

# J-DSP Online Lab Exercise for Speech Signal Formant Estimation and Clustering

April 23, 2021

Vivek Narayanaswamy, Andreas Spanias

Final NSF IUSE Workshop  
Collaborative Research: Integrated Development of Scalable Mobile  
Multidisciplinary Modules for STEM

SenSIP Center, School of ECEE, ASU



The work at Arizona State University is supported in part by the NSF DUE award 1525716 and the SenSIP Center.



## Motivation

- Increasing demand for online education.
- Many areas are seeking knowledge of DSP tools early in their curriculum.
- Exposition to speech processing and applications has become vital.
- Machine learning to infer statistical properties of speech has become ubiquitous.
- Requirements for on-the-fly learning and visualization.



## Background

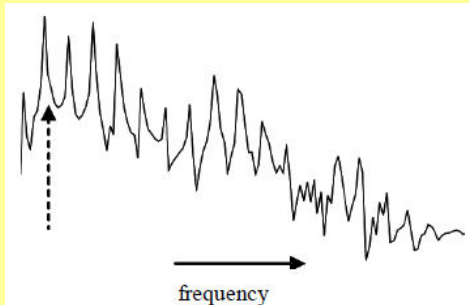
- ***Goals of the Project:***

- Reinforce fundamental concepts in speech processing.
- Describe concepts related to the speech production system / mechanism.
- Illustrate fundamental differences between Fine and Formant structures of speech.
- Introduce Analysis and Synthesis using Linear Prediction
- Introduce Machine Learning concepts such as clustering and prediction with examples
- Utilize J-DSP to cluster and estimate formant frequencies.

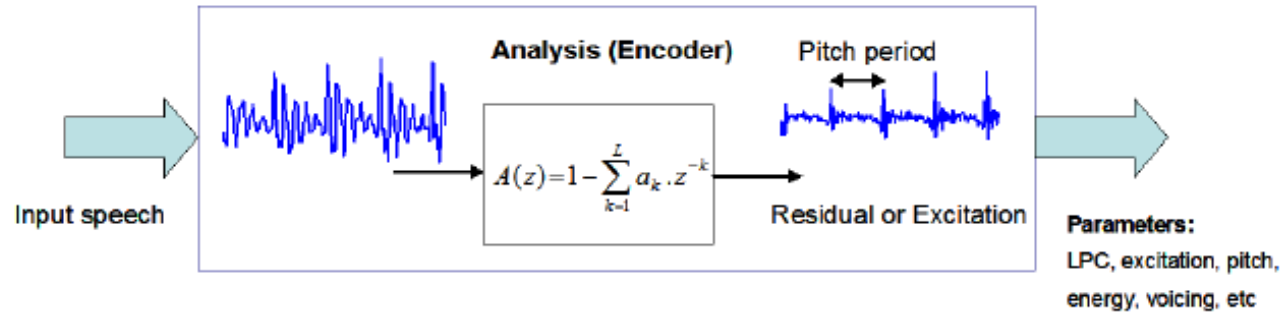


# Introducing Fundamentals of Speech Processing.

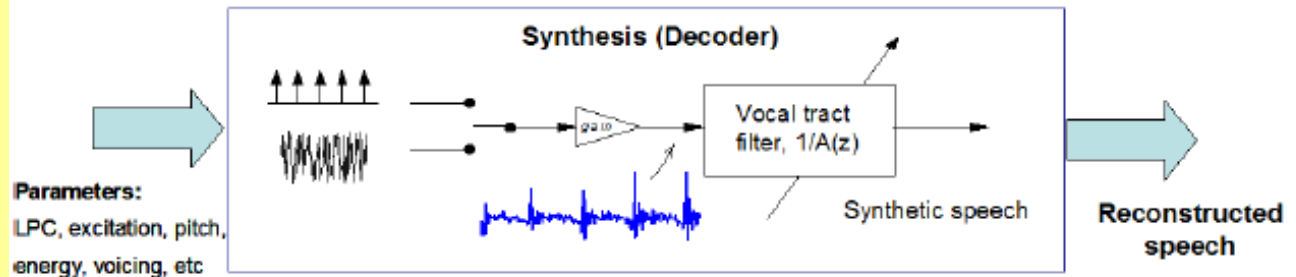
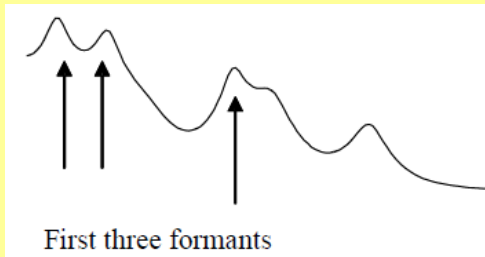
**Fine Structure**



