EE113DB 2015 FINAL PRESENTATION VOICE RECOGNITION AND CONTROL SYSTEM

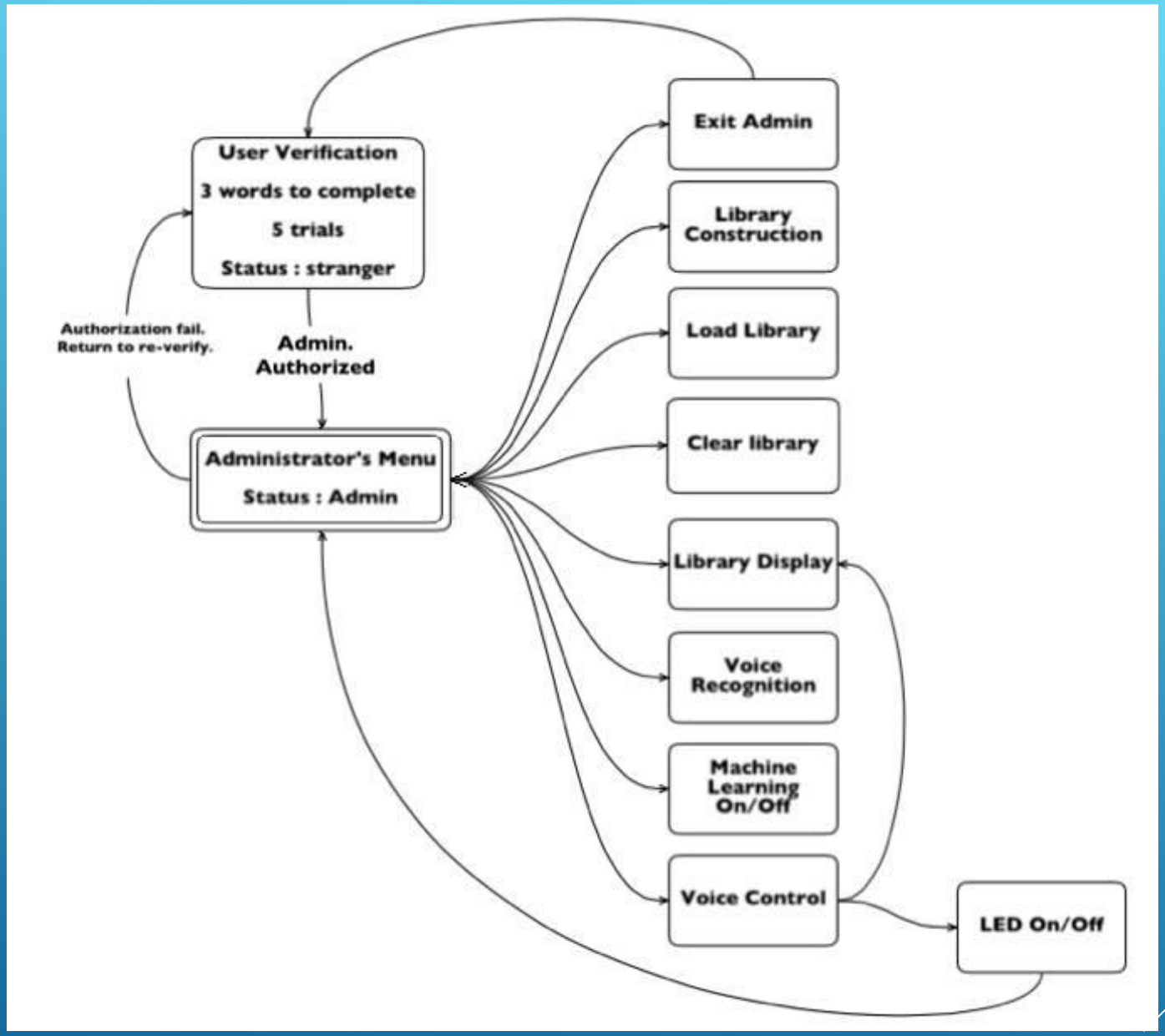
Present By:

- George Li (Zhuoqi Li)
- Ben kim (Bumjoong Kim)

