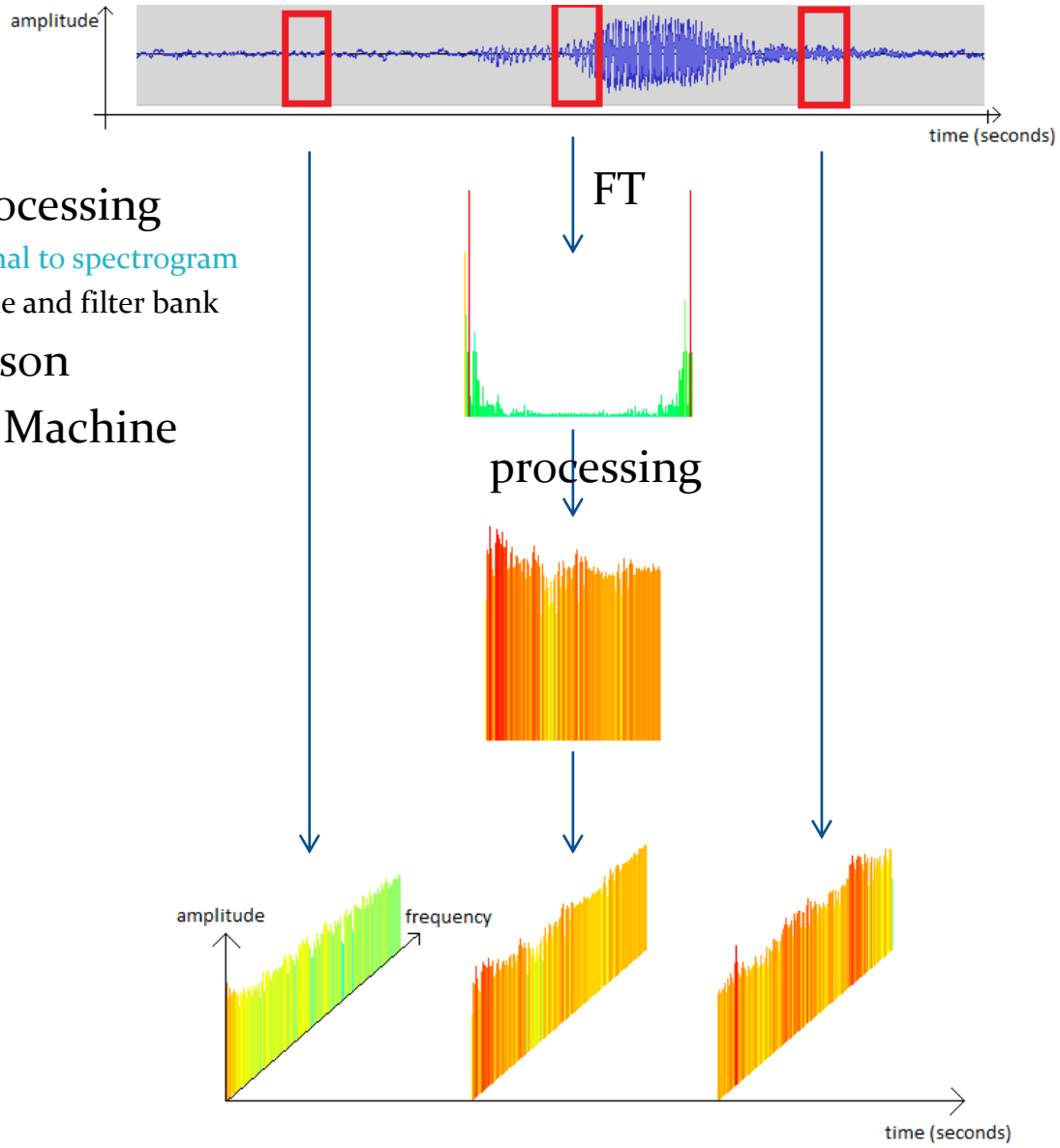
Sound processing

- Sound recording
- Time split
- Treatment of each slice



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

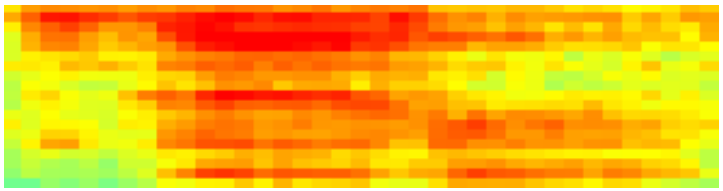
From signal to spectrogram



I. Sound processing
   a. From signal to spectrogram
   b. Mel's scale and filter bank
II. Comparison
III. Human Machine Interface

Spectrogram

- 2D view spectrogram



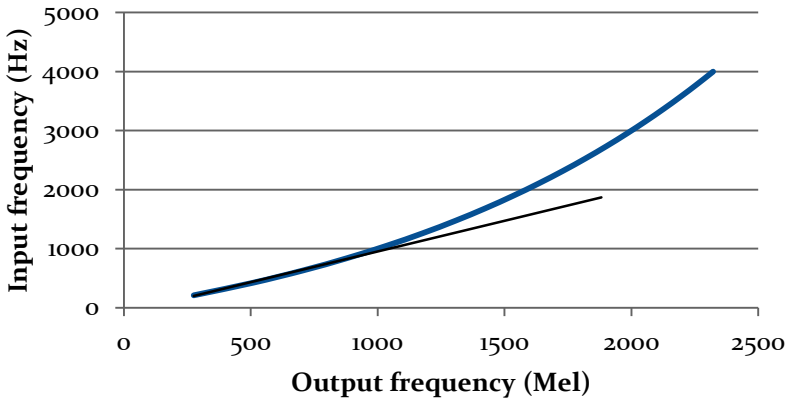
« gauche »

