Participatory Sensing of Cycling Path Safety with Wearable Devices

Kalynn Kosyka[1] | Josh Cherian[2] | Daniel Goldberg[2,3] | Tracy Hammond[2] | Jennifer Horney[4]

[1]Smith College; [2]Sketch Recognition Lab, Texas A&M University; [3]Department of Geography, Texas A&M University; [4]Department of Epidemiology and Biostatistics, Texas A&M University

Introduction

The number of people bicycling in the United States is increasing including commuters. In 2015, the US had 45,000 bicyclists injured with 818 fatalities including 50 in Texas. This represents a six percent increase since 2006.

In 2012, the three major causes of injury were: 1) being hit by a car (29%), 2) falling (17%), 3) roadway or walkway being in disrepair (13%) costing over 4 billion dollars. Studies comparing the potential effects of cycling on mortality have concluded on average, the estimated health benefits of cycling were substantially larger than the risks of cycling relative to driving.

As road quality is the third major cause of injury to bicyclists, this study developed a machine learning classification technique capable of correctly classifying road quality through the use of wearable devices.

Methods

Pebble smartwatches were affixed to the seat post of the bicycle and 2nd Generation **Moto X** smartphones were attached to the back pocket of bicyclists

An **Android application** recorded accelerometer (XYZ) and GPS data from each device during each ride with a total of 117 minutes and 7,020 windows

A total of ten subjects rode on four different surfaces (bricks, bumpy, grass and smooth) at TAMU

A **machine learning** algorithm (J48) was used to train and extract a decision tree to classify road condition through 3-axis acceleration.

Fifty-two features were developed with the window size of 1 second and 0.5 second overlap

Data was divided for training and testing per device

NSF Project# 1560106

TAMU REU SUMMER SYMPOSIUM, 2017

Analysis

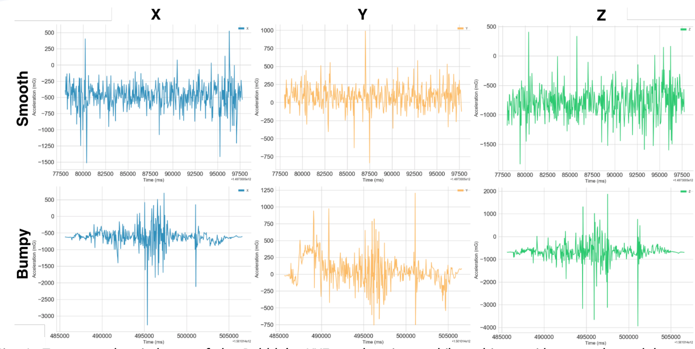

Fig. 1. Two second windows of the Pebble's XYZ acceleration while subjects ride smooth and bumpy roads

Features Extracted for Phase 1

Pebble	Phone
Var. (X,Z)	Var. (Y,Z)
Energy (X)	Avg. Height (Y,Z)
Avg. (X,Y,Z,XY,XZ)	Energy (Y,Z)
Corr. (YZ)	Avg. (X,Z,XY,XZ,YZ)
RMS (X,Z)	Std. Dev. (Y)
Avg. Peaks (X,Z)	Corr. (XZ)
Avg. Valleys (X,Z)	RMS (Z)
	Avg. Peaks (Y,Z)
	Std. Dev. Peaks (Y)
	Avg. Valleys (Z)
	Num. Points

Phases

Phase 1

J48 – a C4.5 algorithm created a decision tree based on the features extracted from all riders

Phase 2

Developed an algorithm to go through a list of classified windows and fix those mislabeled by iterating through them and grouping them in a dynamically-sized window

Phase 1

Train	a	b	b	d
a = bricks	0.933	0.005	0.011	0.055
b = bumpy	0.032	0.894	0.032	0.042
c = grass	0.012	0.042	0.892	0.054
d = smooth	0.009	0.003	0.002	0.987

Table 1. Confusion matrix of the Pebble training data after the decision tree classification in Phase 1

Test	a	b	b	d
a = bricks	0.740	0.036	0.009	0.215
b = bumpy	0.112	0.462	0.161	0.266
c = grass	0.073	0.168	0.565	0.194
d = smooth	0.089	0.065	0.044	0.801

Table 2. Confusion matrix of the Pebble testing data after Phase 1, using decision tree generated from training data

Phase 2

Test	a	b	b	d
a = bricks	0.970	0	0	0.030
b = bumpy	0.094	0.615	0.177	0.115
c = grass	0.008	0.0313	0.938	0.023
d = smooth	0	0	0	1

Table 3. Confusion matrix after the Pebble testing data from Phase 1 went through Phase 2

Fig. 2. The four road surfaces used in study

Results

Phase 1: *J48 with 10-fold cross validation generated tree from training data*

- Pebble training data yielding an accuracy of 92.64% and f-measure of 0.967, whereas for testing data yielded accuracy of 63.93% and f-measure of 0.778
- Phone training data yielding an accuracy of 97.25% and f-measure of 0.982, whereas for testing data yielded accuracy of only 33.14% and f-measure of 0.272

Phase 2: *Further improve classification through iterating data classified from phase 1*

- **Pebble testing data yielded accuracy of 88.06% and f-measure of 0.974**
- Phone testing data yielded accuracy of only 27.95% and f-measure of 0.257

Conclusions

Smart watch yield more accurate surface identification than smartphones. The developed machine learning algorithm was capable of **correctly classifying** among four surface types with an **accuracy of better than 88%** by smart watch.

Future Work

Our results demonstrate that we can effectively classify road surface through smart devices.

In future work, we hope to be able to create a mobile application that allows many users to share their accelerometer data which will then be labeled on a map displaying road surface quality. This will allow fellow cyclists to determine the best routes for user's comfort, as well as draw attention to areas that need infrastructure repair.