SYSTEM OVERVIEW

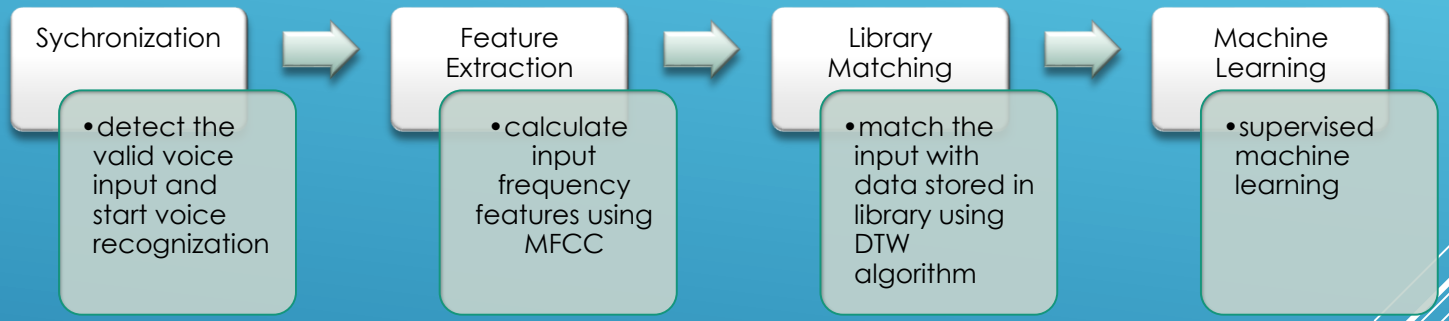
- Goal:
  - interpret spoken words
  - differentiate different people's voices
  - operate according to spoken commands

- Objective of this project: to accurately identify spoken vocabularies as well as their speakers from a 70-word database and operate as told

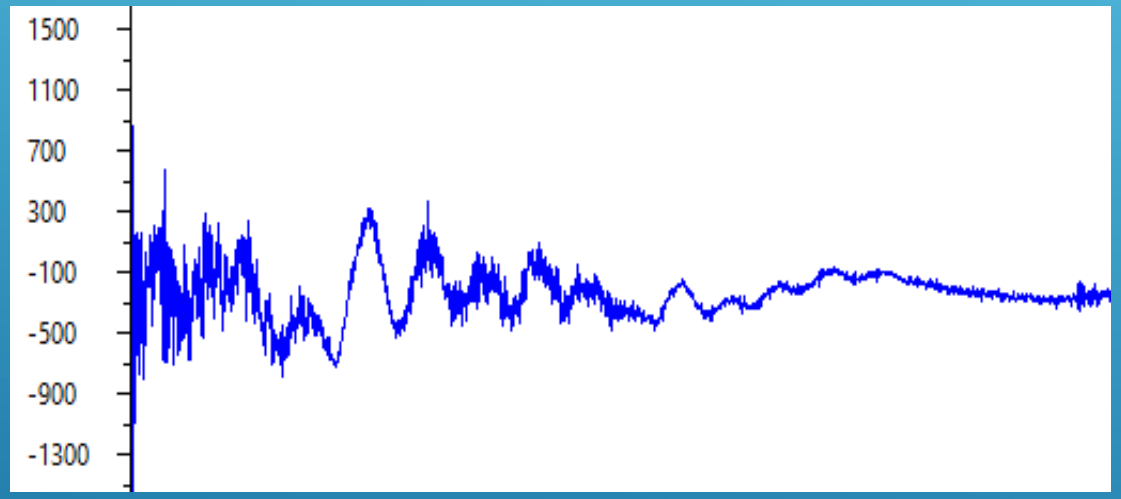

User Verification
3 words to complete
5 trials
Status : stranger
Authorization fail.
Return to re-verify.
Admin.
Authorized
Administrator's Menu
Status : Admin
Exit Admin
Library Construction
Load Library
Clear library
Library Display
Voice Recognition
Machine Learning On/Off
Voice Control
LED On/Off

