A Catchy Title

John Smith
Graduate School of AI
KAIST
Daejeon, South Korea
alice@kaist.ac.kr

Mary Thompson
School of EE
KAIST
Daejeon, South Korea
bpb@kaist.ac.kr

Michael Miller
School of Computing
KAIST
Daejeon, South Korea
carol@kaist.ac.kr

ABSTRACT

How similar are two sound-clips? what is the best rhetorical question you can ask? In this project, we develop *SomeMETHOD*, a fast and effective way of measuring the similarity between two short sound clips.

1. INTRODUCTION

Specify the problem; Give the motivation; List your main contributions

The problem we want to solve is the following:

- GIVEN: a collection of N sound clips, of similar duration, and each having a class label among k=5 classes
- FIND: a clip-to-clip similarity function
- to MINIMIZE: the classification error, in the 1-nearest-neighbor classifier.

This is an important problem, because ... millions of dollars ... millions of human lives ...

The contributions of this project are the following:

- our proposed *someMETHOD* is novel, combining wavelets with a spike-removal preprocessing step
- it is effective, achieving 90% classification accuracy
- it is scalable, being linear on the number of sound-clips N.

2. PROPOSED METHOD

The main motivation behind our method is to handle spikes carefully. Since the input signals are noisy, with bursty noise, traditional methods like time-warping and wavelets will focus on the spikes, and ignore the rest of the signal, giving misleading results.

Our proposed method is as follows: We will use the spike-detection method of ..., to remove spikes, and only then, we use the k strongest Daubechies-4 wavelet coefficients, to compare the two (spike-removed) sound clips.

...

3. EXPERIMENTS

We implemented our method and compared it with the 10 other ones. The results are very promising.

Figure 1 shows our results: Figure 1(a) gives a scatter-plot of the N sound-clips, where the axis are the two main features we propose to use ... Figure 1(b) shows the wall-clock time of our method, versus the size of the database N.

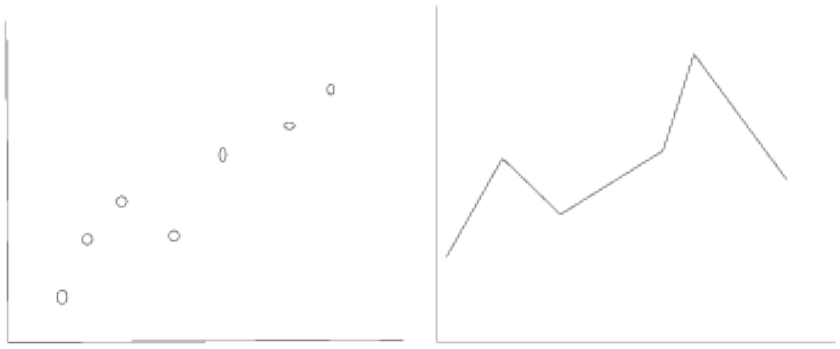
4. CONCLUSIONS

The proposed method *someMETHOD* has the following advantages:

- it gives better classification accuracy than all 10 competitors we tried
- its accuracy is very close to the very best competitor in the *UCR Insect Classification Contest.*
- it is scalable

5. REFERENCES

.



(a) (b)

Figure 1: A fictitious dataset (a) and its performance plot (b)

APPENDIX

A. APPENDIX

A.1 Labor Division

The team performed the following tasks

- Implementation of Daubechies-4 [Smith, Thompson]
- Comparison of Daubechies-4 against euclidean distance [Miller]
- Data collection [all]
- Experiments on the real data [Miller]

A.2 Full disclosure wrt dissertations/projects

A.2.0.1 Smith:.

His dissertation is on a music retrieval system ('query by whistle'). Although related to this class's project, Smith never considered wavelets, AutoRegression, or generalized-time-warping, for his dissertation, that he studied and implemented in this project.

A.2.0.2 Thompson:.

She is not doing any project or dissertation related to this project: her thesis is on phylogenetic trees.

A.2.0.3 Miller:.

He is not doing any project or dissertation related to this project: his thesis is on dark matter discovery.