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Mel's scale and filter bank

- Mel's scale:
  - Reduce the importance of high frequencies
  - From Hz to Mel

Mel's scale

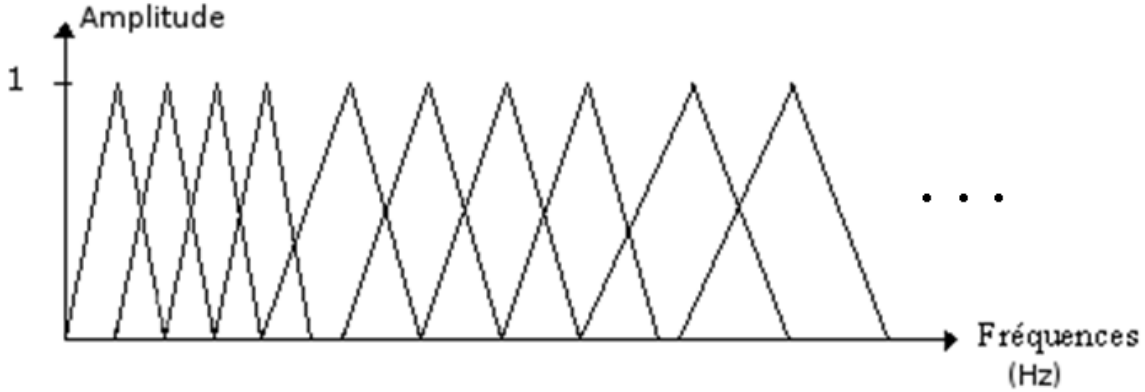


$$M = \frac{1000}{\log(2)} * \log\left(1 + \frac{F}{1000}\right)$$

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Mel's scale and filter bank

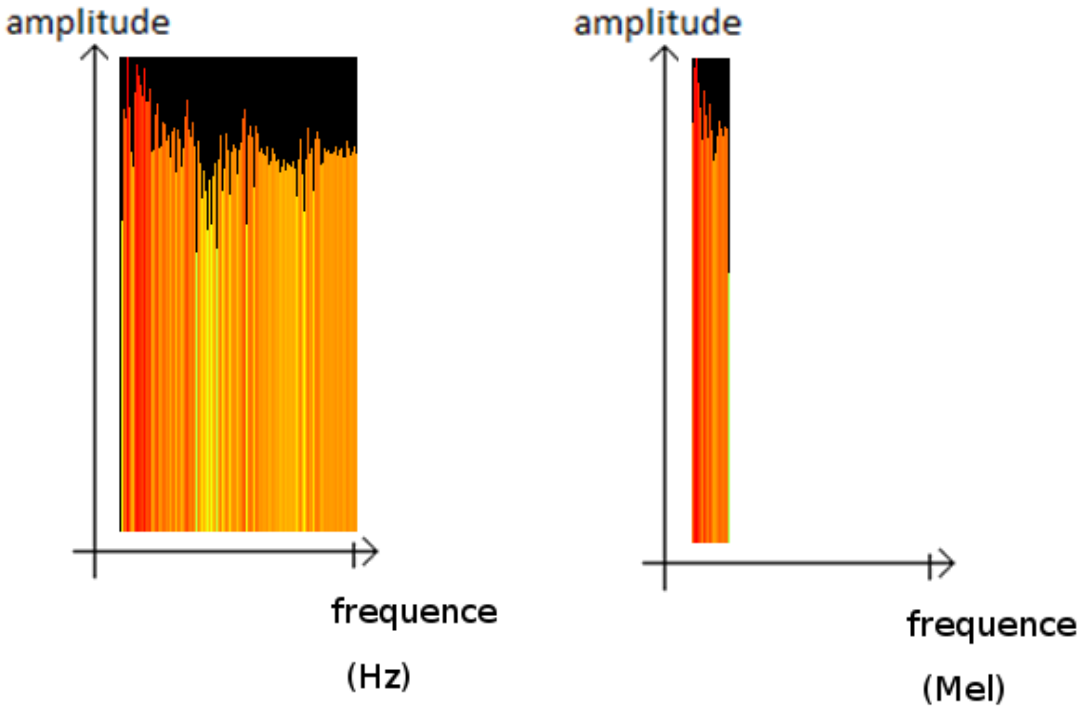
- Filter bank:
  - Reduce the number of frequencies considered
  - Uniform on the Mel's scale, non uniformly on the frequency scale



