

# Automatic Detection of Audio Problems for Quality Control in Digital Music Distribution

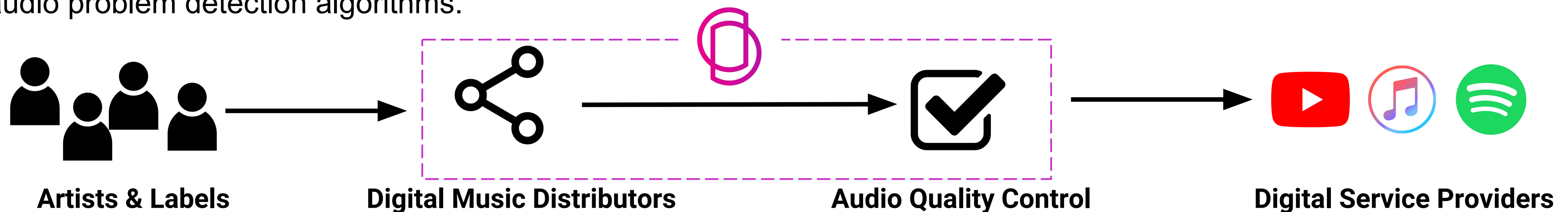
Pablo Alonso-Jiménez<sup>1</sup> Luis Joglar-Ongay<sup>1,2</sup> Xavier Serra<sup>1</sup> Dmitry Bogdanov<sup>1</sup>

Music Technology Group, Universitat Pompeu Fabra<sup>1</sup>  
SonoSuite<sup>2</sup>

Paper 10205

For **Digital Music Distributors**, **Audio Quality Control** is paramount to guarantee the quality of their product.

Given that manual **QC** is very **time** and **resources consuming**, automation is needed in order to develop an efficient and scalable service. This work is a collaboration between **Essentia** and **SonoSuite** in order to create and evaluate audio problem detection algorithms.



**Essentia** is an **C++/Python** library for audio signal processing, developed at the **MTG-UPF** and licensed under **Affero GPL** or **commercial license**.

### Functionalities

- Audio features
  - Spectral features
  - Rhythm and tempo
  - Tonality and melody
  - Fingerprinting
- Real-time **TensorFlow** models support

### Design criteria

- **C++** with **Python** wrappers
- **Large-scale** deployment
- **Real-time** processing
- **Cross-platform**
  - (Linux, MacOS, Win, iOS, Android, JS)

<https://essentia.upf.edu/>

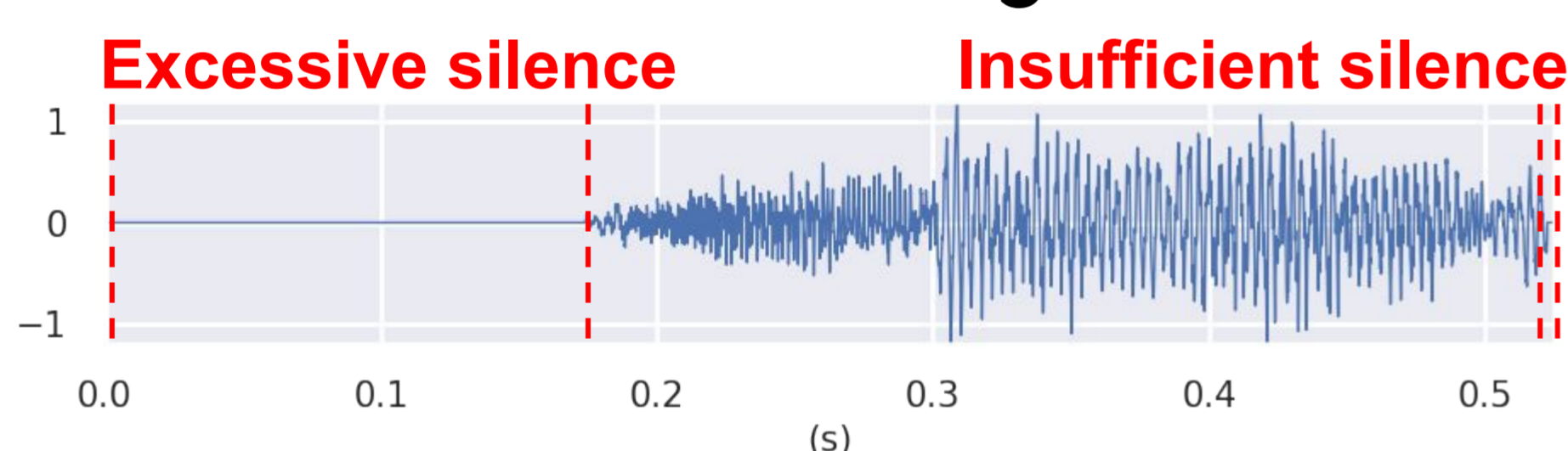
**SonoSuite** is a white label, flexible, scalable and affordable digital music distribution SaaS.