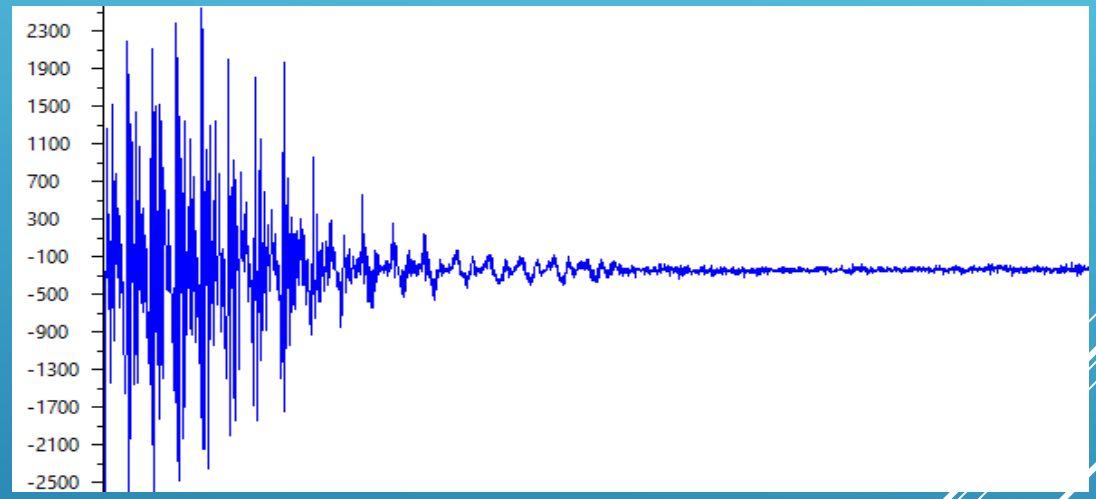
TECHNICAL DESIGN

- Primary principle behind: to model human hearing system
- Based on EE113DA Miniproject 2
- Voice recognition process:
  - Synchronization
  - Feature extraction
  - Library Matching
  - Machine learning
- Voice control process:
  - conditional statements



- Differentiate between environmental noise and human speech
- Utilize the STE (short time energy algorithm)