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

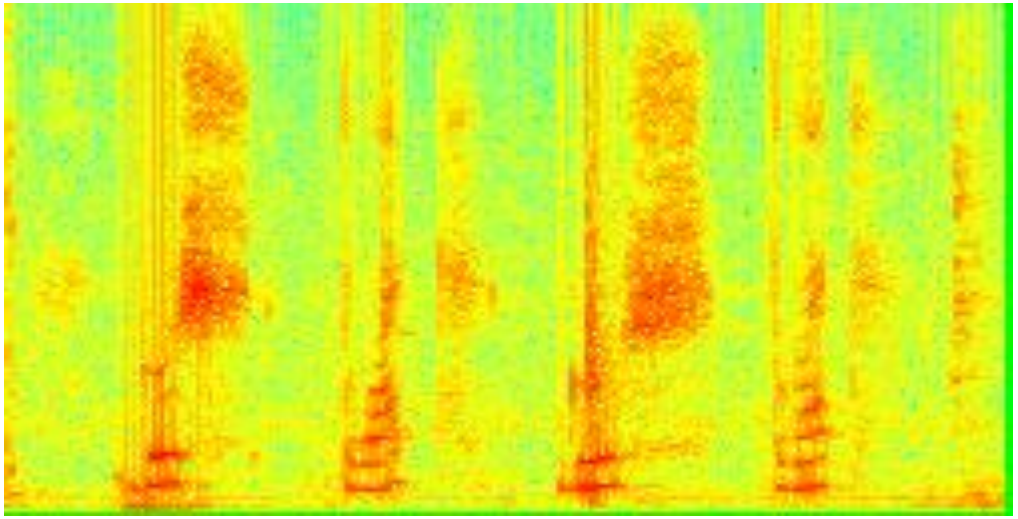
Mel's scale and filter bank

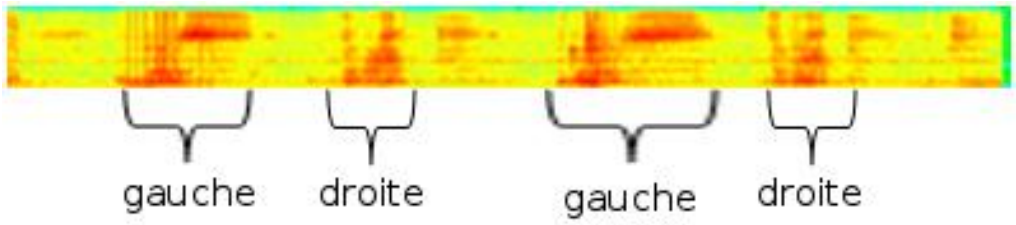
- Example:



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Mel's scale and filter bank





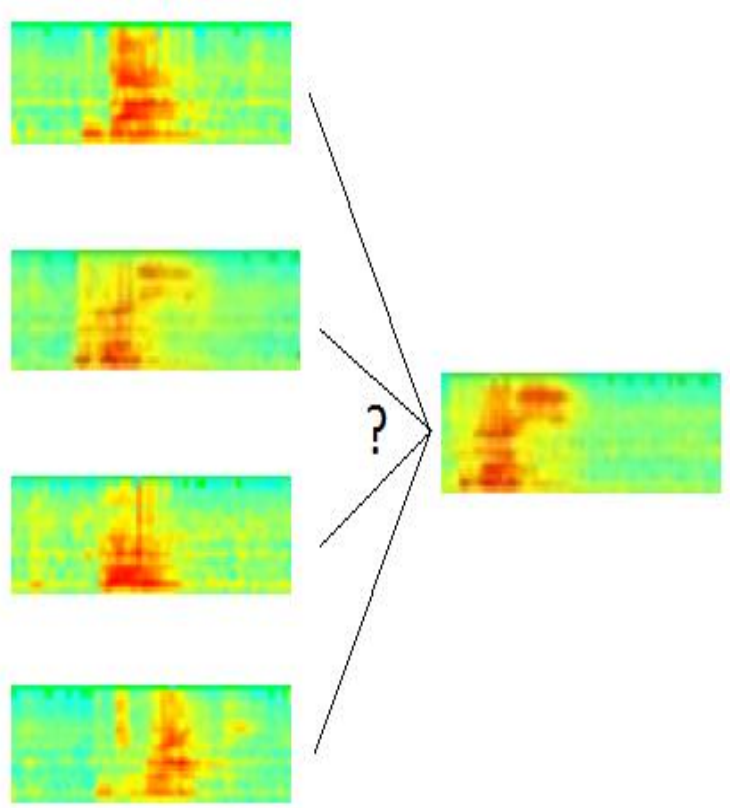
Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

MAM4

POLYTECH NICE-SOPHIA

Comparison

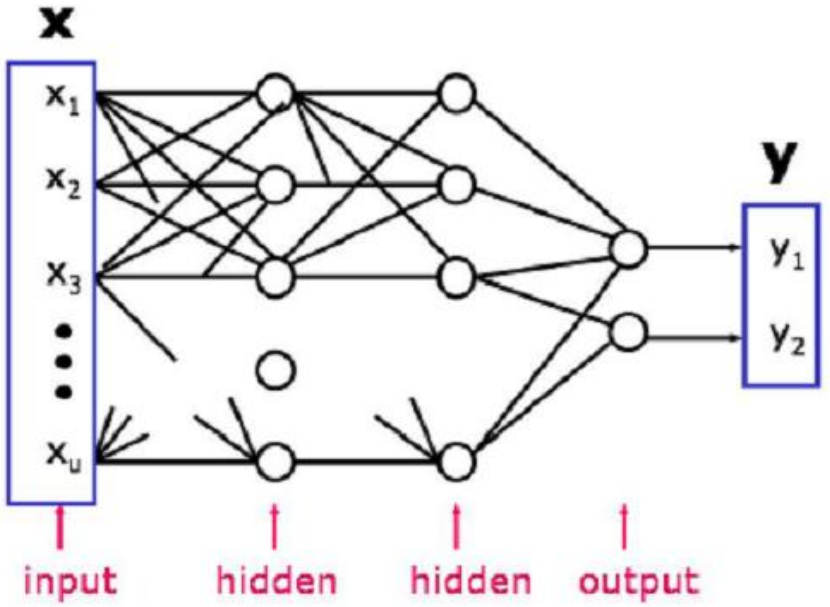
- “Bas”
- “Gauche”
- “Haut”
- “Droite”



