



Ψηφιακή Επεξεργασία Σημάτων

1η Εργαστηριακή Άσκηση

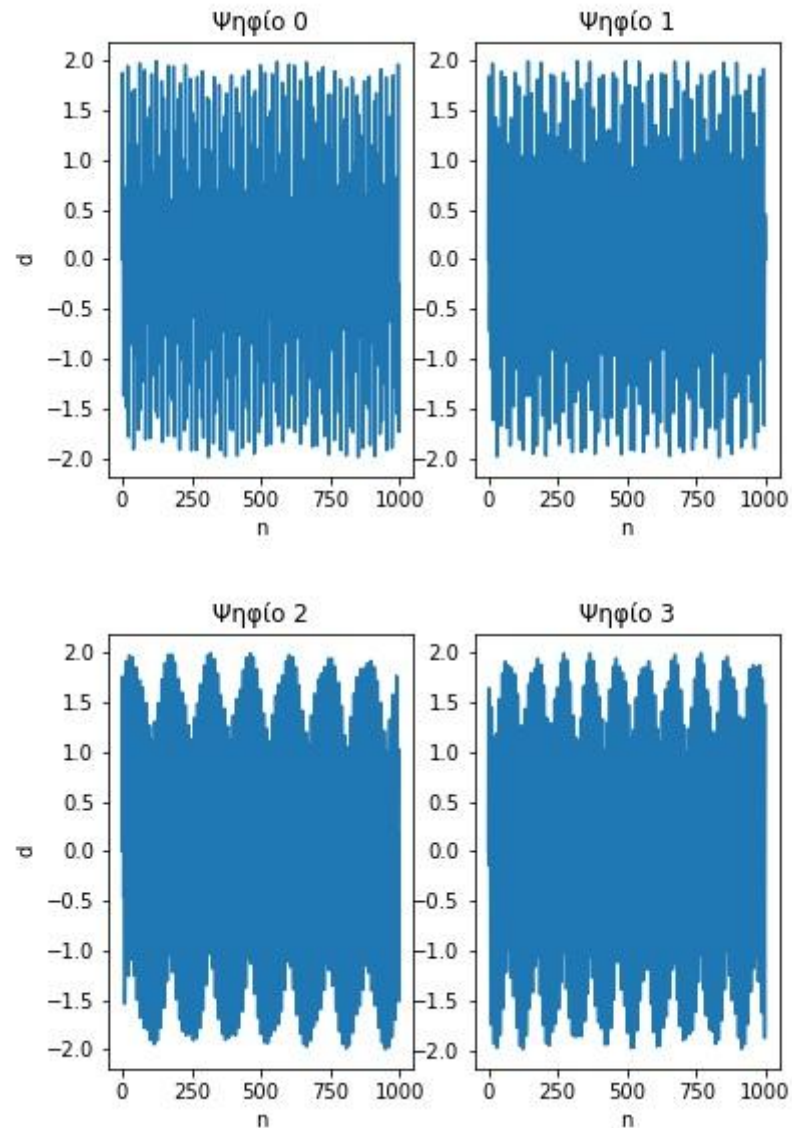
Εισαγωγή στην Ψηφιακή Επεξεργασία Σημάτων με Python και
Εφαρμογές σε Ακουστικά Σήματα

