Title: Using parallel tokenizers with DTW matrix combination for low-resource spoken term detection

Speaker: Wang Haipeng

Date: 08/Mar/2013
Time: 10:30am
Venue: SHIAE, Rm 702, ERB

Abstract:
Recently the posteriorgram-based template matching framework has been successfully applied to query-by-example spoken term detection tasks for low-resource languages. This approach employs a tokenizer to derive posteriorgrams, and applies dynamic time warping (DTW) to the posteriorgrams to locate the possible occurrences of the query term. Based on this framework, we propose to improve the detection performance by using multiple tokenizers with DTW distance matrix combination. The proposed approach uses multiple tokenizers in parallel as the front-end to generate different posteriorgram representations, and combines the distance matrices of the different posteriorgrams into a single matrix. DTW detection is then applied to the combined distance matrix. Lastly score post-processing techniques including pseudo-relevance feedback and score normalization are used for further improvement. Experiments were conducted on the spoken web search datasets of MediaEval 2011 and MediaEval 2012. Experimental results show that combining multiple tokenizers significantly outperforms the best single tokenizer, and that the DTW matrix combination method consistently outperforms the score combination method when more than three tokenizers are involved. Score post-processing techniques show further gains on top of using multiple tokenizers.