LIPSCOMB UNIVERSITY
RAYMOND B. JONES SCHOOL OF ENGINEERING

EECE 3841 Speaker Detection Final Project

By Faris Jugovic, Kristopher Pesnell, Jose De Los Santos

LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Introduction

- Determine if a voice belongs to a nominal user
- Uses FFT and RMSE for detection
- Lightweight, no ML required

LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Method Summary

- FFT transforms voice recordings
- Use files 1–3 to generate average profile
- Compare file 4 using RMSE
- RMSE < 0.005 ⇒ Match

LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

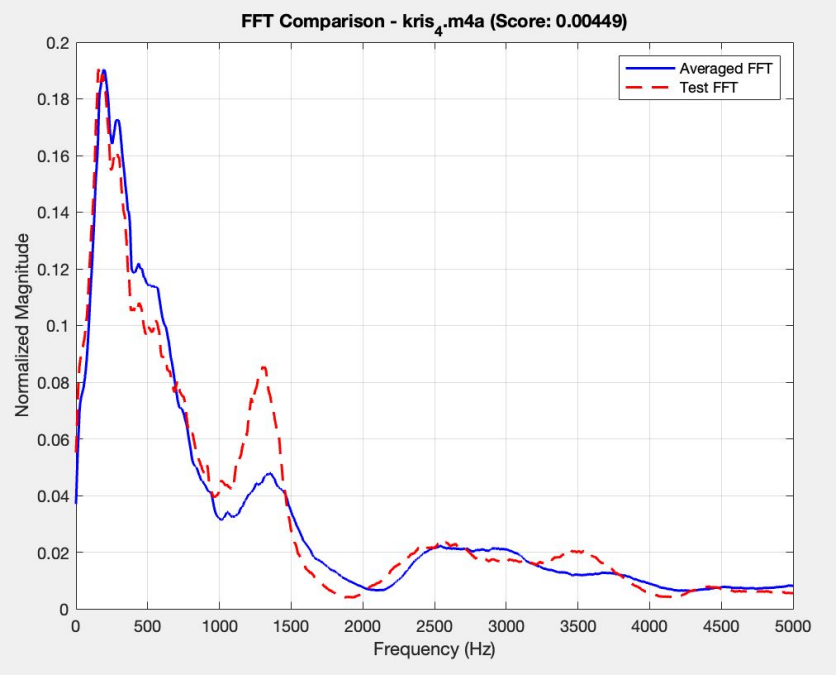
Methodology: Data Collection

- 14 users: 10 nominal, 4 imposters
- Each provided 5 recordings
- First 3 recordings used for user profile
- Consistent hardware and sentence
  - The arsonist has oddly shaped feet and the human torch was denied a bank loan.

LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Methodology: Input Processing

- FFT
- Normalize
- Truncate
- Average
- Smooth
- Compare



LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Methodology: Algorithm

- Create user profile: avg(FFT of files 1–3)
  - FFT → Normalize → Truncate → Average → Smooth → Compare
- Compare sample to user profile
- Return match if RMSE < threshold

LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Training, Validation, and Testing

- Training: files 1–3 per user
- Validation: 4th audio file
- Testing: 5th audio file
- Fully automated in script

Pseudocode

```
For each user:
    avg_profile = average(FFT(file1), FFT(file2), FFT(file3))

    For each test_file in [file4, file5, imposters]:
        processed_test = normalize(smooth(FFT(test_file)))
        score = RMSE(processed_test, avg_profile)

        if score < threshold:
            match = true
        else:
            match = false
```

Pseudocode

```
Set configuation
For each_user [Test Users for self recognition]
For each_user [Test User against each imposter]
Print Results
```

LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Threshold Tuning

- Threshold sweep: 0.0015–0.005
- 0.005 chosen for best results
- Lower ⇒ false rejects, higher ⇒ false accepts

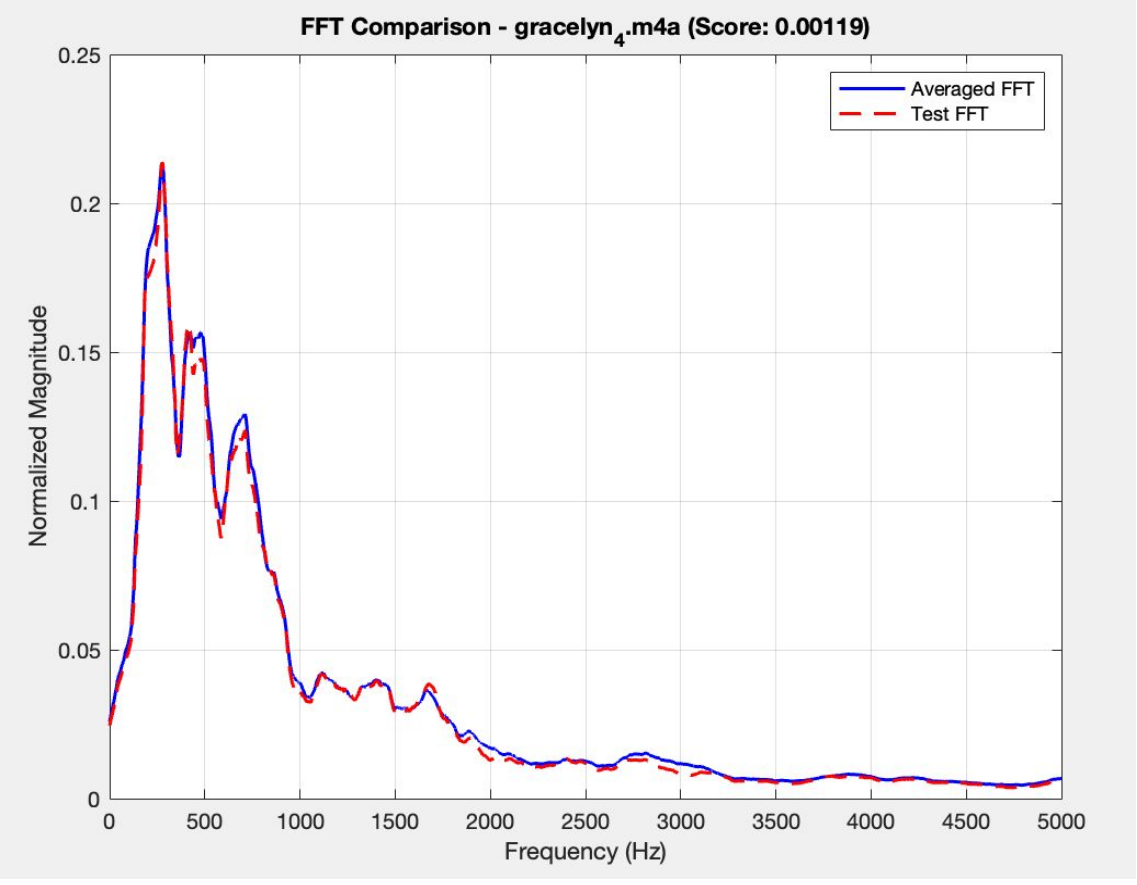
LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Final Results