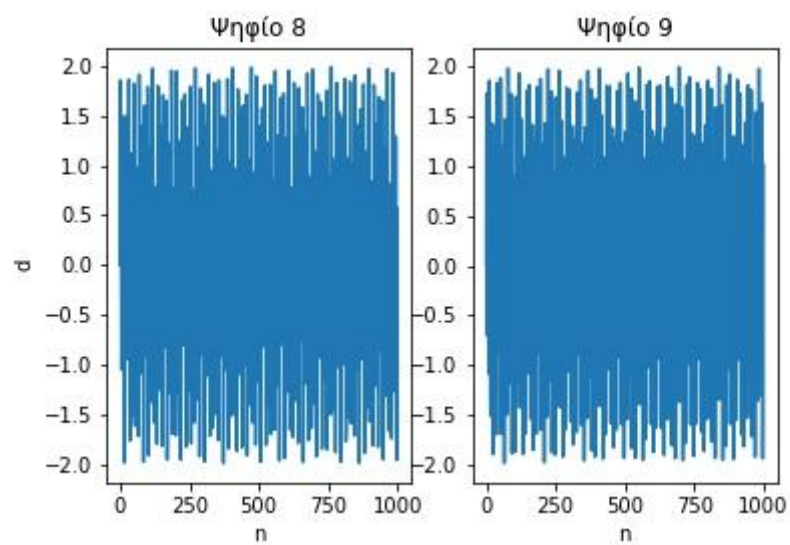
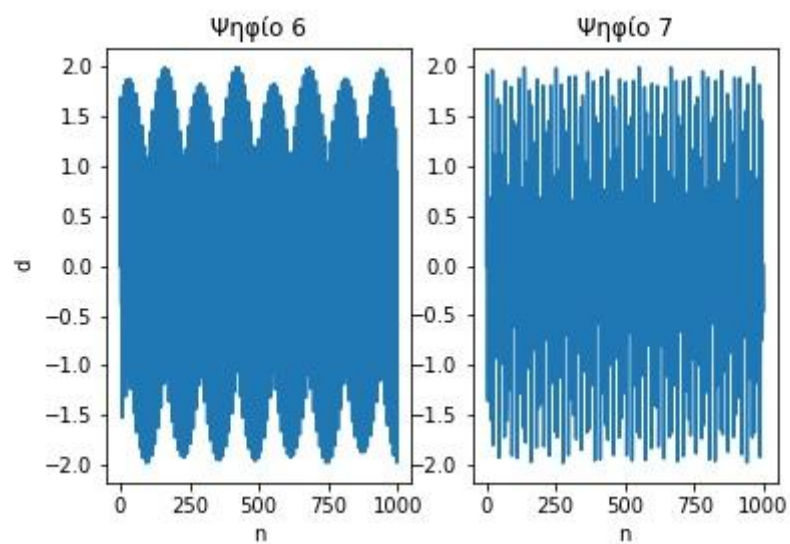
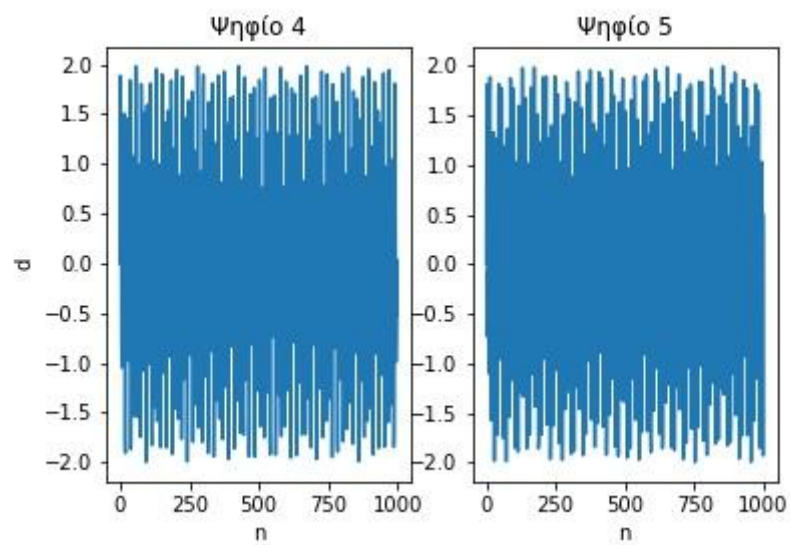
Στέφανος Κοντόπουλος

Μέρος 1: Σύστημα Εντοπισμού Τηλεφωνικών Τόνων

Σαν πρώτο βήμα είναι η κατασκευή των τηλεφωνικών τόνων ως άθροισμα ημιτόνων συχνότητας που προέρχεται από τον δοθέν πίνακα 1.

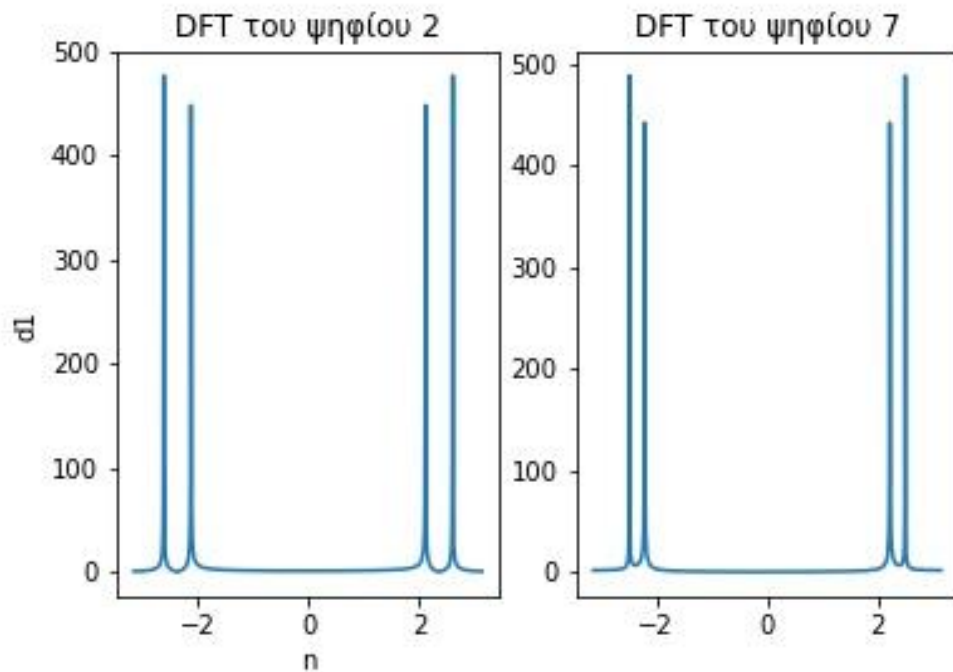
Οπτικοποίηση των τηλεφωνικών τόνων:





