A Neural Network Approach for the Design of the Target Cost Function in Unit-Selection Speech Synthesis

Francisco Campillo Dáaz, Jos Luis Alba, Eduardo Rodríguez Banga

Universidade de Vigo, Spain

Corpus-based speech synthesis performance depends on the skill to model and represent appropriately all the characteristics of the speech units that serve as a basis for concatenation. Although there is usually general agreement in the set of essential features (fundamental frequency, duration, power and phonetic context), it is still an open question the proper way of modelling them and considering their respective contributions to the cost functions, specially with regards to those related to the phonetic context. Precisely, this paper presents a new approach for modeling the phonetic context that also simplifies the hard task of training the corresponding weights to the different features in the target cost function.

Full Paper