Machine Listening In Everyday Life : A Perspective Of A Start-up

CASA Workshop @ DAFx

Boris DEFREVILLE
ORELIA
Sept 23 2011

ORELIA

ircam Centre Pompidou

DAFx 11

Back To My PhD

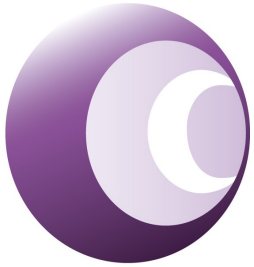
. Find new indicator which goes beyond the sound pressure level (LAeq)

. People refers to sound sources semantics when they recognize, and to physical description when they cannot recognize sound sources

. When you want to assess noise annoyance, you need to detect and recognize the sound sources

. **Tech Starting point** : MIR community

ORELIA Company



The tools we use :

- MARSYAS

- WEKA

- LibSVM

- Sonic Visualiser, Audacity

- GStreamer

Smart Sensor :

AUDIOSENSE 800 : small CPU card with microphone

Real time multiple target detection

Technology :

Classification technology using supervised learning. Binary class. No demixing, just recognizing the most prominent sound source at a moment

Sometimes called biological inspired because we have only a small set a audio features per class and the sets are proper to each class. Features are discriminant and independent.

Primary Function Of Listening

Surviving :

- Prevent from danger

- Detecting anomalies

- Continuously monitoring around

Everyday Listening
#
Musical Listening

Machine Listening Applications

- **Prevent from danger**

Aggression detection (scream, cry for help) and damage for good detection (breaking glass, alarms).

Machine Listening Applications

Detecting anomalies

Predict failure on industrial machines

Supervised and unsupervised classification

Machine Listening Applications

Continuously monitoring around

Noise source recognition (e.g. plane, car, horn, reverse noise) for noise assement/mapping

Challenge n°1 : Diversity Of Sounds

Target sounds + adverse sounds + background noises

Ex : kind of seagull singing in front of the microphone triggering a scream alert. In Paris!!

Interclass confusion : aggression or play?

The key for recognition is the quality of the database and the labels

1/ Target sound :

Make a lot of recordings with different background noise. Don't over-fit your classifier. Quit easy

2/ Adverse sound

Make a lot of recordings but you'll never have it exhaustive. Correct when you have false alarm

Challenge n°2 : Occurrence

Very short sound (1s) occurring once a year!!! (e.g. a break glass, a scream)

Analyze audio every second. 1 false alarm per 12h means a 99.997% accuracy!

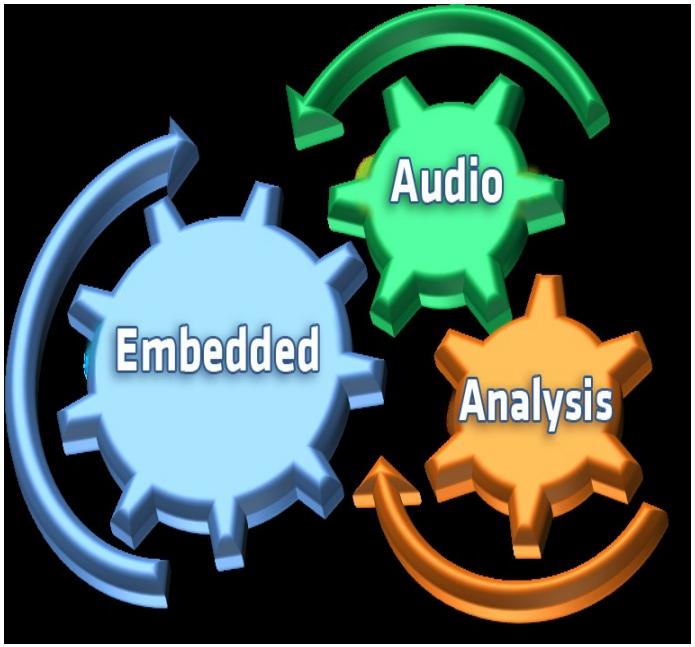
Real criterion is how much false detection for one correct detection

--> make statistics

Challenge n°3 : Privacy = Embedded Analysis

Privacy?