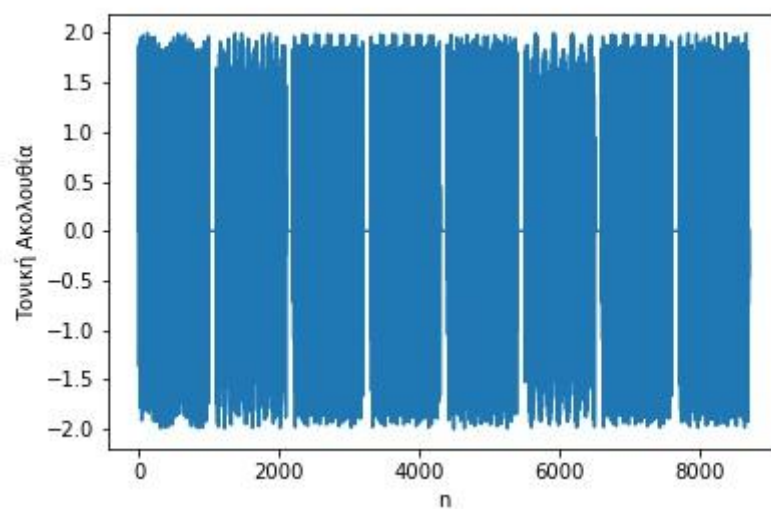
Στην συνέχεια υπολογίζουμε τον DFT των παραπάνω σημάτων.

Οπτικοποίηση των DFT των ψηφίων 2 & 7 αντίστοιχα:



Στην συνέχεια κατασκευάστηκε το ζητούμενο αρχείο `tone_sequence`, που περιέχει την ακολουθία των ψηφίων του αριθμού μητρώου, δηλ. 0 3 1 1 4 6 1 7.

Οπτικοποίηση του `tone_sequence.wav`:



Έπειτα υπολογίστηκαν οι Fourier M/Σ με τετραγωνικά και Hamming παράθυρα και βρήκαμε μια λίστα από δείκτες και από τις αντίστοιχες τους συχνότητες που είναι κοντά στις touch tone.

```
Λίστα k:  
[[115. 163.]  
 [ 85. 148.]  
 [ 85. 163.]  
 [ 85. 180.]  
 [ 94. 148.]  
 [ 94. 163.]  
 [ 94. 180.]  
 [104. 148.]  
 [104. 163.]  
 [104. 180.]]  
  
Λίστα συχνοτήτων:  
[[0.72256631 1.02415921]  
 [0.53407075 0.92991143]  
 [0.53407075 1.02415921]  
 [0.53407075 1.13097336]  
 [0.59061942 0.92991143]  
 [0.59061942 1.02415921]  
 [0.59061942 1.13097336]  
 [0.65345127 0.92991143]  
 [0.65345127 1.02415921]  
 [0.65345127 1.13097336]]
```

Η συνάρτηση ttdecode που περιέχεται στο αρχείο Part1.ipynb παίρνει σαν όρισμα κάποιο τονικό σήμα και επιστρέφει την αντίστοιχη τονική ακολουθία.

