decoding for robust ASR, in GMM-derived uncertainty features for noise robust ASR, [Nathwani et al., 2013].

representations of words and phrases and their compositionality, [Mikolov et al., 2013b] Mikolov, T., Sutskever, I., Chen, T. Corrado, G.S., and Dean, J. Distributed representations in vector space, [Mikolov et al., 2013a] Mikolov, T. Chen, K., Corrado, G., and Dean, J. Efficient estimation of word appearance of words (context of a word) can be used to describe the underlying semantics. The most used method to learn these representations is to predict a word using the context in which this word appears [Mikolov et al., 2013b], Pennington et al., 2014, and this can be realized with neural networks. These representations have proved effective for a series of natural language processing Baroni et al., 2014]. In particular, Mikolov's Skip-gram and CBOW models et al. [Mikolov et al., 2013b, Pennington et al., 2014] have become very popular because of their ability to process large amounts of unstructured text data with reduced computing costs. Efficiency and the semantic properties of these representations motivate us to explore these semantic representations for our task of recognition in noisy conditions.

Societé et atouts du poste

MULTISPEECH is a joint research team between the Université of Lorraine, Inria, and CNRS. Its research focuses on speech processing, with particular emphasis to multisource (source separation, robust speech recognition), multilingual (computer assisted language learning), and multimodal aspects (audiovisual synthesis).

Mission confiée

The goal of this PhD Thesis will be devoted to the innovative study of the taking into account of semantics through predictive representations that capture the semantic features of words and their context. Research will be conducted on the combination of semantic information with information from denoising to improve speech recognition.

Principalités activités

Context and objectives

Under noisy conditions, audio acquisition is one of the toughest challenges to have a successful automatic speech recognition (ASR). One possible approach relies on the ability to attenuate ambient noise in the signal and to take it into account in the acoustic model used by the ASR. Our DNN (Deep Neural Network) denoising system and our approach to exploiting uncertainties have shown their combined effectiveness against noisy speech. To go further and to improve the performance of the automatic speech recognition in noisy conditions, we propose to use semantic or thematic information. The addition of semantic information will remove ambiguities due to the background noise.

Semantic and thematic spaces are vector spaces used for representation number of words, sentences or textual documents. The corresponding models and methods have a long history in the field of computational linguistics and natural language processing [Turney and Pantel, 2010]. Almost all models rely on the hypothesis of statistical semantics which states that: Statistical patterns of appearance of words (context of a word) can be used to describe the underlying semantics. The most common method to learn these representations is to predict a word using the context in which this word appears [Mikolov et al., 2013b], Pennington et al., 2014, and this can be realized with neural networks. These representations have proved effective for a series of natural language processing [Baroni et al., 2014]. In particular, Mikolov’s Skip-gram and CBOW models et al. [Mikolov et al., 2013b], Mikolov et al., 2013a have become very popular because of their ability to process large amounts of unstructured text data with reduced computing costs. Efficiency and the semantic properties of these representations motivate us to explore these semantic representations for our task of recognition in noisy conditions.

Main activities

The goal of this PhD Thesis will be devoted to the innovative study of the taking into account of semantics through predictive representations that capture the semantic features of words and their context. Research will be conducted on the combination of semantic information with information from denoising to improve speech recognition.

The ASR stage will be supplemented by a semantic analysis to detect the words of the processed sentence that could have been misrecognized and to offer similar (at the acoustic level) words that better fit the context. Predictive representations using continuous vectors have been shown to capture the semantic characteristics of words and their context, and to overcome representations based on counting words. Semantic analysis will be performed by combining predictive representations using continuous vectors and information from denoising. This combination could be done by the rescoring component. All our models will be based on the powerful paradigm of DNN. The performances of the various modules will be evaluated on artificially noisy speech signals and on real noisy data.

References


Informations générales

• Ville : Villers-lès-Nancy
• Centre Inria : CRIS Nancy - Grand Est
• Date de prise de fonction souhaitée : 2019-09-02
• Durée de contrat : 3 ans
• Date limite pour postuler : 2019-05-01

Contacts

• Equipe Inria : MULTISPEECH
• Directeur de thèse : Ilirna Illina / ilirna.illina@loria.fr

A propos d’Inria

Inria, l’institut national de recherche dédié aux sciences du numérique, promeut l’excellence scientifique et le transfert pour avoir le plus grand impact. Il emploie 4200 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3000 scientifiques pour relever les défis des sciences informatiques et mathématiques, souvent à l’interface d’autres disciplines. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 160 start-up. L’institut s’efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l’économie.

L’essentiel pour réussir

Application deadline
May 1st, 2019 (Midnight Paris time)

How to apply
Upload your CV on jobs.inria.fr; this should be a pdf file of at most 2Mo.
In addition, send the following documents to ilirna.illina@loria.fr in a single pdf or ZIP file:
- CV
- A cover/motivation letter describing your interest in this topic.
- Your degree certificates and transcripts for Bachelor and Master (or the last 5 years).
- Master thesis (or equivalent) if it is already completed, or a description of the work in progress, otherwise.
- Publications or web links, if any (it is not expected that you have any).
In addition, one recommendation letter from the person who supervised(s) your Master thesis (or research project or internship) should be sent directly by his/her author to stephan.merz@inria.fr.
Applications are to be sent as soon as possible.

Consignes pour postuler

Sécurité défense :
Ce poste est susceptible d’être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L’autorisation d’accès à une zone est délivrée par le chef d’établissement, après avis ministériel favorable, tel que défini dans l’arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l’annulation du recrutement.

Politique de recrutement :
Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.
Compétences

Required skills: Background in mathematics, machine learning (DNN), statistics, natural language processing and computer program skills (Perl, Python).

English writing and speaking skills are required in any case.

Candidates should email a detailed CV with diploma

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération


Monthly salary after taxes: around €1596.05 for 1st and 2nd year. €1678.99 for 3rd year. (medical insurance included).

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n’est pas garanti.