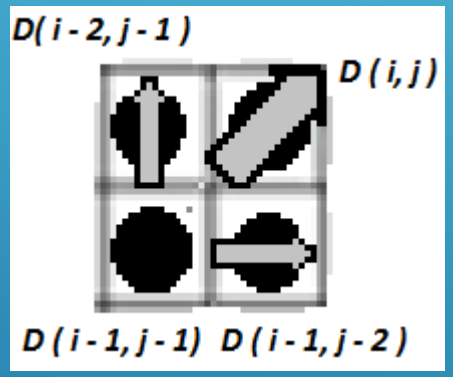
SYNCHRONIZATION

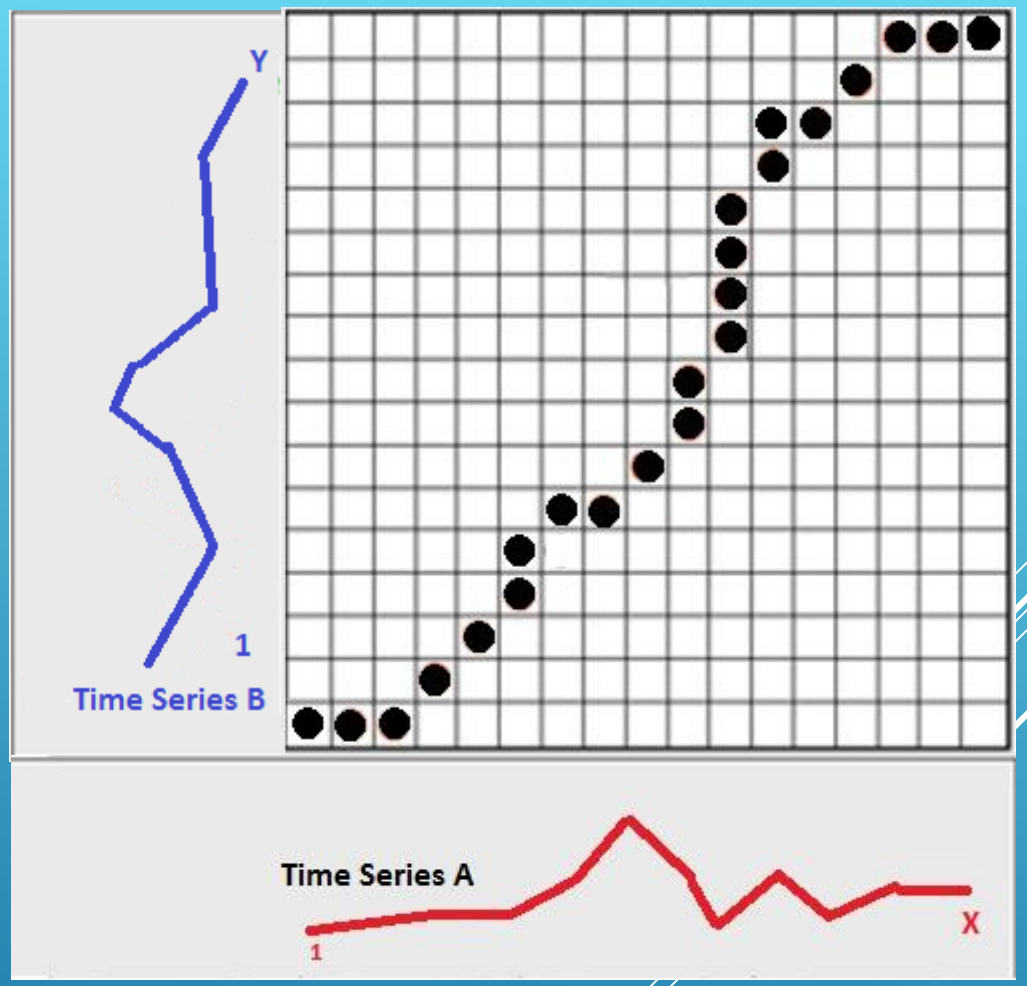
FEATURE EXTRACTION

- Mainly based on EE113DA Miniproject 2
- Mel Frequency Cepstral Coefficient + Delta-delta = WMFCC
- For each speech input, break it into many frames
- Power spectrum information in different frequency range

LIBRARY MATCHING

- Library data saved in txt files
- Library data loaded from txt files when needed
- DTW (Dynamic Time Wrapping)





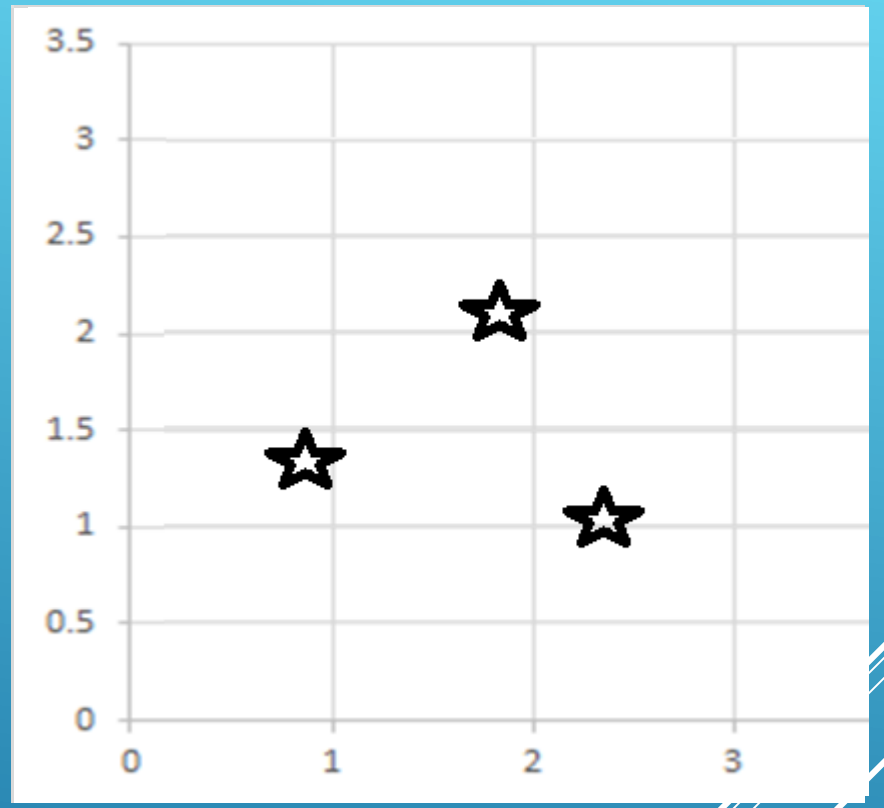
MACHINE LEARNING

* Supervised machine learning algorithm
- user judge the result and give feedback