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Existing methods

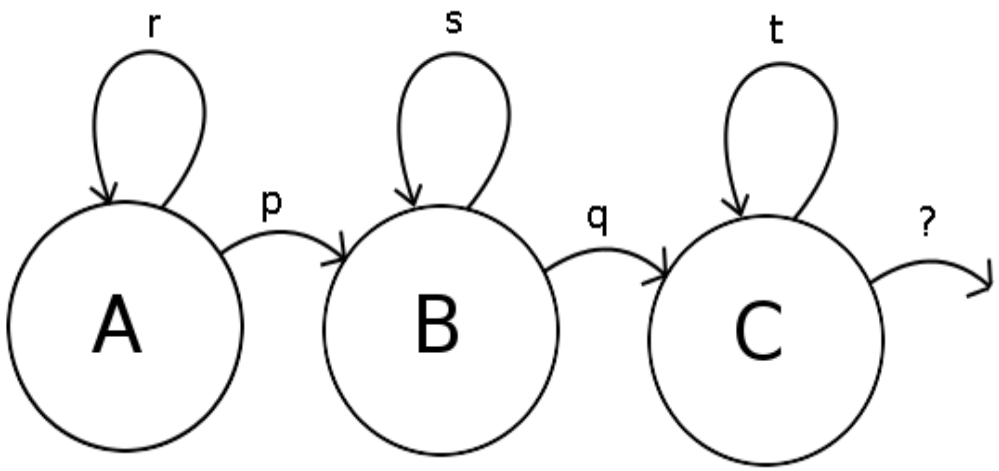
- Artificial neuronal networks:



