2018-00647 - PhD Position / Expressive speech synthesis based on deep learning [S]

Level of qualifications required: Graduate degree or equivalent
Fonction: PhD Position

Context
Team:
   MULTISPEECH (https://team.inria.fr/multispeech/) - Inria Nancy - Grand-Est

Contacts:
   Denis Jouvet (denis.jouvet@loria.fr; https://members.loria.fr/DJouvet/)
   Vincent Colotte (Vincent.colotte@loria.fr; https://members.loria.fr/VColotte/)

Assignment
Scientific Context

Over the last decades, text-to-speech synthesis (TTS) has reached good quality and intelligibility, and is now commonly used in information delivery services, as for example in call center automation, in navigation systems, and in voice assistants. In the past, the main goal when developing TTS systems was to achieve high intelligibility. The speech style was then typically a “reading style”, which resulted from the style of the speech data used to develop TTS systems (reading of a large set of sentences). Although a reading style is acceptable for occasional interactions, TTS systems should benefit from more variability and expressivity in the generated synthetic speech, for example, for lengthy interactions between machines and humans, or for entertainment applications. This is the goal of recent or emerging research on expressive speech synthesis. Contrary to neutral speech, which is typically read speech without conveying any particular emotion, expressive speech can be defined as speech carrying an emotion, or spoken as in spontaneous speech, or also as speech with emphasis set on some words.

Main activities
Missions

Deep learning approaches leads to good speech synthesis quality, however the main scientific and technological barrier remains the necessity of having a speech corpora corresponding to the speaker and the target style conditions, here expressive speech. This thesis aims at investigating approaches to overcome this barrier. More precisely, the objective is to propose and investigate approaches allowing expressive speech synthesis for a given speaker voice, using both the neutral speech data of that speaker, or the corresponding neutral speech model, and expressive speech data from other speakers. This will avoid lengthy and costly recording of specific ad hoc expressive speech corpora (e.g., emotional speech data from the target voice speaker).

Let recall that three main steps are involved in parametric speech synthesis: the generation of sequences of basic units (phonemes, pauses, etc.) from the source text; the generation of prosody parameters (durations of sounds, pitch values, etc.); and finally the generation of acoustic parameters, which leads to the synthetic speech signal. All the levels are involved in expressive speech synthesis: alteration of pronunciations and presence of pauses, modification of prosody correlates and modification of the spectral characteristics.
The thesis will essentially focus on the two last points, i.e., a correct prediction of prosody and spectral characteristics to produced expressive speech through deep learning-based approaches. Some aspects to be investigated include the combined used of only the neutral speech data of the target voice speaker and expressive speech of other speakers in the training process, or in an adaptation process, as well as data augmentation processes.

The baseline experiments will rely on neutral speech corpora and expressive speech corpora previously collected for speech synthesis in the Multispeech team. Further experiments will consider using other expressive speech data, possibly extracted from audiobooks.

**Bibliography**


**Skills**

**Required qualifications**

- Master in automatic language processing or in computer science

**Skills and profile**

- Background in statistics, and in deep learning
- Experience with deep learning tools
- Good computer skills (preferably in Python)
- Experience in speech synthesis is a plus

**Benefits package**

- Subsidised catering service
- Partially-reimbursed public transport
- Social security
- Paid leave
- French courses

**Remuneration**

Gross Salary per month: €1982 (year 1 & 2) and €2085 (year 3)

**General Information**

- **Town/city**: Villers-lès-Nancy
- **Inria Center**: CRI Nancy - Grand Est
- **Starting date**: 2018-10-01
- **Duration of contract**: 3 years
- **Deadline to apply**: 2018-05-01
Contacts

- **Inria Team**: MULTISPEECH
- **Recruiter**: Jouvet Denis / denis.jouvet@inria.fr

The keys to success

Application deadline

May 1st, 2018 (Midnight Paris time)

How to apply

Upload your file on jobs.inria.fr in a single pdf or zip file, and send it as well by email to denis.jouvet@inria.fr. Your file should contain the following documents:

- Your CV.
- A cover/motivation letter describing your interest in this topic.
- A short (max one page) description of your Master thesis (or equivalent) or of the work in progress if not yet completed.
- Your degree certificates and transcripts for Bachelor and Master (or the last 5 years).
- Master thesis (or equivalent) if it is already completed and publications if any (it is not expected that you have any). Only the web links to these documents are preferable, if possible.

In addition, one recommendation letter from the person who supervises(d) your Master thesis (or research project or internship) should be sent directly by his/her author to denis.jouvet@inria.fr.

Applications are to be sent as soon as possible.

Conditions for application

**Defence Security**:
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

**Recruitment Policy**:
As part of its diversity policy, all Inria positions are accessible to people with disabilities.

**Warning**: you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.