RUTGERS THE STATE UNIVERSITY OF NEW JERSEY

Open Ears

Peter Lin, Stephen Shanko, Chris Slezak

{peter.lin, stephen.shanko, chris.slezak}@rutgers.edu

Advisor: Dr. Athina Petropulu

Goal

- ❑ To create an Android Application to alert distracted users of incoming traffic
- ❑ To make it safer for pedestrians to listen to music or to use their mobile phones while walking outside

Motivations and Objectives

❑ **Motivations**

- Modern technology has become very distracting
- According to the CDC, 4,280 pedestrians were killed in traffic accidents in 2010 and 70,000 more were injured

❑ **Objectives**

- Correctly classify samples from a microphone as incoming traffic
- Provide the user with useful feedback to alert them that a car is approaching

Research Challenges

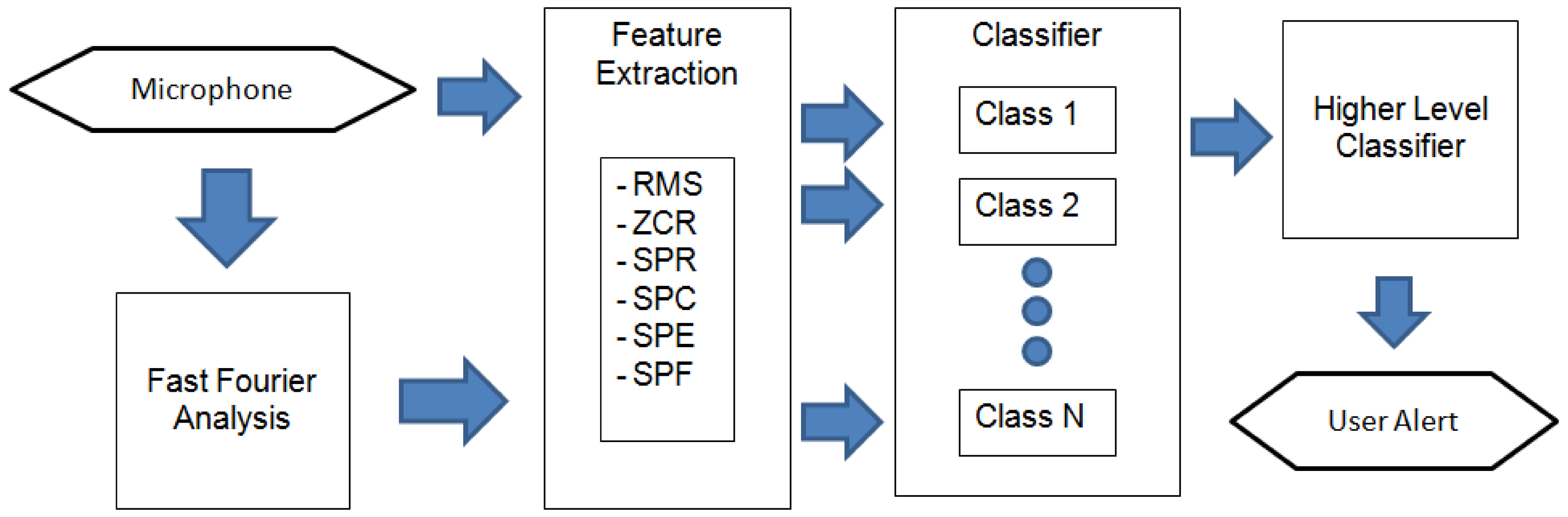
- ❑ Selection of a kernel function that offers increased performance and robustness
- ❑ Deciding on parameters: Audio frame length, Sparsity level, and the size of each dictionary
- ❑ Implementation of the KOMP algorithm and higher level classifier in Android

Acknowledgement

We would like to thank Dr. Petropulu for her guidance and support throughout the project.

