Detecting Distraction

FINAL RESEARCH REPORT

Maxine Eskenazi, Alan W Black, Tim Keller, Ting-Yao Hu

Contract No. DTRT12GUTG11

Distracted driving has become a major cause of crashes and loss of life. While there is legislation prohibiting the use of cellphones while driving, people continue to use them. We reason that if we can create a system that automatically detects when a person is distracted and warns them (even shutting down an application if necessary), then some serious accidents could be prevented.

The goal of this project was to use the data collected in the Detecting Driver Distraction project previously funded by the Center in order to develop automatic distraction detection algorithms based on the characteristics of the subject's speech. The algorithms were trained on the places in the data that had previously been labeled as "distracted", using a set of OpenSmile [OpenSmile 2.1.0] features. Those features included prosodic items such as timing, pitch and intensity and acoustic ones such as spectral tilt. Speech recognition was not used here, nor was information from the gas pedal, the brake pedal or the steering wheel. OpenSmile enabled us to detect hesitations and fillers (sounds used to keep the floor in a dialog). The detector was implemented for a proof-of-concept email reading system that gracefully (remembering context) shut down when it detected that the user was distracted.

We have the following time-synced in the database:

- Driving conditions: hairpin turns, stop sign, red light
- Driving behavior: gas pedal, brake, steering wheel, speed, lane keeping
- Driver information: speech, gaze, head movement (microphone and backward facing camera)
- Visual distractions: signage (close and afar), turns, packages on the side of the road
- Cognitive distractions (additions to perceptual and memory load): emails, texts, phone calls
- Nature and content of the WoZ actions.

The cognitive distractions (email, phone calls, text messages) have been constructed to vary in both perceptual and memory load. For example, email varies in memory load from Mom's "how are you?" to a friend's "tell me five things you want for your birthday and I'll choose one and surprise you". From the perceptual load perspective, mails vary from Mom's mail to a friend asking, "How can I drive to downtown taking those street closings and one ways into account?", where the driver is likely to mentally picture the streets and/or a map while responding.

The detector was run over the speech of all 50 subjects in the database. Figure 1 shows the results for one of them. In this figure we see that there were six hand-labeled instances of distraction (this number varied from driver to driver). Although the detector performed correctly, it also showed several false positives; instances that we believe, after inspection, were related to other driver states, possibly caused by the memory load of the specific email they were dealing with.

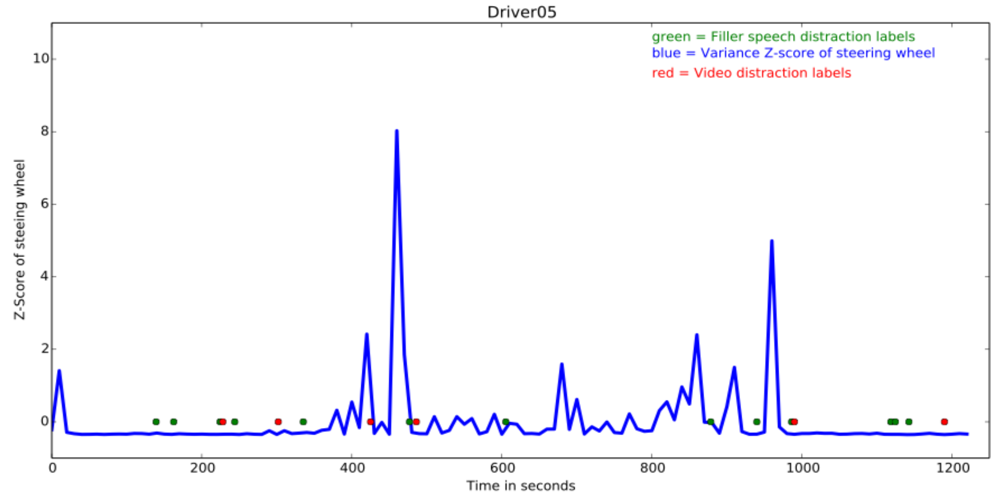

Figure 1. Distraction detection for *one driver* over a ~23 minute course in a driving simulator. The X axis is time in seconds, the Y axis is the Z-score of the change in steering wheel direction. The blue line indicates the use of the steering wheel (there were hairpin turns), the red dots are where an observer noted that the driver was distracted. The green dots (lighter grey in a black and white printout) indicate where the automatic distraction detector noted distraction.

