

Proposal for a postdoctoral position at INRIA, Bordeaux, France

Title: *Sparse predictive models for the analysis and classification of pathological speech*

Keywords: Pathological speech processing, Sparse modeling, Optimization algorithms, Machine learning, Parkinsonian disorders, Respiratory diseases

Contact and Supervisor: Khalid Daoudi (khalid.daoudi@inria.fr)

INRIA team: GEOSTAT (geostat.bordeaux.inria.fr)

Duration: from 01/11/2021 to 31/12/2022 (could be extended to an advanced or a permanent position)

Salary: 2653€ / month

Profile: PhD degree obtained after August 2019 or to be defended by the end of 2021. High quality applications with a PhD obtained before August 2019 could be considered for an advanced research position.

Required Knowledge and background: A solid knowledge in speech/signal processing; A good mathematical background; Basics of machine learning; Programming in Matlab and Python.

Scientific research context

During this century, there has been an ever increasing interest in the development of objective vocal biomarkers to assist in diagnosis and monitoring of neurodegenerative diseases and, recently, respiratory diseases because of the Covid-19 pandemic. The literature is now relatively rich in methods for objective analysis of dysarthria, a class of motor speech disorders [1], where most of the effort has been made on speech impaired by Parkinson's disease. However, relatively few studies have addressed the challenging problem of discrimination between subgroups of Parkinsonian disorders which share similar clinical symptoms, particularly in early disease stages [2]. As for the analysis of speech impaired by respiratory diseases, the field is relatively new (with existing developments in very specialized areas) but is taking a great attention since the beginning of the pandemic.

On the other hand, the large majority of existing processing methods (of pathological speech in general) still heavily rely on a core of feature estimators designed and optimized for healthy speech. There exist thus a strong need for a framework to infer/design speech features and cues which remain robust to the perturbations caused by (classes of) disordered speech. The first and main objective of this proposal is to explore the framework of sparse modeling of speech which allow a certain flexibility in the design and parameter estimation of the source-filter model of speech production. This exploration will be essentially based on theoretical advances developed by the GEOSTAT team and which have led to a significant impact in the field of image processing, not only at the scientific level [3] but also at the technological level (www.inria.fr/fr/i2s-geostat-un-innovation-lab-en-imagerie-numerique).

The second objective of this proposal is to use the resulting representations as inputs to basic machine learning algorithms in order to conceive a vocal biomarker to assist in the discrimination between subgroups of Parkinsonian disorders (Parkinson's disease, Multiple-System Atrophy, Progressive Supranuclear Palsy) and in the monitoring of respiratory diseases (Covid-19, Asthma, COPD).

Both objectives benefit from a rich dataset of speech and other biosignals recently collected in the framework of two clinical studies in partnership with university hospitals in Bordeaux and Toulouse (for Parkinsonian disorders) and in Paris (for respiratory diseases).

Work description

As stated above, the work to be carried is decomposed in two parts. The main part consists in developing new algorithms, based on sparse modeling, for the analysis of a class of disordered speech. The second part consists in exploring machine learning tools to develop vocal biomarkers for the purpose of (differential) diagnosis and monitoring of the diseases under study.

1. Sparse modeling for disordered speech analysis

The first task will be to investigate sparsity in the framework of linear prediction modeling of speech. The latter is indeed one of the building blocks for the estimation of core glottal, phonation and articulatory features. Sparse linear prediction (SLP) has been recently investigated in a convex setting using the L1-norm and applied, essentially, to speech coding [4]. We will start by investigating the potential of this convex setting in disordered speech analysis. We will then explore the use of non-convex penalties that allow sparsity control and a better decoupling the vocal tract filter from excitation source. We will study the spectral properties of the different models and revisit a set of acoustic features which are not robust to perturbations raising in dysarthric speech. We will then explore the potential of SLP in designing new features which could be informative about dysarthria. The algorithmic developments will be evaluated using a rich set of biosignals obtained from patients with Parkinsonian disorders and from healthy controls. The biosignals are electroglottography and aerodynamic measurements of oral and nasal airflow as well as intra-oral and sub-glottic pressure.

After dysarthria analysis, we will study speech impairments caused by respiratory deficits. The main goal here will be to automatically identify respiratory patterns and to design features to quantify the impairments. The developments will be evaluated using manual annotations, by an expert phonetician, of speech signals obtained from patients with respiratory deficit and from healthy controls.

Depending on the work progress and time constraints, we may also explore sparsity beyond the linear prediction model through existing nonlinear representations of speech. It is well known indeed that the linear source-filter model of speech cannot capture several nonlinearities which exist in the speech production process, particularly in disordered speech.

2. Machine learning for disease diagnosis and monitoring

Using the outcomes of the first part, the (experimental) objective of the second part is to apply basic machine learning algorithms (LDA, logistic regression, decision trees, SVM...) using standard tools (such as Scikit-Learn) to conceive robust algorithms that could help, first, in the discrimination between Parkinsonian disorders and, second, in the monitoring of respiratory deficit.

3. Work synergy

- The postdoc will interact closely with an engineer who is developing an open-source software architecture dedicated to pathological speech processing. The validated algorithms will be implemented in this architecture by the engineer, under the co-supervision of the postdoc.
- Giving the multidisciplinary nature of the proposal, the postdoc will interact with the clinicians participating in the two clinical studies.

References:

- [1] J. Duffy. *Motor Speech Disorders Substrates, Differential Diagnosis, and Management*. Elsevier, 2013.
- [2] J. Ruzs et al. *Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy*. *Journal of Neurology*, 262(4), 2015.
- [3] H. Badri. *Sparse and Scale-Invariant Methods in Image Processing*. PhD thesis, University of Bordeaux, France, 2015.
- [4] D. Giacobello et al. *Sparse Linear Prediction and Its Applications to Speech Processing*. *IEEE Transactions on Audio Speech and Language Processing*, (20)5, 2012.