**Analysis & Synthesis Systems for Speech**



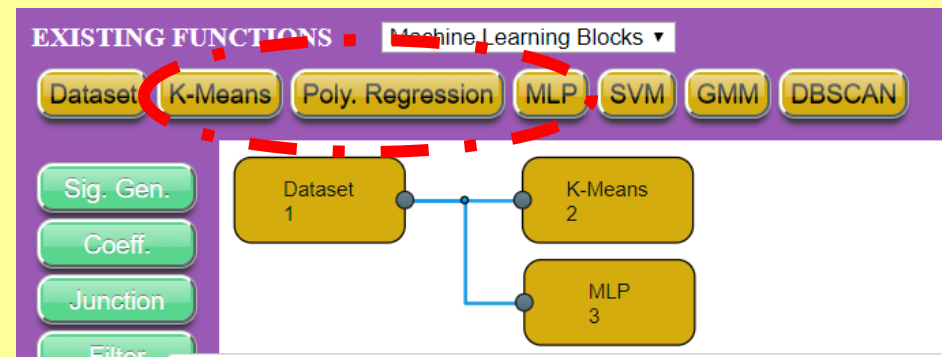
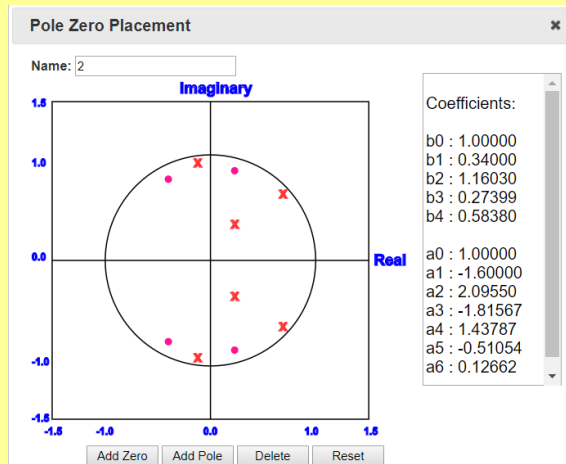
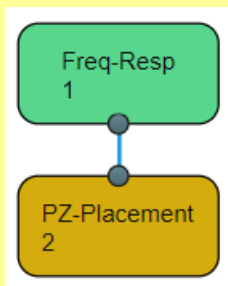
**Formant Structure**



**The students are provided dedicated lectures, notes and references to obtain a deeper understanding of the concepts involved**

## Introducing J-DSP Online Laboratory

- DSP-related courses are part of the accredited online engineering degrees offered by the ASU Online/EdPlus network .
- Real-time DSP operations such as Frequency Response, Pole-Zeros Placement, Filtering, Filter Designing can be easily implemented by students.
- Machine Learning blocks for Clustering / Prediction are available in a pick and place format

