

Efficient Voice Trigger Detection for Low Resource Hardware

Kisun You, Erik Marchi, Siddharth Sigtia, Rob Haynes, Hywel Richards, John Bridle

This talk will describe the architecture of an always-on keyword spotting (KWS) system for battery-powered mobile devices used to initiate an interaction with the device. An always-available voice assistant needs a carefully designed voice keyword detector to satisfy the power and computational constraints of battery powered devices. We employ a multi-stage system that uses a low-power always-on primary stage to decide when to turn on a main processor and run a more accurate (but more power-hungry) secondary detector or checker. We describe a straightforward (DNN/HMM) primary detector and explore variations that result in very useful reductions in computation (or increased accuracy for the same computation). By reducing the set of target labels from three to one per phone, and reducing the rate at which the acoustic model is operated, the compute rate can be reduced by a *factor of six* while maintaining the same accuracy.