We noted that the subjects had varying levels of tolerance of distraction. Figure 2 shows two aspects of driving behavior, use of the gas pedal and of the steering wheel. Here we see that when distracted, some users demonstrate irregular use of the gas pedal (GasPeaks, Subject 4) while others steer more erratically (Subject 1). Figure 3 shows that audio evidence alone did a good job of detecting where Subject 2 was distracted, but the same results did not hold for the other drivers shown here.

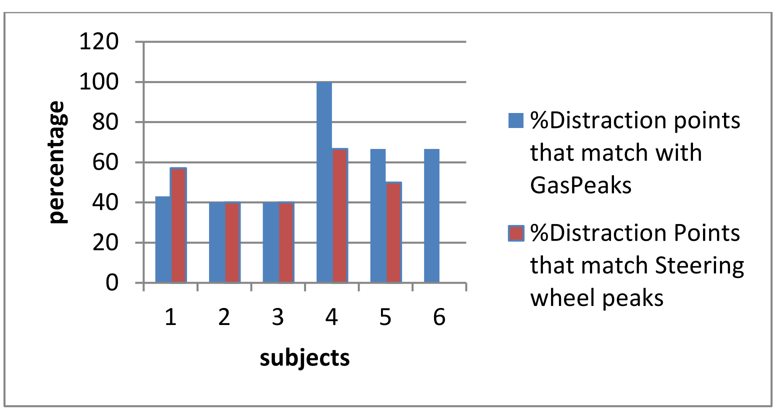

Figure 2. Distraction and individual behavior. The X axis shows six different subjects. The blue bars (on the left for each subject) show their irregular use of the gas pedal and the red bars show their irregular used of the steering wheel over the whole driving course. The Y axis is the percent of times that a point that was hand-labeled as a place where the driver was distracted coincided with either irregular gas pedal or steering wheel use.

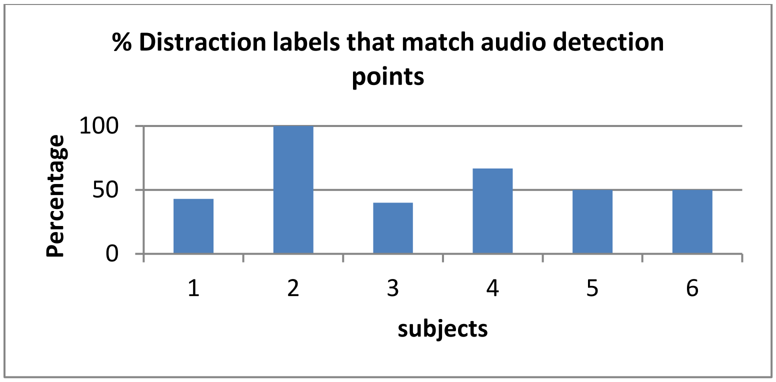

Figure 3. Automatic distraction detection and hand-labeled distraction. The X axis is for the same 6 subjects as in Figure 3. The Y axis is the percent of time that the automatic distraction detector correctly detected the hand labels.

Our results mainly helped us to determine what elements a distraction detection algorithm can use. They also highlighted the importance of modeling subjects individually due to the high variability of behavior between subjects. Finally, close inspection of the videos showed that the use of multiple modaliites rather than just the speech modality would give the detection algorithms more precision.

Milestones

- implementation of distraction detection algorithms based on speech from the driver

Products

- a set of distraction detection algorithms

References

OpenSmile https://audeering.com/technology/opensmile/

DISCLAIMER