Automatic karaoke generation.

Subject proposed by Nathan Souviraà-Labastie (Nahimic / Villeneuve d'Ascq)
Master internship, Lille (France)
Expected start : September - October 2019

Advisors

— Nathan Souviraà-Labastie, R&D engineer (Nahimic) nathan.souviraa-labastie@nahimic.com
— Raphaël Greff, R&D Director (Nahimic) raphael.greff@nahimic.com

Company description

Nahimic (a.k.a. A-Volute) is a company based in Villeneuve d'Ascq (France) that publishes audio enhancement software for the gaming industry, in particular the Nahimic software on MSI laptop. Nahimic has developed a solution for digital and real-time 3D sound. The suite of audio effects proposed by Nahimic includes effects to improve multimedia content (music or movie), as well as microphone effects for communication (noise reduction).

Conundrum

A Karaoke version of a music piece is a version where the singer's voice is no longer present in the song. Generally, such a version of the music is presented with subtitles of the lyrics allowing the user to sing to the rhythm of the "instrumental" piece.

Most of the time, these Karaoke versions are generated ("mastered") by hand by a sound engineer. Entertainment companies already have large databases of this type of content. However, they can not cope with the amount of songs created every day, especially by amateur musicians, and must focus on the most famous songs. Thus, an automatic Karaoke generation tool would allow the general public to access a potentially infinite database of Karaoke. Similarly, in the case of streamed content, an automatic (and real-time) tool would also be required.

Approach

This internship will focus on the automatic generation of subtitles for such content, with the subtitles being synchronized with the music pieces.

A first axis of work would be the adaptation of a state of the art speech-to-text method for singing voice.

A second axis of work would be to use the lyrics that are available online in plain text version. It would then be a matter of synchronizing and displaying this version based on a comparison with the version produced by the algo of speech-to-text.

Skills

Who are we looking for ? Preparing an engineering degree or master's degree, or even a PhD (3 month visit), you preferably have knowledge in the development and implementation of advanced algorithms for digital audio signal processing. Skills or experience in Natural Language Processing (NLP) or symbolic data processing would be a plus.

In addition, notions in the following various fields would be appreciated : - Audio, acoustics and psychoacoustics - Audio effects in general : compression, equalization, etc. - Machine learning and artificial neural networks. - Statistics, probabilist approaches, optimization. - Programming language : Matlab, Python.

as well as experiences in the following areas : - Sound spatialization effects : binaural synthesis, ambisonics, artificial reverberation. - Voice recognition, voice command. - Voice processing effects : noise reduction, echo cancellation, antenna processing. - Virtual, augmented and mixed reality. - Computer programming and development : Max/MSP, C/C+++/C#. - Video game engines : Unity, Unreal Engine, Wwise, FMod, etc. - Audio editing software : Audacity, Adobe Audition, etc. - Scientific publications and patent applications. - Fluent in English and French. - Demonstrate intellectual curiosity.

Other offers

https://nahimic.welcomekit.co/
https://www.welcometothejungle.co/companies/nahimic/jobs