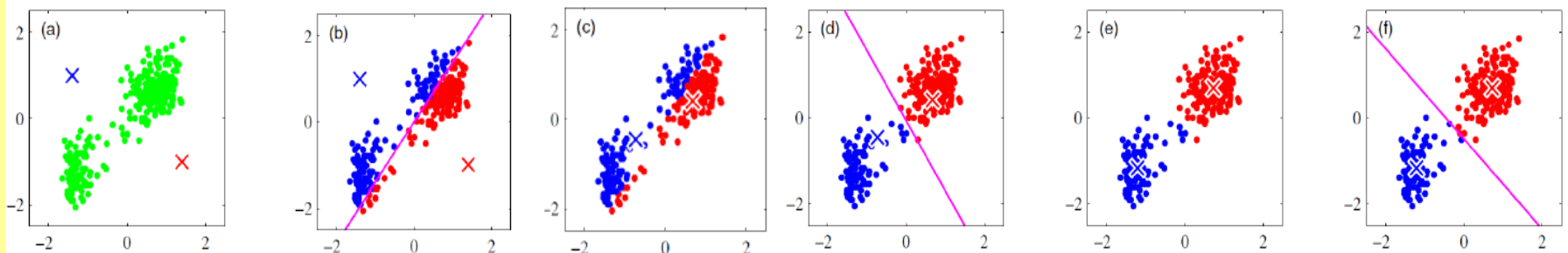


## Introducing K-Means Clustering

- We provide an overview of K-means clustering algorithm to the students .
- Cover concepts related to unsupervised machine learning.
- Provide mathematical descriptions and introduce methods of measuring convergence of the algorithm.

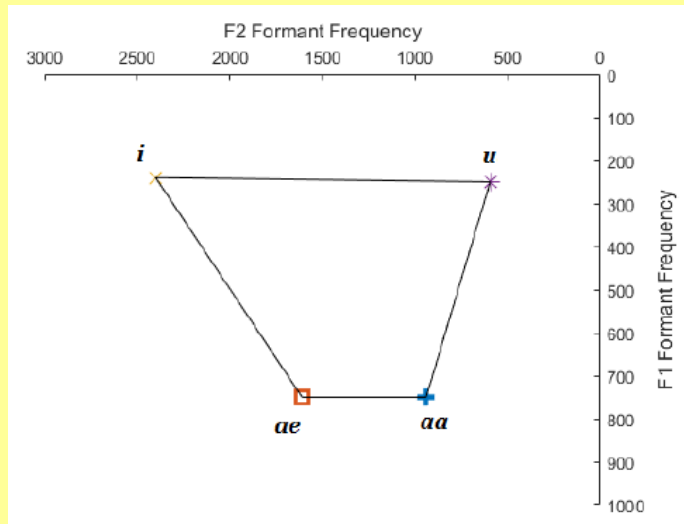
$$\sum_{i=0}^m \min_{\mu_{(j)} \in C} \|x_{(i)} - \mu_{j(i)}\|^2 = \sum_{k=1}^K \sum_{i \in C} \|x_{(i)} - \mu_{(k)}\|^2$$

- Visual explanation of the clustering process



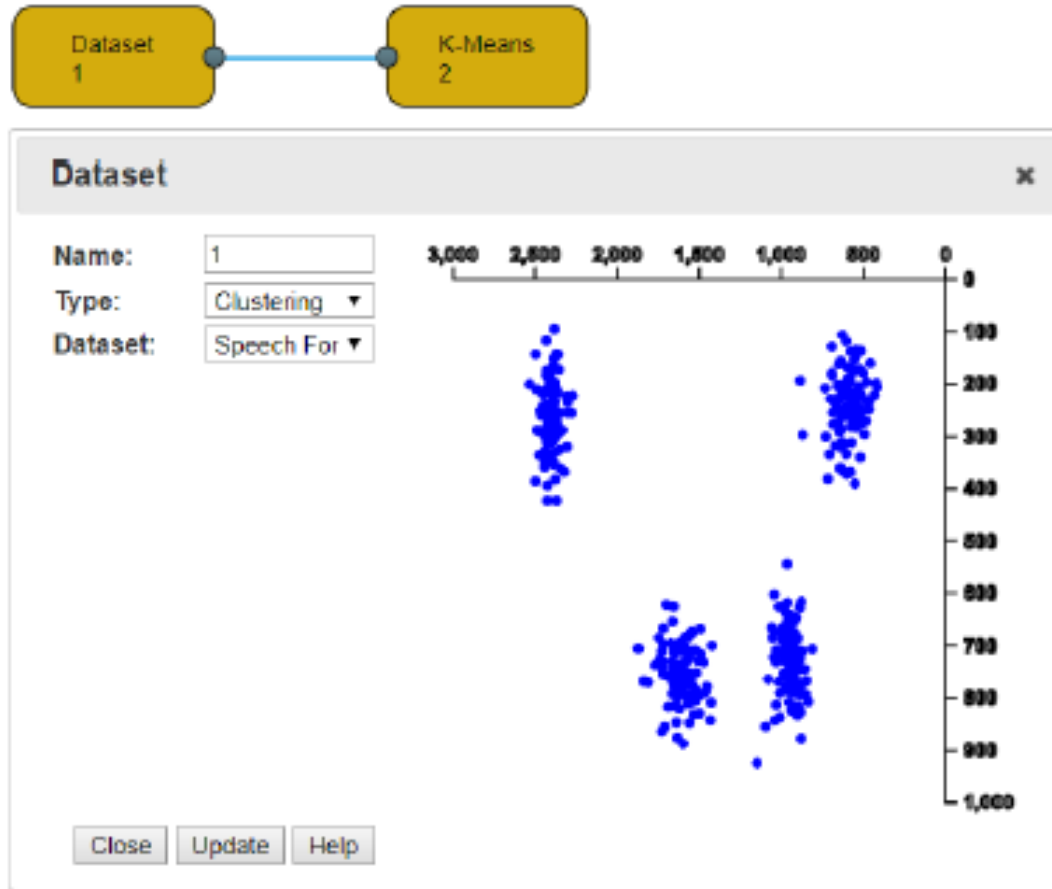
## Formant Dataset used for the JDSP Exercise