No recordings, no human listening



Solution :

Embed the analysis

Challenge :

Reduced CPU

Challenge n°4 : The Diversity Of Audio Format

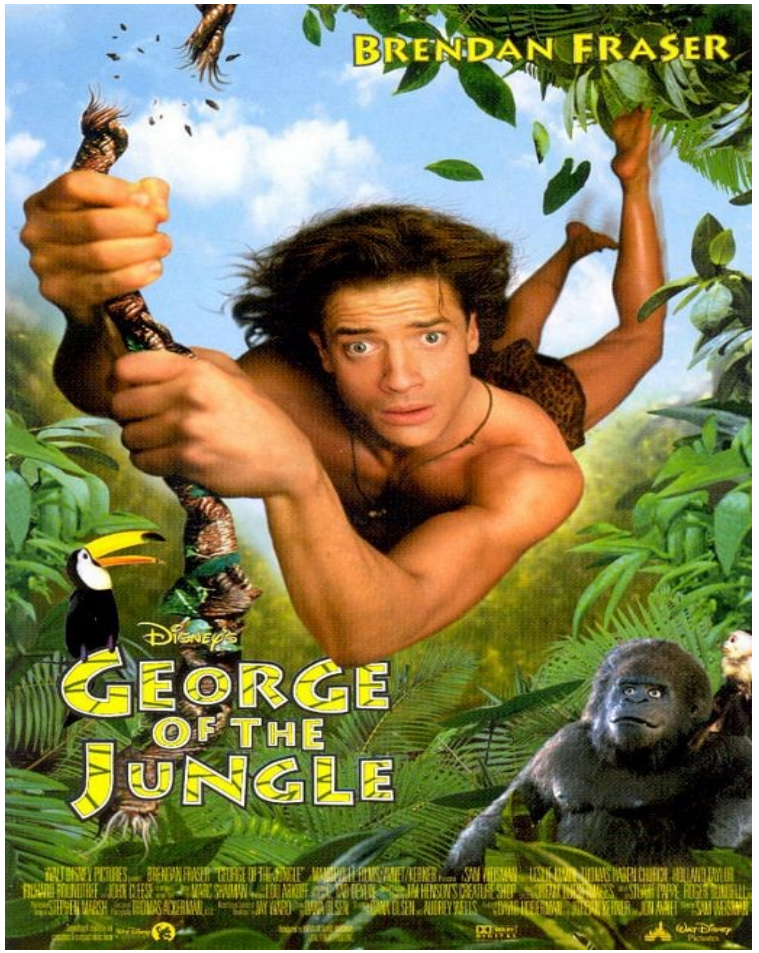
. Encoding (aac, mp3, ulaw, GSM...)

. Frequency response

. Number of bits

. DSP like AGC, Voice Enhancement, noise limiter

-->transform your database and make a new feature selection + learning phase



The Future Of ML Applications

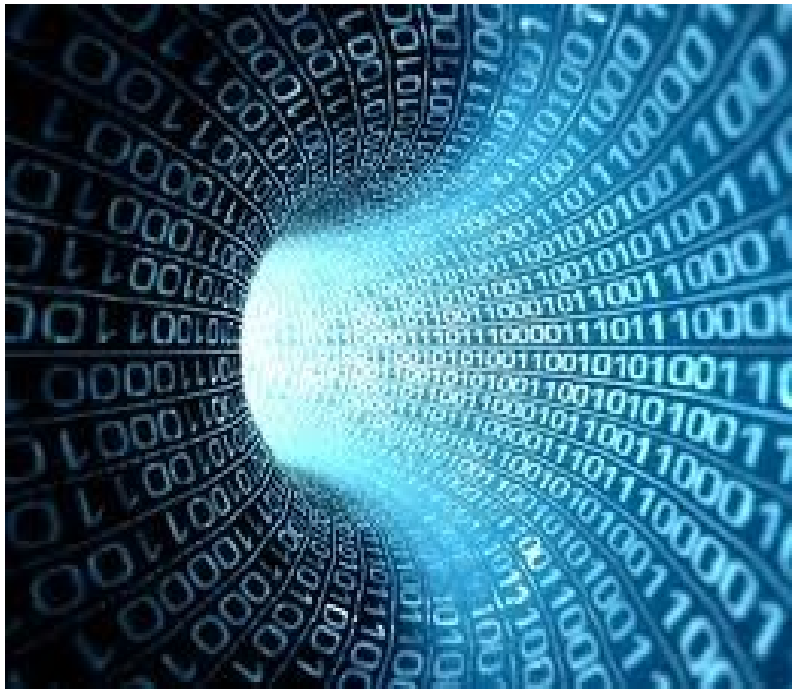
CPU is increasing, microphones are everywhere

Embedded analysis on existing devices

Smartphones, cameras, intercom...

Need to swich from X86 to ARM architecture

Power consumption is a real challenge



ORELIA SAS

20 chemin des prés
77810 Thomery, FRANCE
+33 (0) 951568853
contact@orelia.fr