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

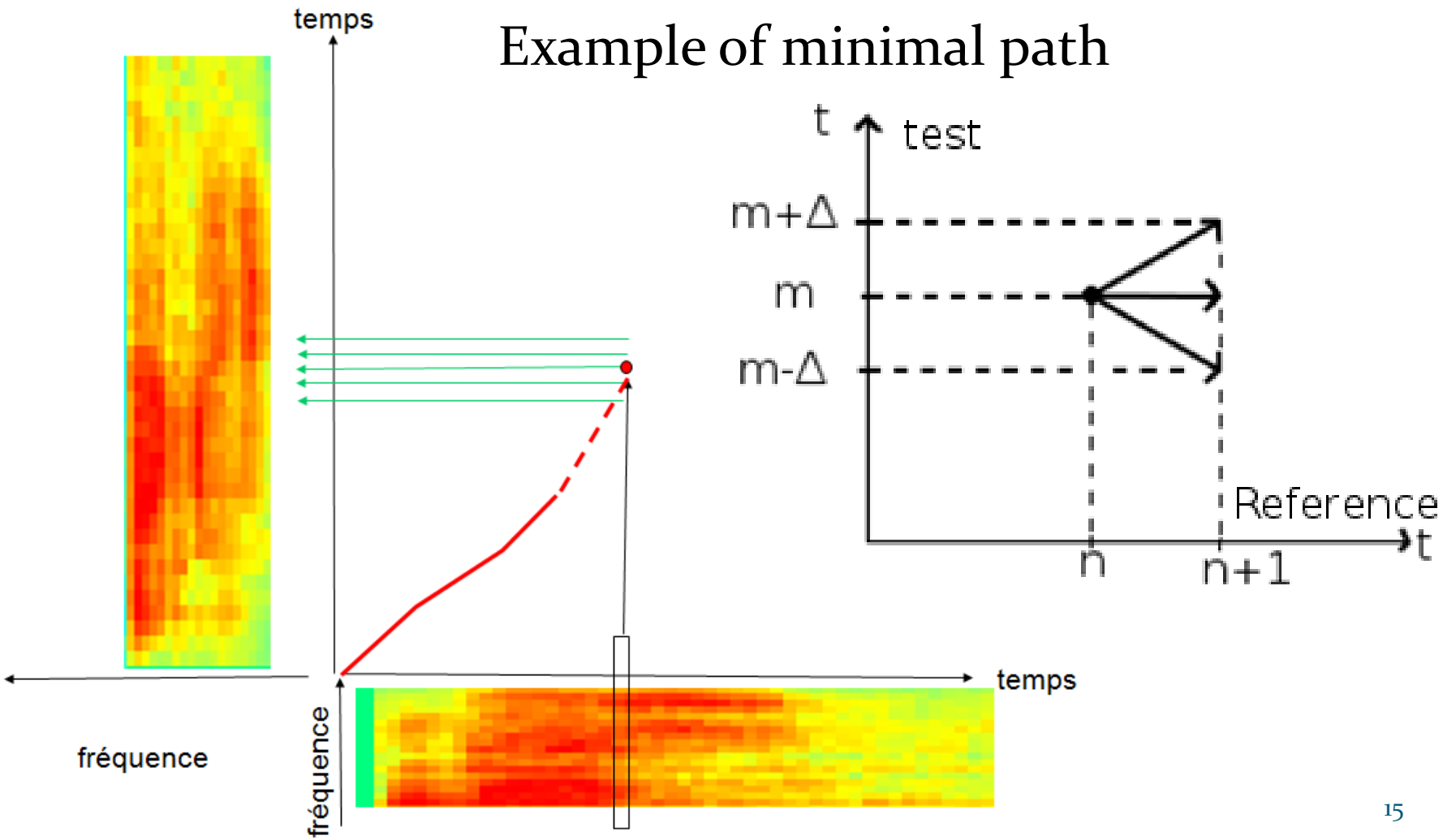
Existing methods

- Hidden Markov model



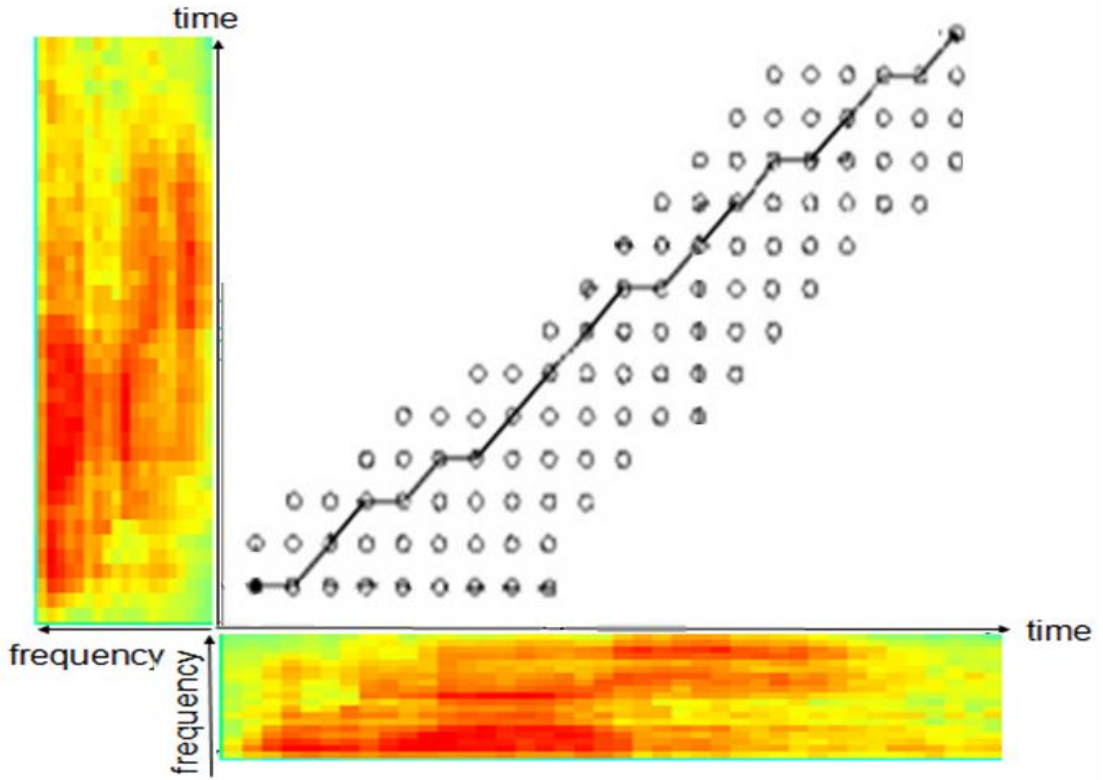
Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Dynamic Time Warping