- **Dataset:** We collect speech signals from 100 different speakers (male and female) and extracted the formants F1 and F2 for the following vowels only: i, u, aa, and ae. Therefore, there are 100 different data points for each of the vowels.



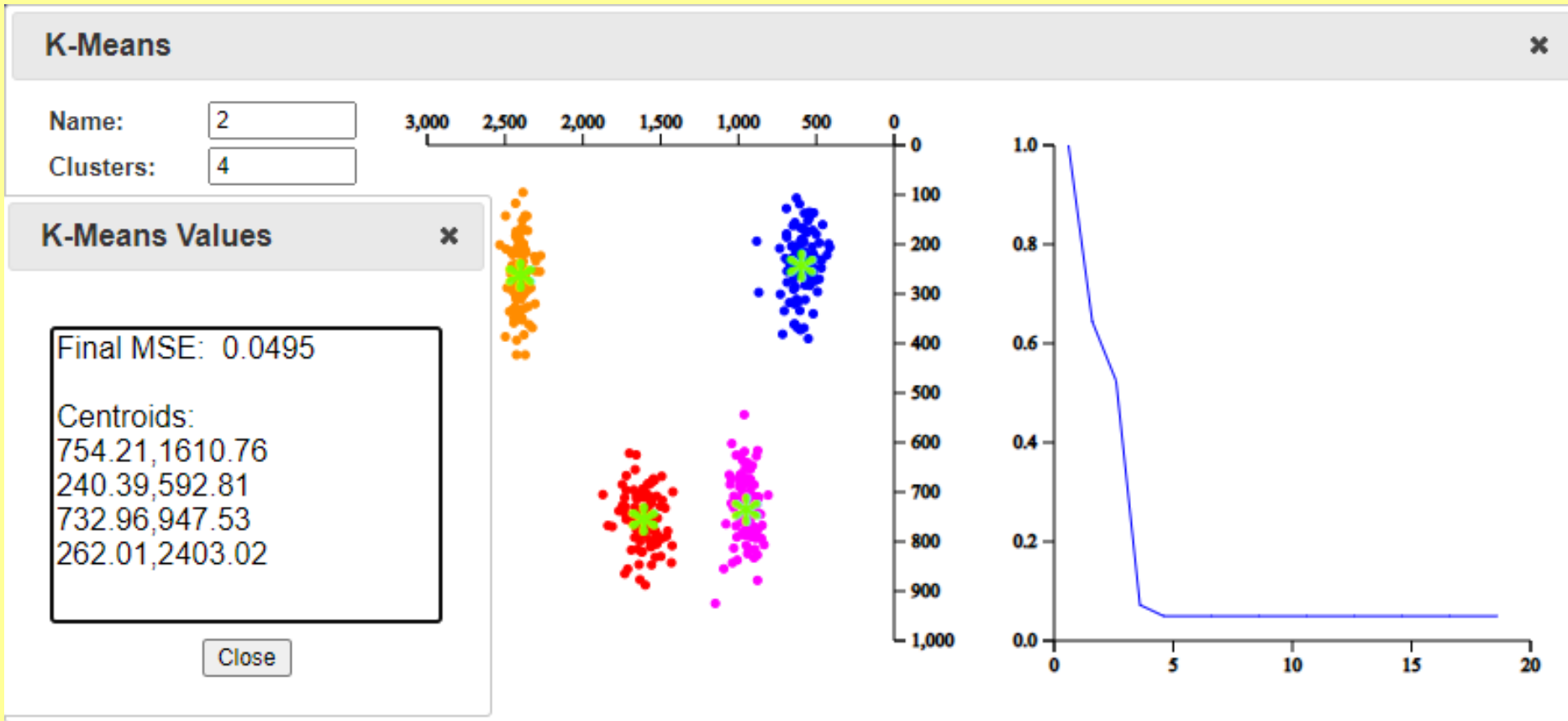
- Introduce students to vowel formant representation using a *vowel chart* or a *vowel space diagram*.
- It is a schematic arrangement of the vowels based on first two formant frequencies.
- Vowel chart usually takes the form of a quadrilateral as shown below.