* K-mean cluster algorithm

* Update the library data every time

* Can be turned off when the library is fully updated

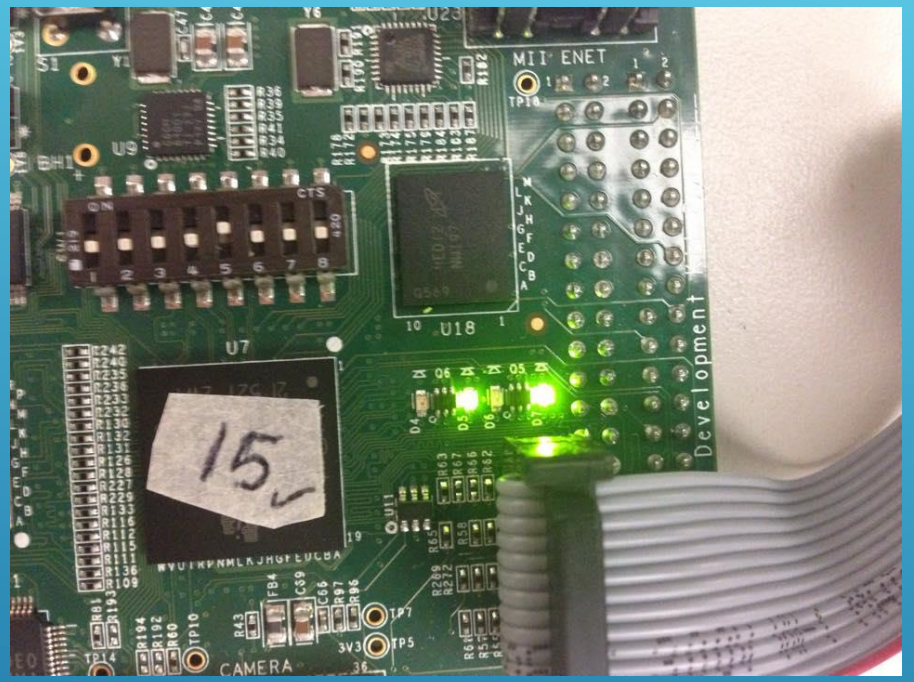


- Display Library
- Turn on LED on the DSP chip

Console | Available Products | Advice

voice_recongization_project:CIO

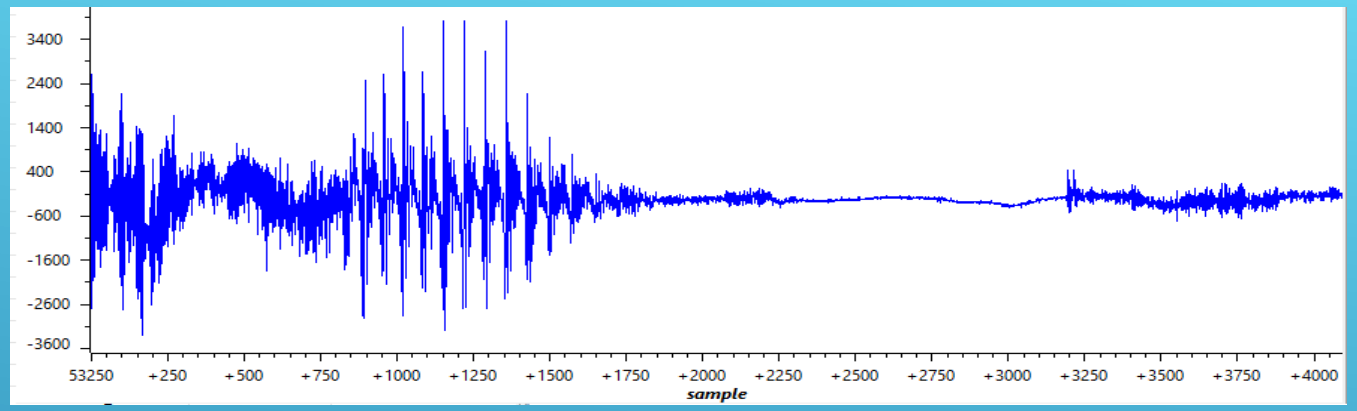
```
Finished

Your sentence is:
two Crazy cough Display Library Finished Library slot #       Vocabulary        Speaker        Correction
      1                 Cat                 Ben            5
      2                 all                 Ben            1
      3                 No                 Ben            2
      4                 Fit                 Ben            3
      5                 Cute                 Ben            2
      6                 Crazy                 Ben            2
      7                 share                 Ben            9
      8                 sky                 Ben            2
      9                 cough                 Ben            8
      10                 thank                 Ben            1
      11                 Cat                George            7
      12                 No                George            6
      13                 All                George            4
      14                 Cute                 George            5