- True Positives: 10
- False Negatives: 0
- False Positives: 0
- True Negatives: 40
- Precision: 100%
- Recall: 100%

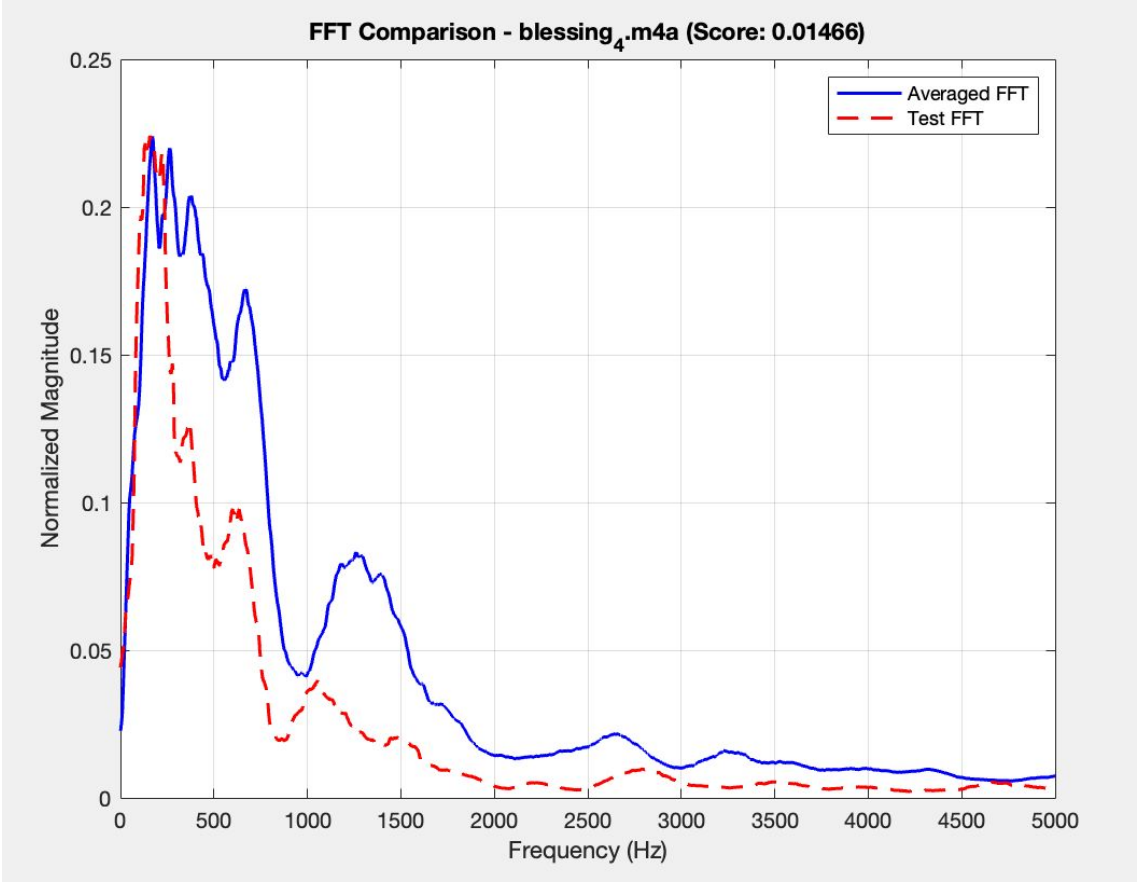
LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Match Case



LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Rejected Case



LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Runtime Analysis

- Total runtime: 10.27s
- test user.m: 0.21s per call
- Profile generation: 0.18s
- FFT comparison: 0.05s

LIPSCOMB UNIVERSITY
RAYMOND B. JONES
SCHOOL OF ENGINEERING

Discussion and Lessons Learned

- Strengths: efficient, accurate, simple
- Weaknesses: sensitive to noise and users with similar FFT’s
- Lessons: preprocessing and tuning are critical