Dynamic Time Warping

Example of minimal path



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

MAM4

POLYTECH NICE-SOPHIA

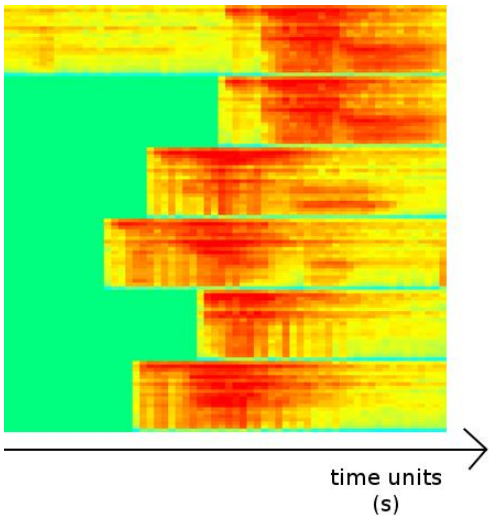
Dynamic Time Warping

- Limits:
  - Euclidean distance
  - Slower for big vocabularies
  - Word width

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Method amelioration

- DTW parameters modifications
- Word beginning detection



- Global approach on DTW

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Results

I. Sound processing

II. Comparison

a. Existing methods

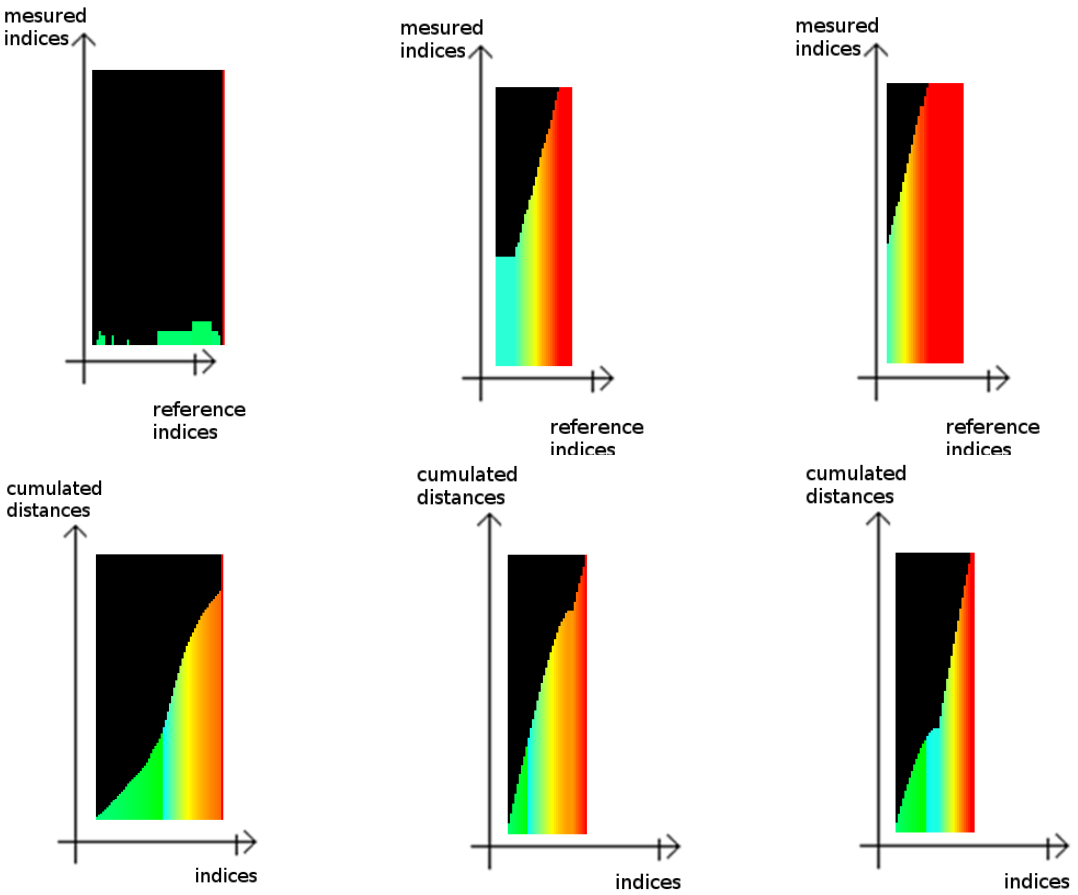
b. Dynamic Time Warping

c. Method amelioration

d. Results

III. Human Machine Interface

- Local DTW / Median DTW / Global DTW



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

POLYTECH NICE-SOPHIA

Results

I. Sound processing

II. Comparison

a. Existing methods

b. Dynamic Time Warping

c. Method amelioration

d. Results

III. Human Machine Interface

Words / Method ...	"gauche", "droite", "haut", "bas"	"Bonjour", "Hello", "Maison", "Placard"	"vacherin", "tiramisu", "moelleux", "bûche"	"Riri", "Fifi", "Loulou", "toto"
DTW local	62,5%	80%	70%	70%
DTW local and beginning detection	62,5%	80%	70%	70%
DTW median and beginning detection	85%	85%	98%	65%
DTW global and beginning detection	95%	90%	98%	85%