      15                 Fit                George            9
      16                 Share                 George            4
      17                 Thank                 George            5
      18                 Crazy                 George            9
      19                 Sky                George            3
      20                 Cough                 George            3
      21                 turn                ben            4
      22                 LED                ben            1
      23                 on                ben            2
      24                 off                 ben            2
      25                 one                 ben            3
      26                 three                 ben            2
      27                 two                 ben            3
      28                 four                 ben            1
      29                 Finished                 ben            1
      30                 Display                 Ben            3
      31                 Library                 Ben            0
```



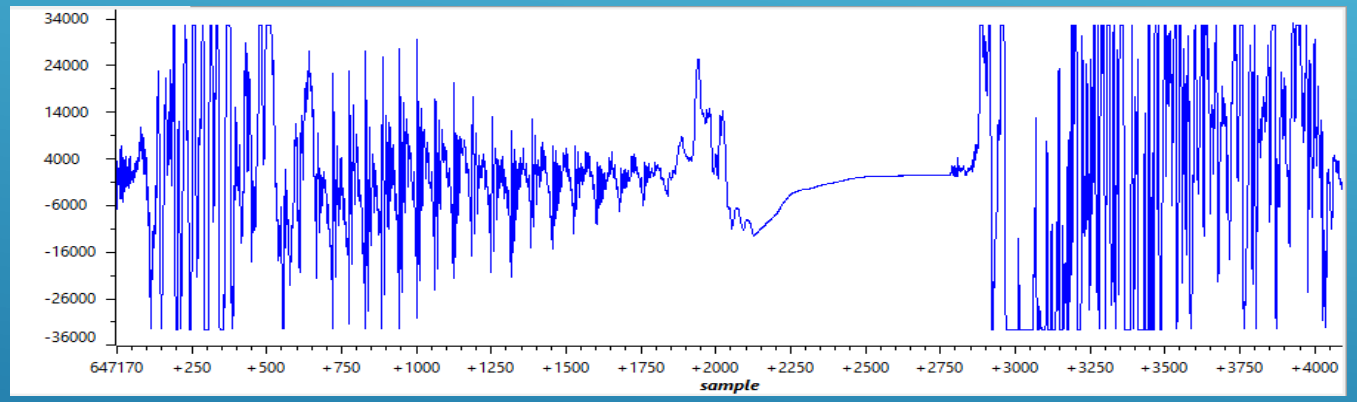
VOICE CONTROL

TECHNICAL & GENERAL ISSUES

- Coding challenges
- Equipment uncertainty
- Processing speed optimization
- Uncertain nature of human voice
  - intonation effects
  - Volume of the voice
  - length of the speech