Our team invests in R&D to develop state of the art technologies in the digital music industry.

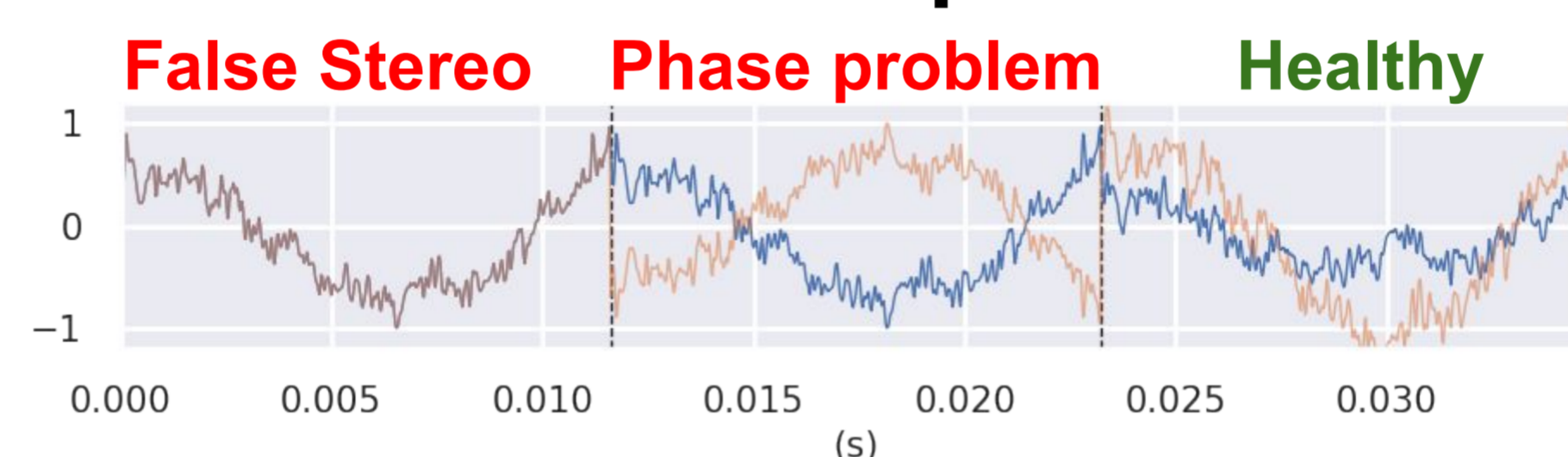
Our vision is to empower media creators around the world to allow them to develop their catalog and manage their distribution, marketing, licensing and royalties accounting activities.