Methodology

- ❑ Android device microphone is always recording
- ❑ Features are extracted from each frame
- ❑ Energy and zero crossing rate in time domain
- ❑ Spectral roll-off, centroid, entropy, and flux in frequency domain
- ❑ Classes are represented by dictionaries learned in MATLAB and transferred to the phone
- ❑ Residual for each class is calculated using KOMP and used to classify each input
- ❑ Higher level classifier decides when to alert



Input: A signal $\mathbf{z}$, a kernel function κ, $\mathbf{A}$, and a sparsity level T_0.
Task: Find a coefficient vector $\mathbf{x} \in \mathbb{R}^K$ with at most T_0 non-zero coefficients such that $\Phi(\mathbf{Y})\mathbf{A}\mathbf{x}$ approximates $\Phi(\mathbf{z})$.
Initialize: $s = 0$, $I_0 = \emptyset$, $\mathbf{x}_0 = \mathbf{0}$, $\hat{\mathbf{z}}_0 = \mathbf{0}$
Procedure:

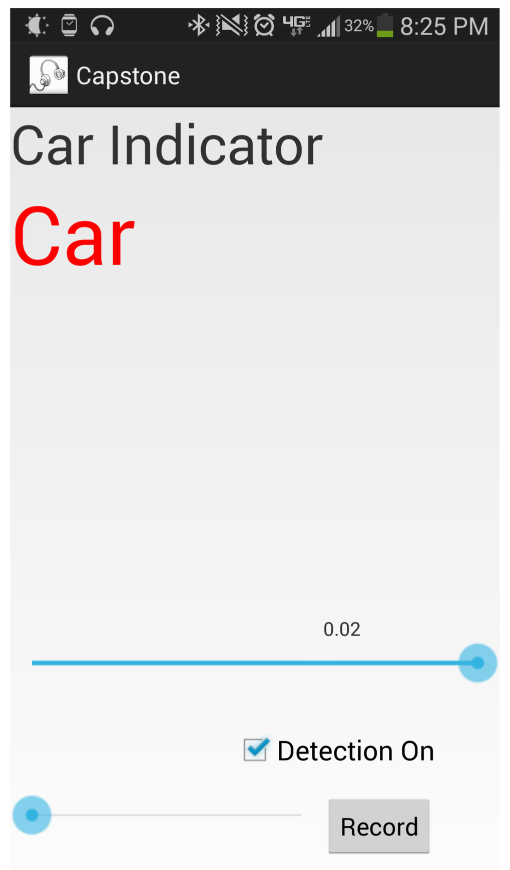
1. $\tau_i = \left(\mathbb{K}(\mathbf{z},\mathbf{Y}) - \hat{\mathbf{z}}_s^T \mathbb{K}(\mathbf{Y},\mathbf{Y})\right)\mathbf{a}_i, \ \forall i \notin I_{s-1}$
2. $i_{max} = \arg\max_i |\tau_i|, \ \forall i \notin I_{s-1}$
3. Update the index set $I_s = I_{s-1} \bigcup i_{max}$
4. $\mathbf{x}_s = \left(\mathbf{A}_{I_s}^T \mathbb{K}(\mathbf{Y},\mathbf{Y})\mathbf{A}_{I_s}\right)^{-1} (\mathbb{K}(\mathbf{z},\mathbf{Y})\mathbf{A}_{I_s})^T$
5. $\hat{\mathbf{z}}_s = \mathbf{A}_{I_s}\mathbf{x}_s$
6. $s \leftarrow s + 1$; Repeat steps 1–6 T_0 times

Output: Sparse vector $\mathbf{x} \in \mathbb{R}^K$ satisfying $\mathbf{x}(I_s(j)) = \mathbf{x}_s(j), \forall j \in I_s$ and zero elsewhere.

