Location-Dependent Vocabularies and Speaker Style Personalization for Accurate Mobile Diet Recognition

Xiaochen Huang and Emmanuel Agu
Computer Science Dept, Worcester Polytechnic Institute, Worcester, MA

I. INTRODUCTION

Obesity increases the risk of many health conditions and obesity rates have more than doubled since the 1970s [1]. Diet tracking is an important behavior change strategy for controlling obesity. However, manually recording foods eaten is tedious. Many diet tracking approaches have been proposed but have limitations. Recognizing food from photographs fails for complicated foods. Scanning food barcodes works but not all foods have barcodes.

Speech is one of the most natural methods of interaction. However, in practice, factors such as environmental noise and individual pronounciation styles result in low speech recognition accuracy. Our work focuses on accurately recognizing food orders as users order food at restaurants. Over 25 percent of Americans consume fast food daily [2]. We improve the accuracy and speed of speech recognition by exploiting two main concepts:

1. *Location-dependent speech recognizer vocabularies:* Just before a user orders food at a restaurant, we limit the vocabulary of "legal spoken words" (recognition range) by pre-populating the speech recognizer's corpus with only menu items from nearby restaurants (within 500 yards). Evaluations using 121 food items from Dunkin Donuts and McDonalds menus showed that limiting the recognizer's range improved accuracy (fig 1).
2. *Personalization:* Since the speech recognizer may pick up the speech of other customers and restaurant staff, we trained the recognizer on the speaking style of the smartphone owner.

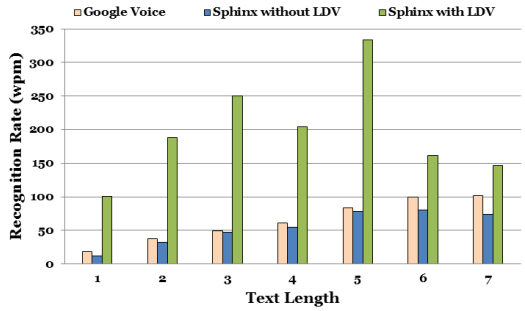

Figure 1. Speech recognition rate of CMU Sphinx vs Google Voice for different text lengths

II. THE DIETRECORD RESTAURANT ORDER RECOGNITION APP

Leveraging location-dependent speech vocabularies and recognizer personalization, we propose *DietRecord*, a smartphone app that automatically recognizes and records foods ordered by smartphone users at restaurants. DietRecord's interfaces also allows users to browse their food order history and food nutrition information. DietRecord inserts recognized foods into appropriate meals (breakfast, lunch, dinner, snack), based on the order's time. Fig 2 shows the DietRecord system architecture and figure 3 shows screenshots of DietRecord.

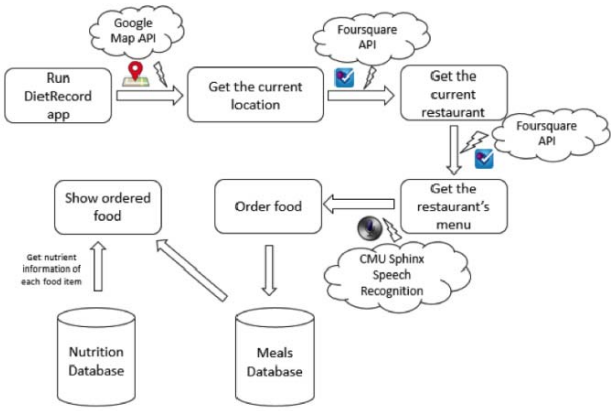

Figure 2. DietRecord System Architecture

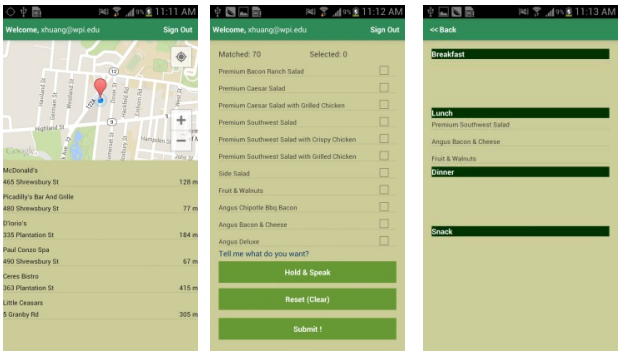

Figure 3. (a) Restaurant detection screen (b) Screen showing recognized foods (c) Screen showing nutrition information.

III. DIETRECORD EVALUATION

In a survey of 23 subjects, 82 percent of them were concerned about their daily calorie intake and over 50 percent of them wanted a mobile app for automatic diet tracking. DietRecord accurately recognized 86% of orders made by subjects who ordered 40 different items from McDonald's fast food restaurant. Recognition accuracy rate was mostly influenced by whether the word was a compound word (e.g. Big mac) or not and not by the length of the food name.

REFERENCES

[1] National Center for Health Statistics, http://www.cdc.gov/nchs/.
[2] Food Research and Action Center (FRAC), Overweight and Obesity in the US, http://frac.org/initiatives/hunger-and-obesity/obesity-in-the-us/.

978-1-4673-9548-9/15 $31.00 © 2015 IEEE
DOI 10.1109/ICHI.2015.67