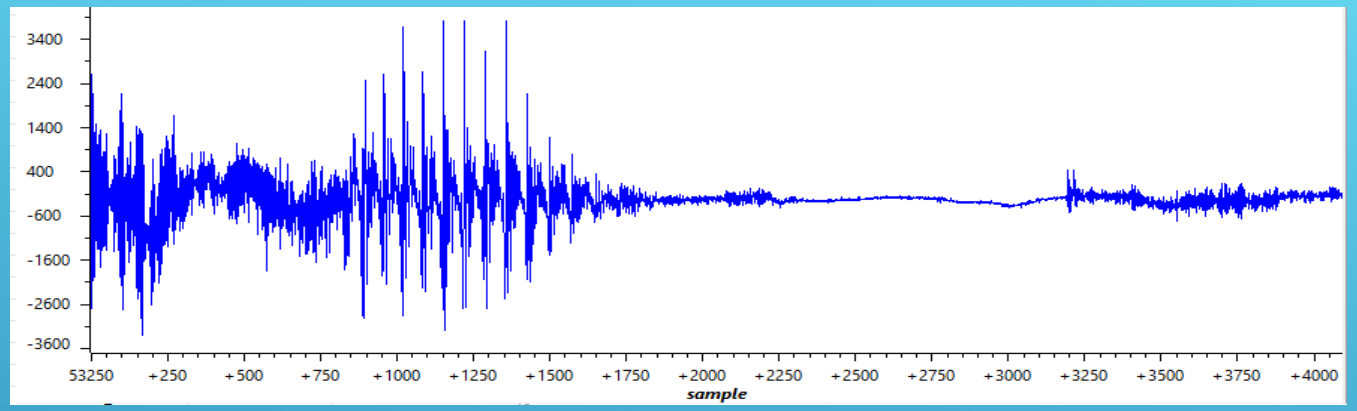


‘Cat’ with normal volume

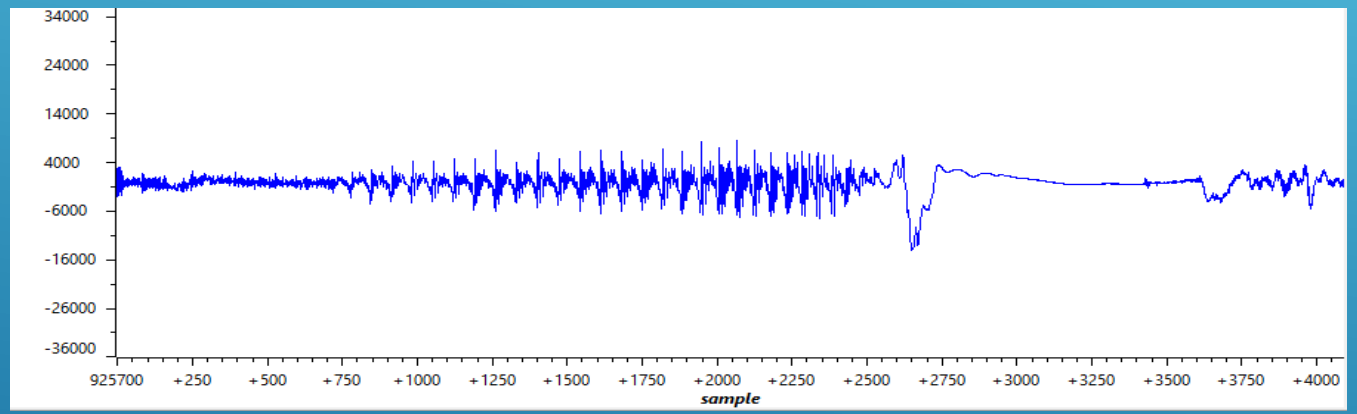


‘Cat’ with high volume

TECHNICAL & GENERAL ISSUES – SPEECH VOLUME



‘Cat’ with exclamatory intonation



‘Cat’ with interrogative intonation

TECHNICAL & GENERAL ISSUES - INTONATION

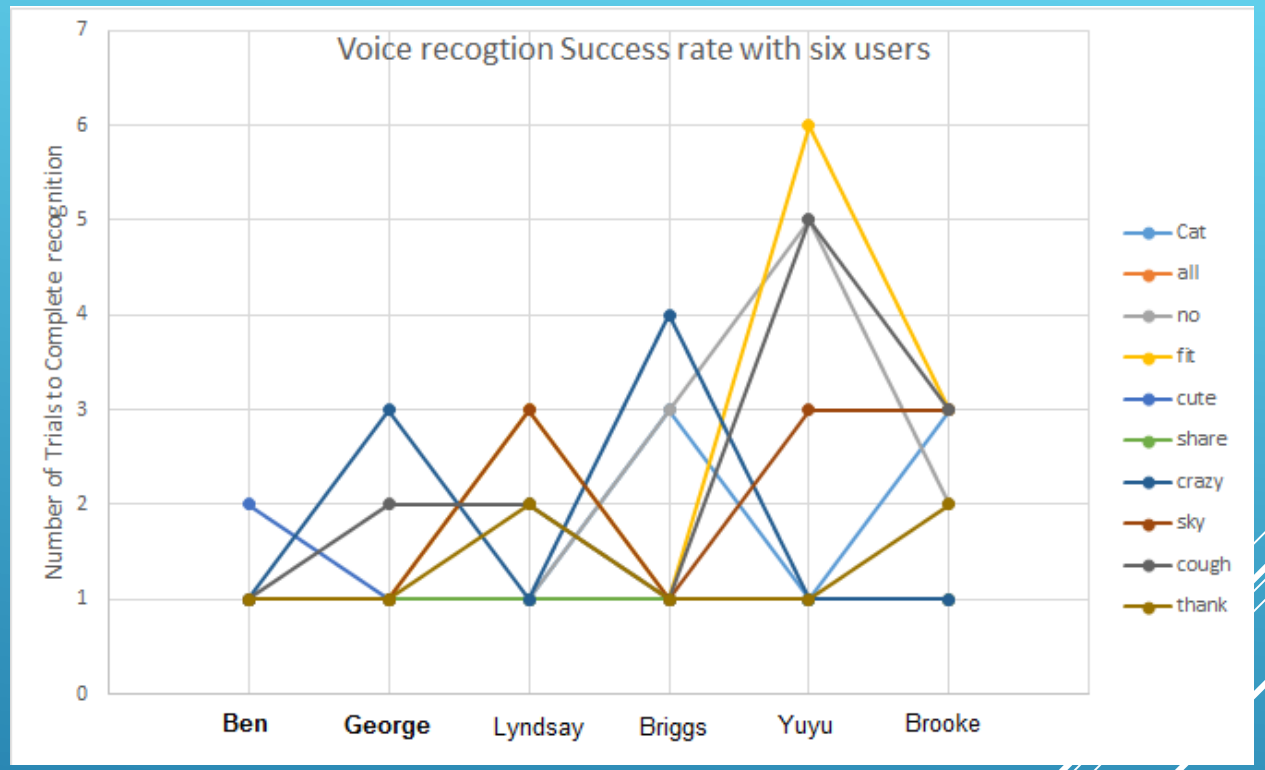
EXPERIMENTAL RESULT

- Experimental procedure
  - Words with clear vowels: cat, no, all fit, cute
  - Words with obvious consonants: share, crazy, sky, cough, thank
- Experimental results:

	Version 1.3	Version 1.4	Version 1.8
Total Recognition rate	65.0%	77.5%	85.7%
Speaker Recognition rate	75.0%	87.5%	85.7%
Vocab Recognition rate	90.0%	87.5%	100.0%
Trial times	40	40	42

Optimization level	Processing Time (s)
0	22.5
1	14.5
2	5.45
3	4.5

Ver 2.0 (05192015)	Ben	George	Lyndsay	Briggs	Yuyu	Brooke
cat	1	1	1	3	1	3
all	1	1	1	1	1	1
no	1	1	1	3	5	2
fit	1	1	3	1	6	3
cute	2	1	1	1	1	1
share	1	1	1	1	1	1
crazy	1	3	1	4	1	1
sky	1	1	3	1	3	3
cough	1	2	2	1	5	3
thank	1	1	2	1	1	2



EXPERIMENTAL RESULT

CONCLUSION

- vocabulary recognition but human voice identification needs to be improved
- MFCC algorithm works good but the system is still not robust enough
- More voice control operation could be added
- A more efficient machine learning algorithm could be employed