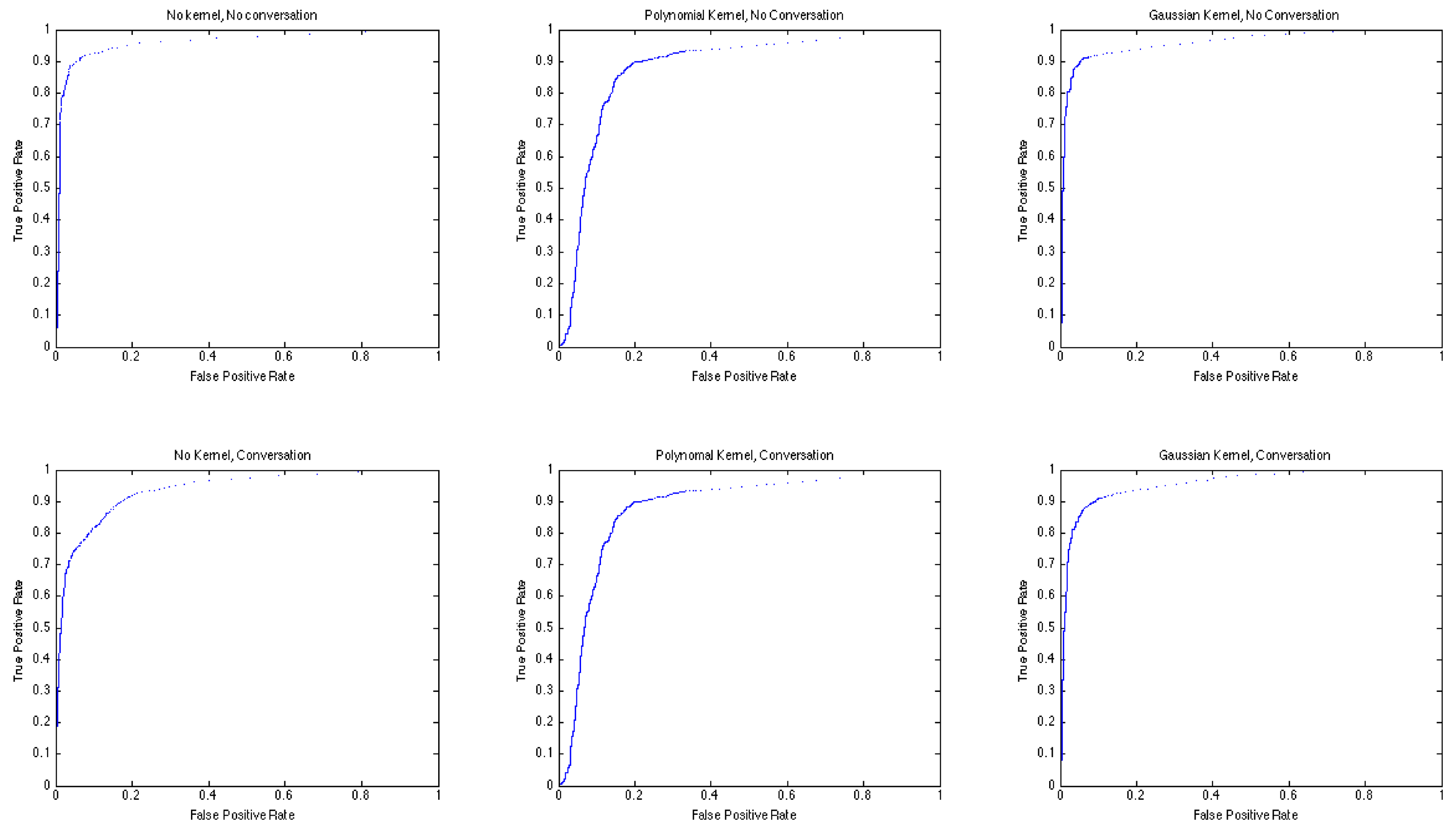
Fig. 2. The KOMP algorithm.

Results

- ❑ App performance was tested on Hoes Lane West with a sample of 21 cars
- ❑ The following results were obtained
- ❑ 20 detections
- ❑ 1 miss
- ❑ 8 false detections
- ❑ Due to the critical safety risks involved, we preferred a false positive to a false negative



- ❑ Performance was tested in MATLAB for a variety of kernels and scenarios. Use of a kernel made the system more robust when the user is having a conversation



References

[1] Nguyen, H.V.; Patel, V.M.; Nasrabadi, N.M.; Chellappa, R., "Kernel dictionary learning," Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on , vol., no., pp.2021,2024, 25-30 March 2012

[2] Smaldone, Stephen, et al. "Improving Bicycle Safety through Automated Real-Time Vehicle Detection." (2010)