Improving Wordspotting Performance with Limited Training Data

by

Eric I-Chao Chang

Submitted to the Department of Electrical Engineering and Computer Science
on May 16, 1995, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis addresses the problem of limited training data in pattern detection problems where a small number of target classes must be detected in a varied background. There is typically limited training data and limited knowledge about class distributions in this type of spotting problem and in this case a statistical pattern classifier cannot accurately model class distributions. The domain of wordspotting is used to explore new approaches that improve spotting system performance with limited training data. First, a high performance, state-of-the-art whole-word based wordspotter is developed. Two complementary approaches are then introduced to help compensate for the lack of data. Figure of Merit training, a new type of discriminative training algorithm, modifies the spotting system parameters according to the metric used to evaluate wordspotting systems. The effectiveness of discriminative training approaches may be limited due to overtraining a classifier on insufficient training data. While the classifier's performance on the training data improves, the classifier's performance on unseen test data degrades. To alleviate this problem, voice transformation techniques are used to generate more training examples that improve the robustness of the spotting system. The wordspotter is trained and tested on the Switchboard credit-card database, a database of spontaneous conversations recorded over the telephone. The baseline wordspotter achieves a Figure of Merit of 62.5% on a testing set. With Figure of Merit training, the Figure of Merit improves to 65.8%. When Figure of Merit training and voice transformations are used together, the Figure of Merit improves to 71.9%. The final wordspotter system achieves a Figure of Merit of 64.2% on the National Institute of Standards and Technology (NIST) September 1992 official benchmark, surpassing the 1992 results from other whole-word based wordspotting systems.

Thesis Co-Supervisor: Richard P. Lippmann
Title: Senior Technical Staff

Thesis Co-Supervisor: David H. Staelin
Title: Professor of Electrical Engineering