Universidad del País Vasco Euskal Herriko Unibertsitatea | MDE Master eta Doktorego Eskola Escuela de Máster y Doctorado Master and Doctoral School

Programa de Doctorado / *PhD Programme*

Tecnologías de la Información y Comunicaciones en Redes Móviles
Mobile Network Information and Communication Technologies

SEMINAR:

LANGUAGE-INDEPENDENT ACOUSTIC CLONING OF HTS VOICES

INVITED SPEAKER:

Carmen Magariños, is a PhD student at the Multimedia Technology Group (GTM) of the University of Vigo, Spain. She received her Telecommunication Engineering degree in 2011, and her M.Sc. degree in 2014, both from the University of Vigo. In 2011 she joined the GTM where she has worked as a research engineer under several projects. Her research interests are focused on speech technology, mainly on HMM-based speech synthesis and speaker adaptation.

OUTLINE:

Text-to-Speech (TTS) systems have traditionally required a costly process of recording new voices in order to synthesize new speakers, speaking styles, emotions or languages. However, the emergence of statistical parametric speech synthesis made possible the modification of the speech characteristics and/or speaker identity in a quite flexible way, avoiding the need of long additional recordings. In particular, HMM-based speech synthesis is able to make such modifications by means of speaker adaptation techniques. Nevertheless, up to now, most of the work in this field has been focused on *intra-lingual* speaker adaptation (source and target speakers speak the same language) while the *cross-lingual* paradigm (source and target speakers speak different languages) seems to be less explored. In this talk, a new method for cross-lingual speaker adaptation will be presented. Within the standard HMM-based synthesis framework, the proposed method combines the language-dependent structure of a synthesis voice model with the acoustic characteristics of another model trained for a different language. Unlike classical adaptation techniques, where the source model is transformed to fit some input data from a target speaker, this method transforms the source model to be acoustically closer to another model that conveys the identity of the target speaker. As a result, a third model is built, which is able to produce synthetic speech in the same language as the source model while sounding like the target speaker.

DATE AND LOCATION:

Thursday May, 5th 12:00 - 13:00 Location: Seminar **P3B2S**	Faculty of Engineering Alda. Urquijo S/N 48013 Bilbao

REGISTRATION AND CONTACT INFO:

For registration details and additional information, contact inma.hernaez@ehu.eus. Doctoral students have priority at registration.