
Assignment MAL: Recording Speech, Fourier Transform, Cluster Analysis, and Speech recognition

Table of Contents

.....	1
1. Recording Speech	1
2. Method 1	1
3. Method 2	2
4. Comparing the methods	2

The assignments below should be solved and documented as a mini-project that will form the basis for the examination. When solving the exercises it is important that

- you document all relevant results and analyses that you have obtained/performed during the exercises
- try to relate your results to the theoretical background of the methods being applied.

The documentation should be integrated (by adding new code/markdown cells) in this script which contains the exercises. This script is to be converted into a .pdf by using the markdown option in Matlab. Please see the following documentation for more information about markdown options

https://se.mathworks.com/help/matlab/matlab_prog/marking-up-matlab-comments-for-publishing.html

The mini-project must be uploaded to Studinet as a single pdf-file. You can create a .pdf of your script/markdown by saving the script as an m file (e.g. 'MyFile.m') and then running the following code:

```
publish('MyFile.m','pdf')
```

You can continuously check your output by clicking the "Publish" button under the 'Publish' tab. This will give you a quick html version of your document. Also, see the above documentation above on how to include images, graphs, tables, etc. in your file. If you have any questions about the exercises, you are strongly encouraged to have talk with your fellow students first. If you are reading the pdf version of this document, please note that you must find the script version and complete the assignments in the script.

1. Recording Speech

Using the MATLAB recorder that has been provided during the course, Record 100 speech signals of each of these words: Dog, Cat, Bird, Horse, Cow.

You have to extract features and analyze these words using two different methods, and compare the performance of the two methods in the end.

2. Method 1

1. Calculating the FFT

- Calculate the FFT of your signals.
- Compute the two-sided and one-sided spectrum of your recordings.
- Go through the one-sided spectrum in bins of 100Hz (0Hz – 100Hz, 101Hz-200Hz etc.), and find the most prominent peak (and index) in each of these frequency bins.
- From the most prominent peaks that you just extracted, take the 10 highest peak and the corresponding frequencies. The frequencies are the features that we will use in the following cluster analysis.

2. Cluster Analysis

- Use either K-means or hierarchical cluster analysis to cluster the data.
- Pick the parameters that you use for the cluster analysis and explain your choices.
- Show the "elbow" plot for the cluster analysis.
- Determine the specificity and sensitivity of the clustering.

3. Method 2

1. Calculating the MFCC

- Calculate the MFCCs of your signals. Extract 10 feature values for each.

2. Model Generation

- Use your different arrays of features to generate a GMM of each word.
- Determine the specificity and sensitivity of the models, by comparing each recording against all models in a posterior analysis.

4. Comparing the methods

Use the calculated specificity and sensitivity for each method to compare their performance - which method performs the best? Describe why you think that this method is better than the other. List the pros and cons for each of these methods when doing speech recognition.

Published with MATLAB® R2016b