# Poster: Gait-based Smartphone User Identification

Matthew Boyle
mboy@bu.edu

Avraham Klausner
aviclaws@bu.edu

David Starobinski
staro@bu.edu

Ari Trachtenberg
trachten@bu.edu

Hongchang Wu
wuhc@bu.edu

## Categories and Subject Descriptors

I.5.4 [**Computing Methodologies**]: Pattern Recognition—*signal processing*

## General Terms

Experimentation, Human Factors, Verification

## 1. INTRODUCTION

Biometric signatures based on a user's physical features form a natural approach to authentication, as they are often high in entropy and hard to forge or forget. Gait, the rhythmical body movements of human beings as they walk, is a particularly appealing biometric because of its distinguishability [2]. In this work, we provide a proof-of-concept implementation of gait-based user identification using the tri-axial accelerometer sensor of the Motorola Droid phone. It is envisioned that this means of user identification could be utilized as a component of a larger recognition system, one that might fully amalgamate all relevant sensed data to generate an idea of whether a legitimate user has possession of the phone at any given instant. [1]

## 2. METHODOLOGY

Our system draws raw $x$, $y$, and $z$-axis accelerometer readings from the phone and computes the Euclidean norm of the readings, removing orientation dependence. Linear interpolation is used to provide the regular sampling intervals required by our discrete signal processing. A wavelet transform [1] is then applied to the raw data to extract useful features in the time and frequency domains.

Wavelet analysis decomposes a signal into approximate spaces and detail spaces. For our analysis we use the Daubechies 4 tap wavelet family and a 4-level multi-resolution decomposition. This allows the data to be described by five arrays of numbers; we then use the Euclidean norm of each array as a component in a five-dimensional feature vector to represent the unique characteristics of a particular user's gait.

Our system uses the $k$ Nearest Neighbors ($k$-NN) algorithm as a classifier, which computes the five-dimensional feature vectors of any newly acquired data and calculates the $k$ closest (in terms of five-dimensional distance) feature vectors originating from any provided learning data. A majority vote using the $k$ nearest feature vectors' sources determines the probable source of any newly acquired data.

[1]This work was funded in part by BU UROP and the NSF under grants CCF-0729158, CCF-0916892, and CNS-1012910. All authors are with the BU ECE department.


*MobiSys'11,* June 28–July 1, 2011, Bethesda, Maryland, USA.
ACM 978-1-4503-0643-0/11/06.

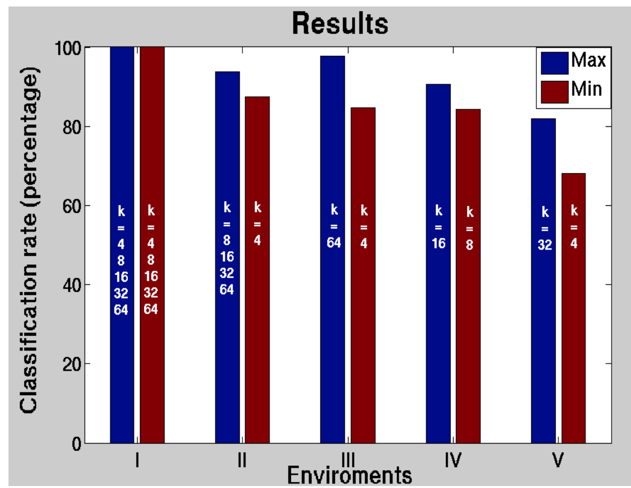


**Figure 1: Results of the five experiments.**

## 3. RESULTS

Our system is implemented and run for different values of $k$, i.e., $k = 4, 8, 16, 32$, and 64. Fig. 1 depicts the results of each of the experiments, with the max and min bars respectively representing the highest and lowest classification rates obtained over all tested values of $k$. In Environment I, the system is tested against two users walking a straight path at 1.57 m/s. Environment II involves two users walking straight at 1.18 m/s. Two users walk at 1.18 m/s in square and circular paths in Environment III. Environment IV is a classification against four users, all walking in a straight path at 1.18 m/s. In environment V, the two users' speeds are inconsistent between trials, which explains the drop in classification ability (our system makes little distinction between "different speeds" and "different people").

## 4. REFERENCES

[1] S. G. Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, second edition, 1999.

[2] M. S. Nixon, T. Tan, and R. Chellappa. *Human Identification Based on Gait.* Springer, 2005.