

For each student, the following audio clips are recorded:

1. LT1: Long Train 1
2. LT2: Long Train 2
3. ST1: Short Train 1
4. ST2: Short Train 2
5. TE: Test

Suppose there are N students in total, and you are using LPC coefficients with an order of M . Your frame size per sample is t , the audio clip length is T . Here is what you need to do for HW#3:

1. Choose one from long train and short train. I.e., your data set is either (LT1, LT2, TE) or (ST1, ST2, TE). No matter what set you use, denote them as (T_1, T_2, TE) .
2. Partition all of the clips into frames, for each student i ($i=1,2,\dots,N$), the amount of frames is $(N_{T_1^i}, N_{T_2^i}, N_{TE^i})$ for (T_1^i, T_2^i, TE^i) respectively. The value of $(N_{T_1^i}, N_{T_2^i}, N_{TE^i})$ depends on your framing approach. E.g., if you use non-overlapped framing, you will have $(N_{T_1^i}, N_{T_2^i}, N_{TE^i}) = \left(\frac{T_{T_1^i}}{t}, \frac{T_{T_2^i}}{t}, \frac{T_{TE^i}}{t} \right)$. If you use overlapped framing, the number is larger correspondingly. For all students, you have $\hat{N} = \sum_{i=1}^N (N_{T_1^i} + N_{T_2^i} + N_{TE^i})$ samples.
3. For each of these \hat{N} frames, calculate the LPC coefficients. For each frame there are M coefficient $\mathbf{f} = (f_1, f_2, \dots, f_M)$. Therefore, for each student i ($i=1,2,\dots,N$), you get three coefficient matrix, i.e., three coefficient set: $\mathbf{F}_{T_1^i}, \mathbf{F}_{T_2^i}, \mathbf{F}_{TE^i}$, where

$$\mathbf{F}_{X^i} = \begin{pmatrix} \mathbf{f}_{X^i}^1 \\ \mathbf{f}_{X^i}^2 \\ \dots \\ \mathbf{f}_{X^i}^N \end{pmatrix}, X = T_1, T_2, \text{ or } TE.$$

4. Calculate the following:

$$ID_{T_2}^{i_*} = ID_{T_1}^{\hat{i}}, \hat{i} = \arg \min_i D(\mathbf{F}_{T_2^{i_*}}, \mathbf{F}_{T_1^i})$$

$$ID_{TE}^{i_*} = ID_{T_1}^{\hat{i}}, \hat{i} = \arg \min_i D(\mathbf{F}_{TE^{i_*}}, \mathbf{F}_{T_1^i})$$

where ID is the identity. i_* is your index in all of the students. D is the Hausdorff distance. $ID_{T_2}^{i_*}$ means the identity verification result of the data set $T_2^{i_*}$. $ID_{TE}^{i_*}$ is similar.

This is the procedure to identify your test set (T2 and TE) conducting Nearest Neighbor method. You may also create a larger anti-self set by combing of the feature set from other students.