Human Machine Interface

I. Sound processing

II. Comparison

III. Human Machine Interface

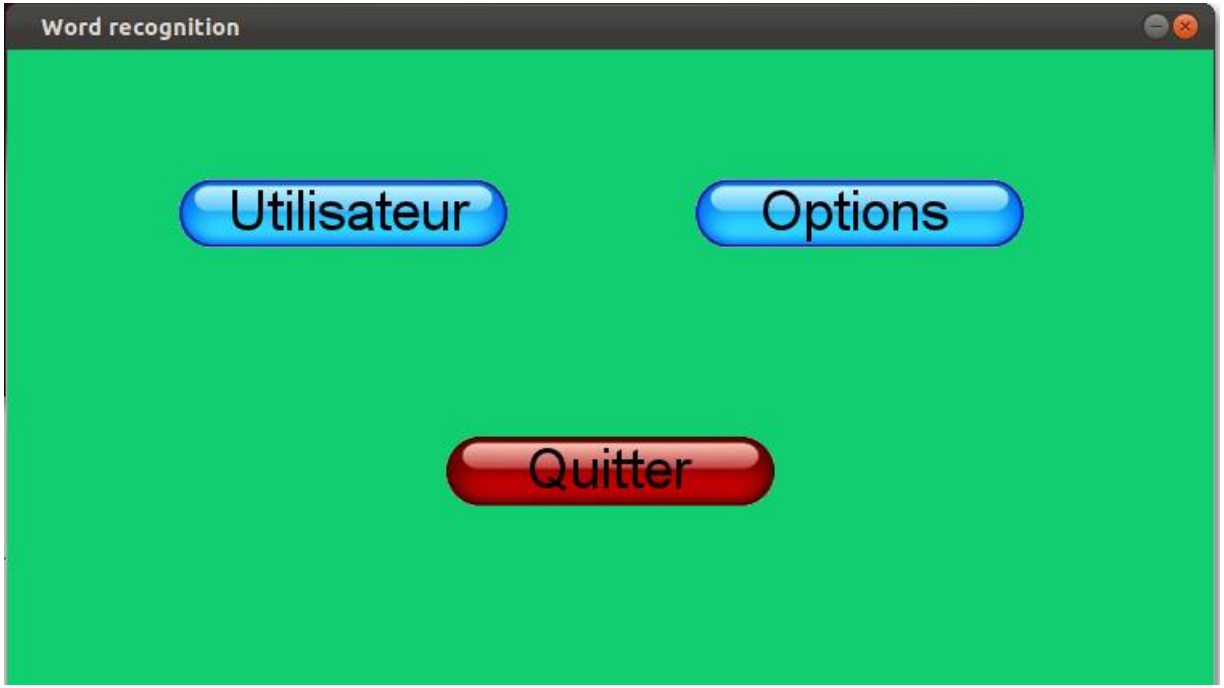
a. Menu

b. How to play

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Menu

I. Sound processing
II. Comparison
III. Human Machine Interface
   a. Menu
   b. How to play



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

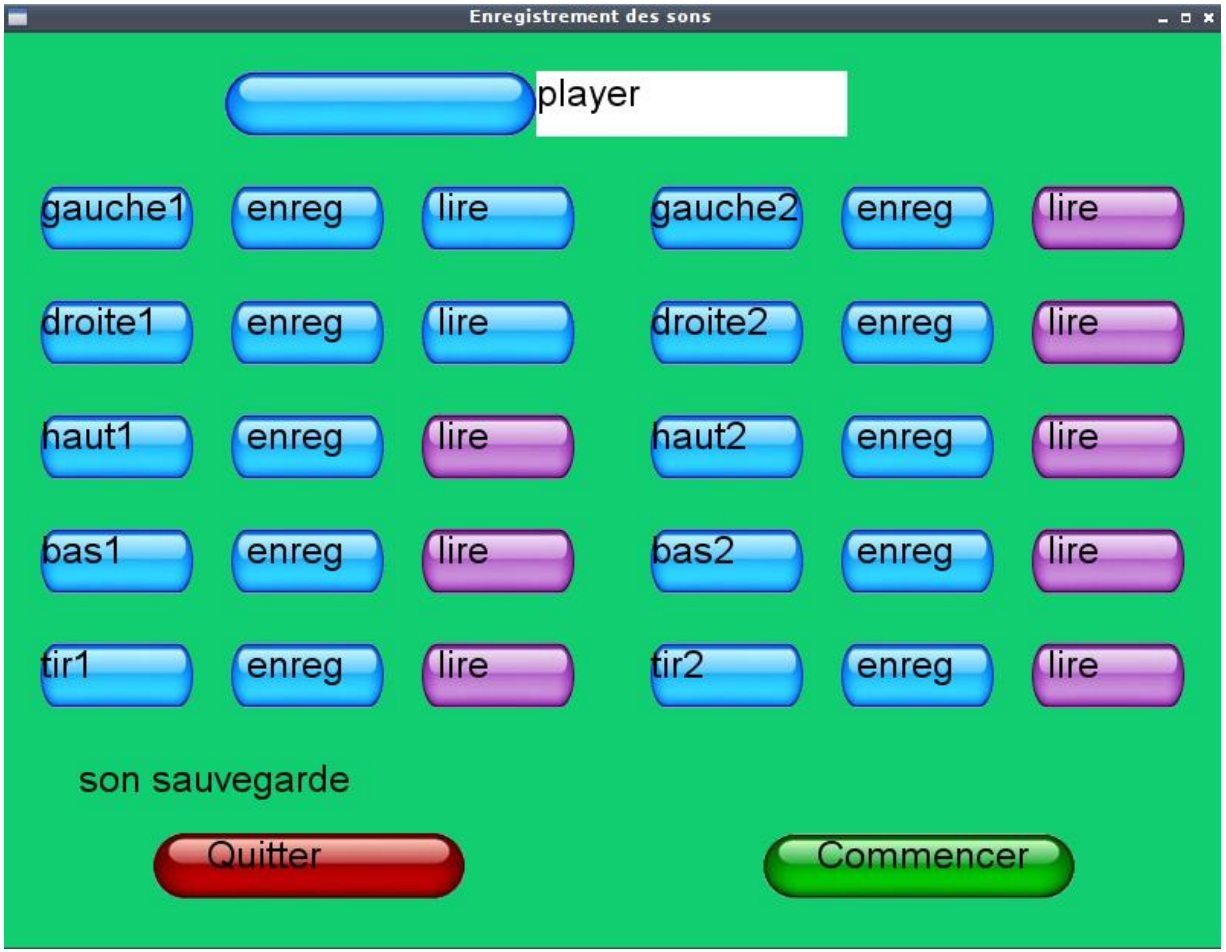
Menu

I. Sound processing
II. Comparison
III. Human Machine Interface
   a. Menu
   b. How to play

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Menu

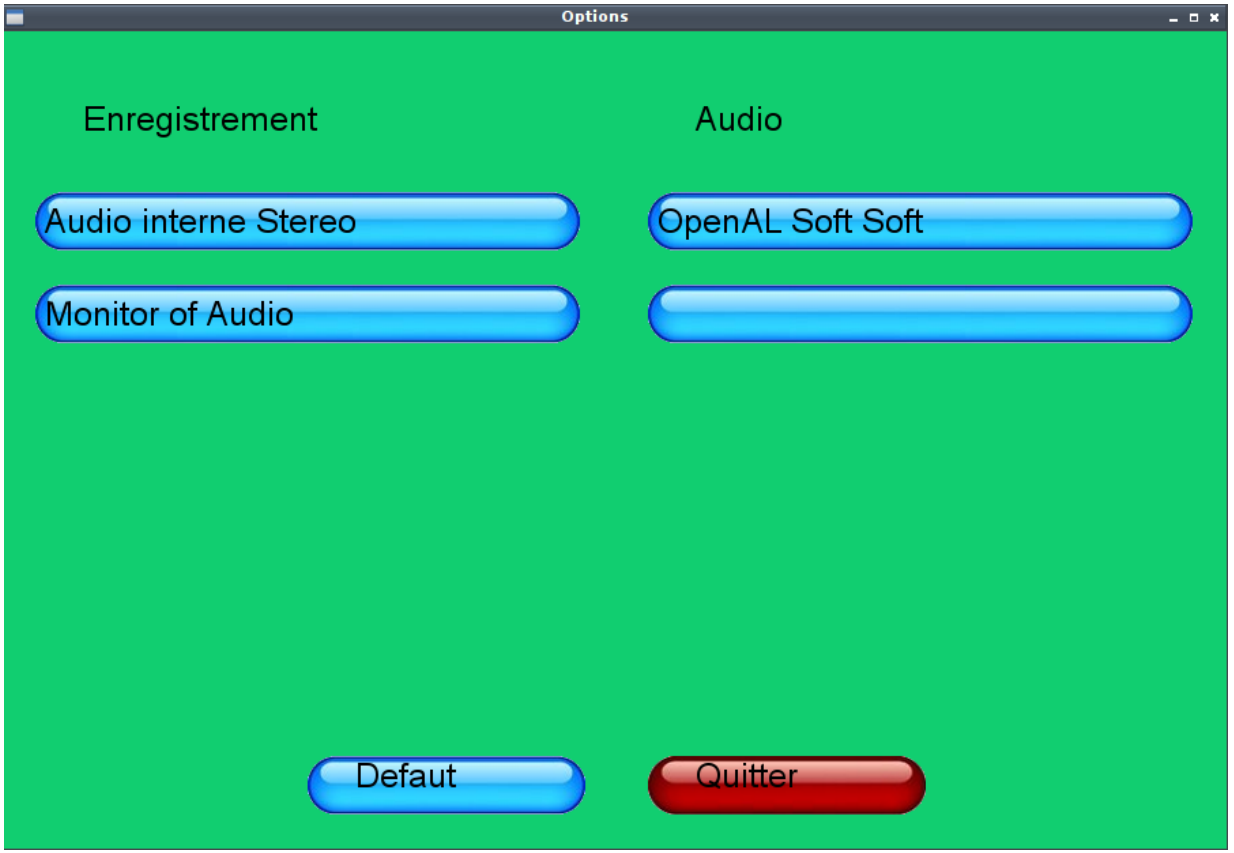
I. Sound processing
II. Comparison
III. Human Machine Interface
   a. Menu
   b. How to play



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Menu

I. Sound processing
II. Comparison
III. Human Machine Interface
   a. Menu
   b. How to play



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Human Machine Interface

I. Sound processing

II. Comparison

III. Human Machine Interface

a. Menu

b. How to play

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

MAM4

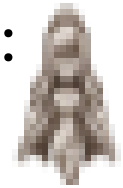
POLYTECH NICE-SOPHIA

Human Machine Interface

I. Sound processing
II. Comparison
III. Human Machine Interface
   a. Menu
   b. How to play

Way in:

Way out:



Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

Human Machine Interface

I. Sound processing
II. Comparison
III. Human Machine Interface
  a. Menu
  b. How to play

Moving:

Waiting:

Wall:

Fire:

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

POLYTECH NICE-SOPHIA

Human Machine Interface

I. Sound processing

II. Comparison

III. Human Machine Interface

a. Menu

b. How to play

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique

MAM4

POLYTECH NICE-SOPHIA

Conclusion

- Interesting project
- Apply our courses(TNS, Infographie, C++)
- Deepen our knowledge
- Others upgrades?

References

- J. Leroux « Dynamic time warping », « HMI », TNS
- Wikipedia
- J. Mariani « Advances and trends in automatic speech recognition », p245-252, 1990
- …

Let’s practice !

Do you have any question?

Guénon Marie
Achard Jean-Paul
Favreau Jean-Dominique