

# Automatic speech segmentation

Jocelyn MAGNE (France)    jmagnus@voila.fr

## What it is about

This worksheet demonstrates the use of the Forward-Backward Divergence model (FBD) in Automatic Speech Segmentation, and how it detects discontinuities in the voice signal. It illustrates in the example below how it is possible to enlarge some segments of the speech (vowels enlargement for instance). To realize this result, it is possible to visually and acoustically perceive the stationary segments of the speech signal.

## More explanations

## Click on the following buttons 1, 2 then 3

(please respect the following advices to choose the wav speech file:

Don't use a wav file containing blanks at the beginning or at the end; don't use a file with background music.

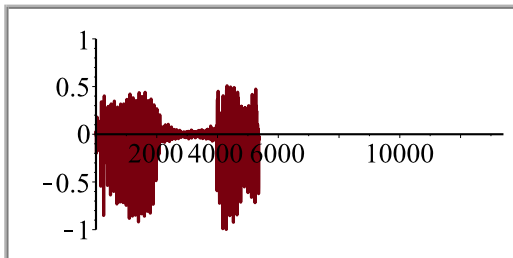
**Some minutes may be required to perform the treatment number 2. To reduce the waiting time, please use a lower sampling frequency file (8KHz), and a sample of voice containing few words (one or two words): for instance the treatment of a 20 Ko wav file needs 3 minutes on a computer with 2.4 GHz intel Core duo processor and 2 Go memory)**

1) [Click Here to Load a Wav speech File](#)

2) [Click Here to Perform the Treatment](#)

3) [Click Here to Save the Wav File Result](#)

Initial wav file



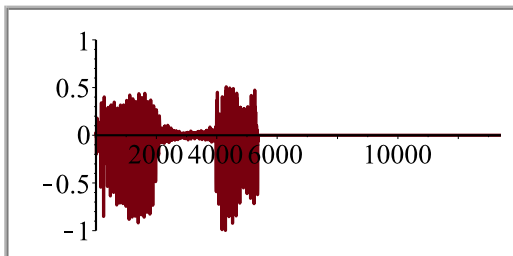
Sampling frequency: 12.80 kHz

File name: /Users/macintosh/assa.wav

Duration: 0.42 seconds

Number of samples: 5365

Result wav file



Sampling frequency: 12.80 kHz

File name: /Users/macintosh/assaaaaa.wav

Duration: 1.14 seconds

Number of samples: 14625

► Settings

## ▼ Conclusions

Although we did not use the complete forward backward divergence model (FBD) as described in the literature, the use of a simple "forward model" was enough for the purpose and give a good result.

## ▼ References

André-Obrecht 93] R. André-Obrecht, "Segmentation et parole?" Habilitation à diriger des recherches, Université de Rennes I, IRISA, Juin 1993.

Sra Sánchez Calle - Curso de procesamiento de la voz - "determinación de la frecuencia fundamental en una señal de voz mediante parametros en el domino del tiempo", Universidad de Granada (ETSII), avril de 2000

François Lonchamp - "cours de phonologie", Université de Nancy 2, 1998

Alain Le Stang - "Cours Maple - Apprendre Maple" <http://alamanya.free.fr/cours.html>

Samir Khan - "Spectrogram Generator" - Maple document from the application center

Julien RACHEDI - "Reconnaissance et classification de phonèmes" Mémoire de master, UPMC, Mars-Août 2005

Pure Data: free open source software for musicians, visual artists, performers, researchers, and developers

## ► See Also

*Legal Notice: © Jocelyn Magne*