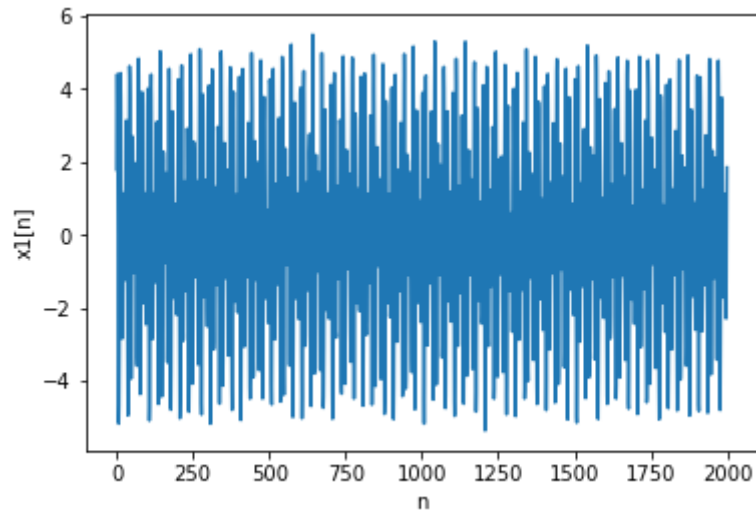
Έγινε έλεγχος βάζοντας ως είσοδο το tone_sequence.wav και πήραμε στην έξοδο την σωστή ακολουθία, δηλ: [0, 3, 1, 1, 4, 6, 1, 7]

Το κάλεσμα της συνάρτησης με εισόδους τα σήματα easySig.npy και hardSig.npy

```
easySig: [8, 1, 0, 3, 9, 6, 3, 8]  
έδωσε: hardSig: [4, 8, 1, 9, 2, 1, 5, 3, 6, 3]
```

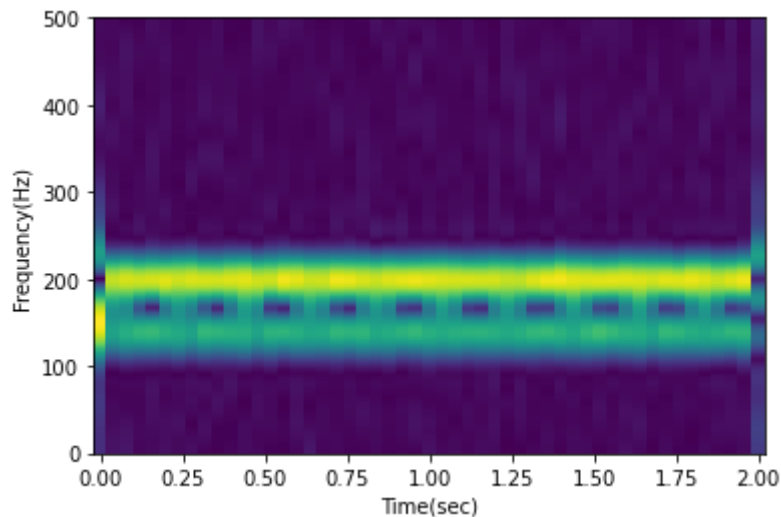
Μέρος 2: Φασματική Ανάλυση Ημιτονοειδών και Ανίχνευση Απότομων Μεταβάσεων με τον Μ/Σ Fourier Βραχέος Χρόνου και τον Μ/Σ Wavelets

Αρχικά υπολογίζουμε το δειγματοληπτημένο σήμα $x_1[n]$:

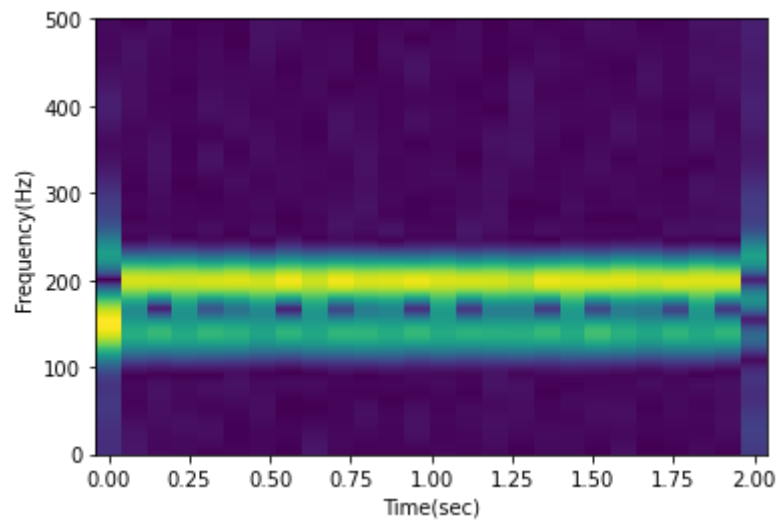


Αφού υπολογίσουμε τον STFT, αναπαριστούμε το πλάτος του με χρήση της συνάρτησης `pcolormesh`.