## Audio Problems

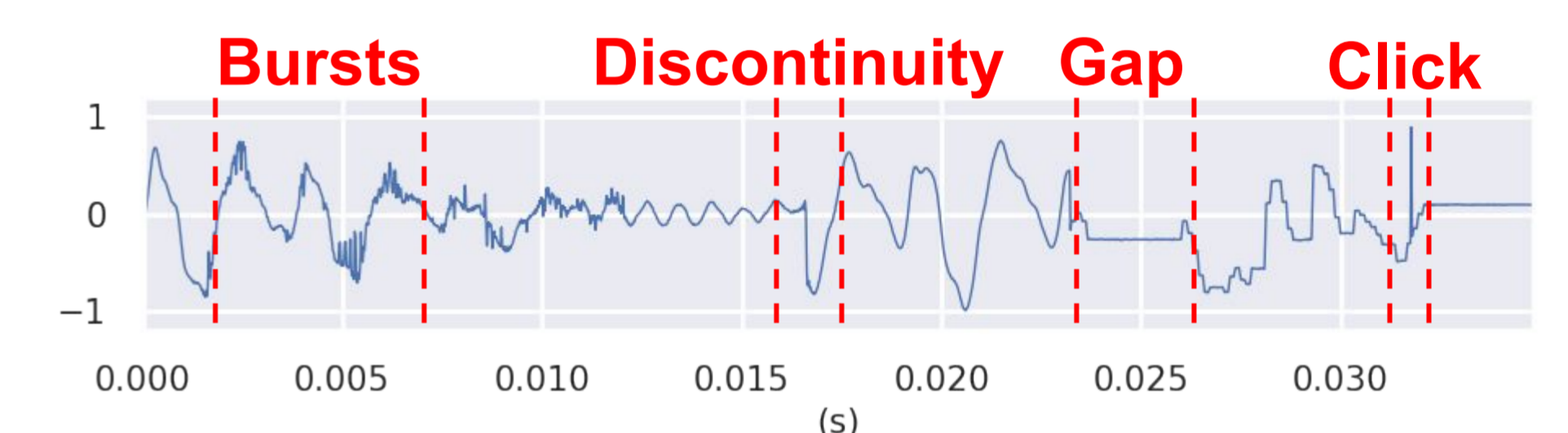
### Incorrect margins



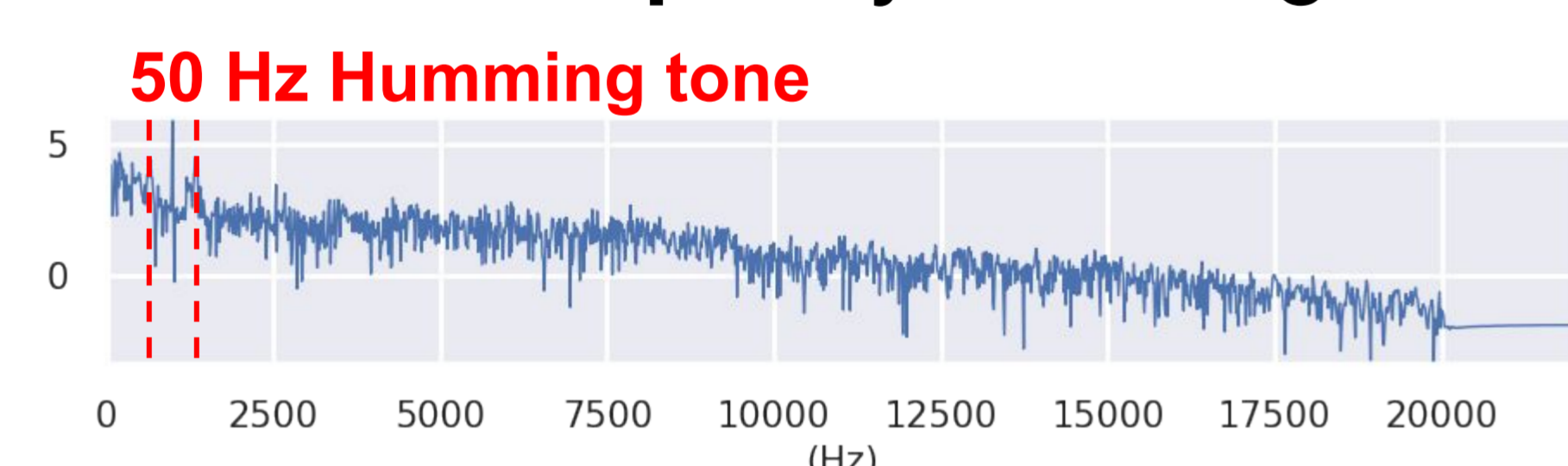
### Phase/Stereo problems



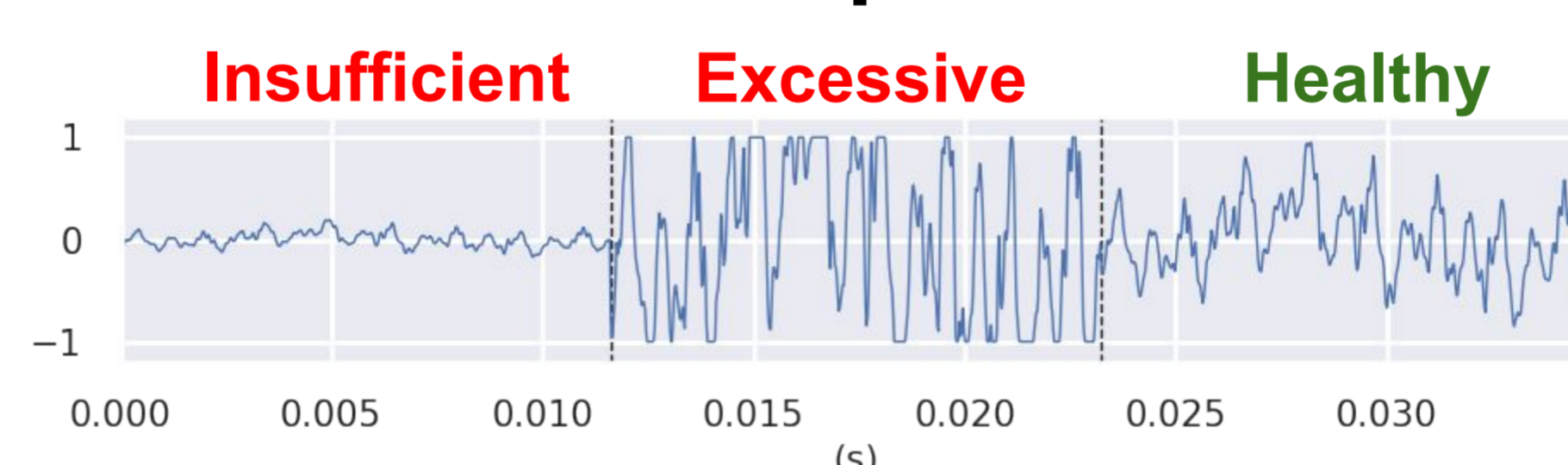
### Audio artifacts



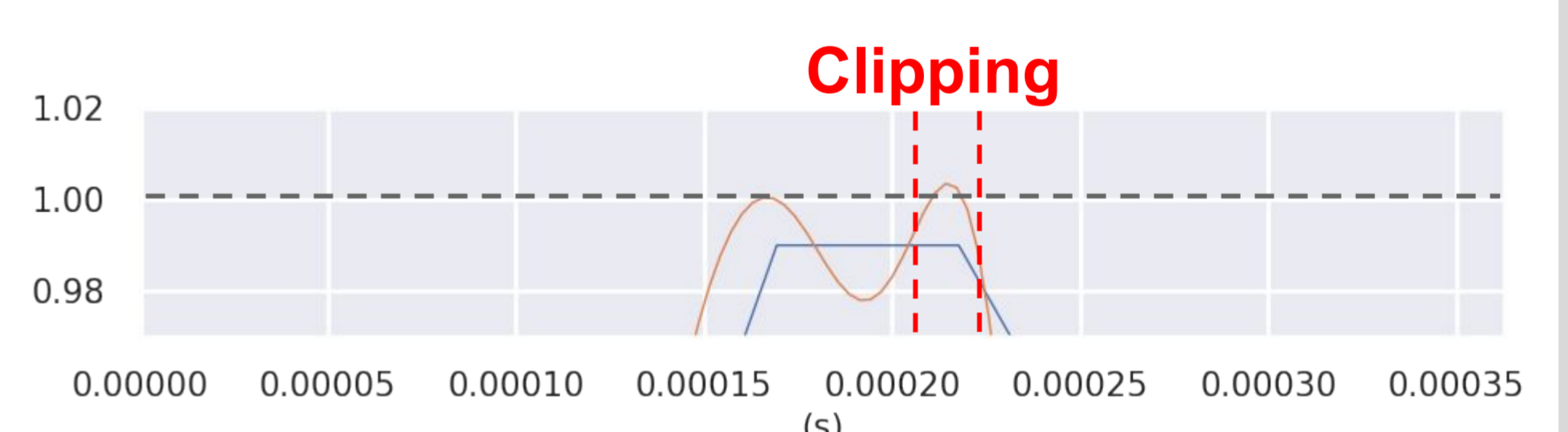
### Low frequency humming



### Loudness problems



### True-Peak detection



## Algorithms

### Incorrect margins

- Energy and time thresholds

### Phase/Stereo problems

- Pearson corr. coefficient

### Audio artifacts

- **Gaps:** audio envelope
- **Clicks:** LPC/matched filters
- **Discontinuities:** LPC
- **Noise bursts:** derivatives
- **Clipping:** threshold on the oversampled signal

### Loudness & Saturation

- **Loudness:** measured following EBU-R128
- **Saturation:** energy and derivatives of the signal

### Noises

- **Humming tones:** stability of the frequency bins in time
- **SNR:** Noise PSD estimated at the beginning of the song

Code at <https://github.com/MTG/essentia>

## Evaluation

### Dataset

- **300k** tracks from **SonoSuite's** collection
- Music from **DMD** all around the world

### Method

- Problem detection rates

### Conclusions

- We have developed and deployed a toolkit for audio problems detection
- These tools are been used in a real industrial scenario

Excessive start silence	0.07%
Excessive end silence	5.32%
Insufficient start silence	16.32%
Insufficient end silence	9.10%
False stereo	3.10%
Phase problems	0.93%
Gaps	0.75%
Clicks and pops	12.64%
Discontinuities	2.63%
Clipping	49.09%
Loudness problems	19.39%
Noise bursts	24.77%
Saturation	1.63%
Humming	62.63%