# Formant Dataset Clustering using K-Means on JDSP



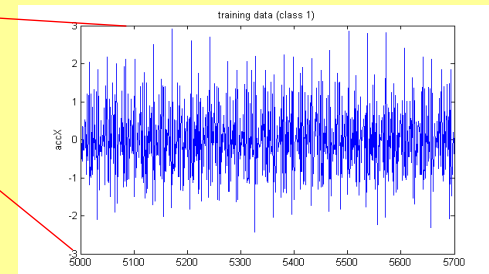
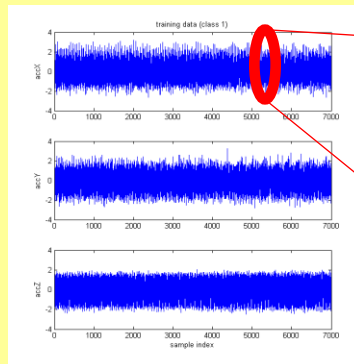


# Visualizing and Evaluating K-Means on JDSP



## Envisioned Speech And Audio Processing Functions

- By using HTML5, large sets of data can be processed at a time.
- Due to this advantage, we plan on using a function in our new software that can acquire speech samples in real-time using mobile devices like Android phones and iPhones .
- This will give our software an ability to perform real-time operations over real-life speech signals
- Through this function, we'll be able to capture and track the changes in parameters of the speech (Formant Analysis, Type of Speech etc)



## Assessment

- J-DSP assessment will be made through online forms and feedback from the students using the software in courses.
- Pre and Post Assessment Quizzes
- The users fill out and submit such forms instantaneously. The feedback data can also be accessed by links provided on-line.



## Conclusion

- The J-DSP online labs for performing speech formant clustering and estimation was elucidated.
- J-DSP allows users to interact with a more user-friendly environment that supports a variety of functionalities required by current academic practices.
- J-DSP provides simple and easy to visualize methodologies that aid in accelerated learning.
- The lab exercise was assigned as a course project for EEE407 and REU students working on speech signal analysis



## REFERENCES

- A. Spanias, Digital Signal Processing; An Interactive Approach – 2nd Edition.
- "HTML5 Differences from HTML4 – APIs". World Wide Web Consortium.
- Huang, C.; Thiagarajan, J. J.; Spanias, A.; Pattichis, C.;, "A Java-DSP interface for analysis of the MP3 algorithm," Proc. of the IEEE DSP/SPE Workshop, pp.168-173, 4-7 Jan. 2011.
- Dixit, A., Katoch, S., Spanias, P., Banavar, M., Song, H., & Spanias, A. (2017, October). Development of signal processing online labs using HTML5 and mobile platforms. In 2017 IEEE Frontiers in Education Conference (FIE) , pp. 1-5.
- Shanthamallu, U. S., Rao, S., Dixit, A., Narayanaswamy, V. S., Fan, J., & Spanias, A. (2019, May). Introducing Machine Learning in Undergraduate DSP Classes. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7655-7659). IEEE.
- Ted Painter and Andreas S. Spanias, "Perceptual Coding of Digital Audio," *Proceedings of the IEEE*, pp. 451-513, Vol. 88, No.4, April 2000.

**The work at Arizona State University is supported in part by the NSF DUE award 1525716 and the SenSIP Center. The work at Clarkson University is supported in part by the NSF DUE award 1525224.**