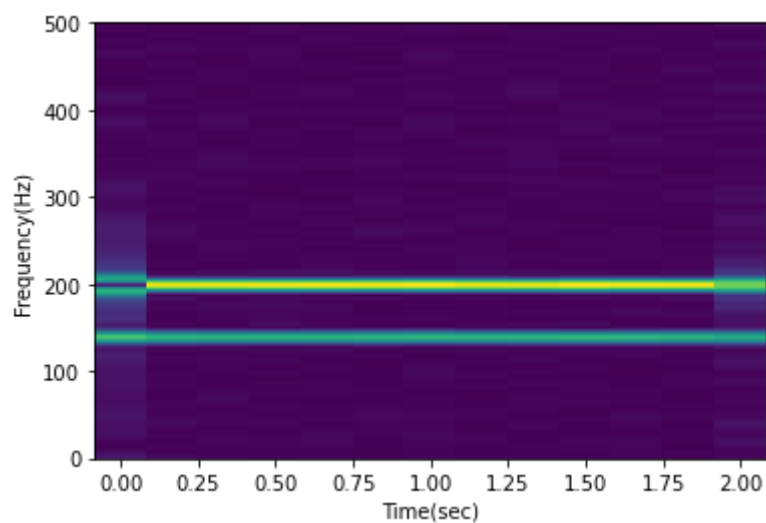
Για μήκος παραθύρου 0.04sec και επικάλυψη 0.02sec:



Για μήκος παραθύρου 0.08sec και επικάλυψη 0.04sec:

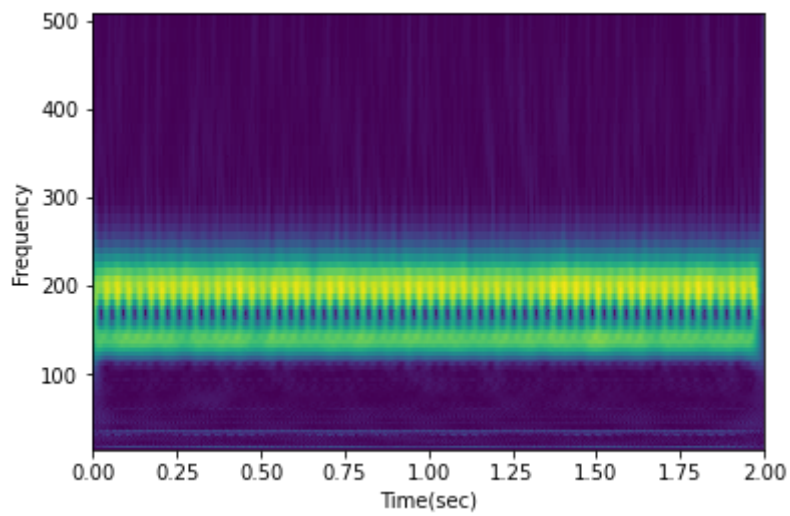


Για μήκος παραθύρου 0.16sec και επικάλυψη 0.08sec:

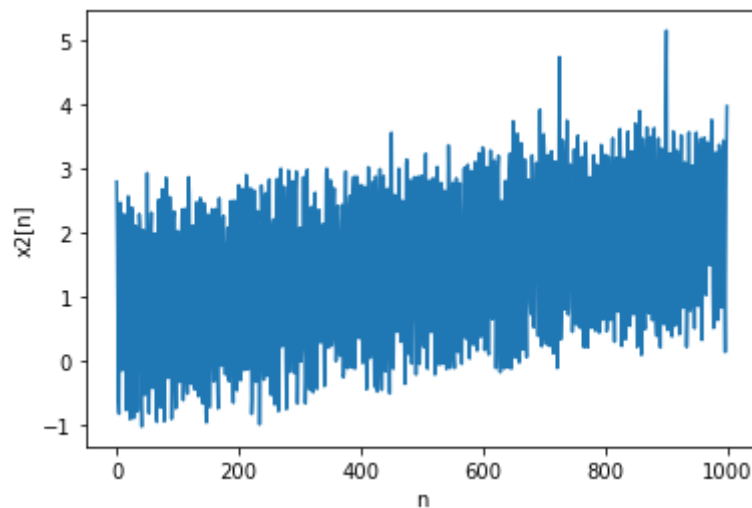


Από τις διαφορετικές τιμές των παραθύρων που χρησιμοποιήσαμε στον υπολογισμό του STFT, παρατηρούμε ότι τα πλάτη κυμαίνονται στην ίδια περιοχή και πως πρακτικά έχουμε καλύτερη ανάλυση του πλάτους. Αυξάνοντας το μέγεθος του παραθύρου παίρνουμε καλύτερη πληροφορία στο πεδίο της συχνότητας, με κόστος να χάσουμε στο πεδίο του χρόνου.

Αφού υπολογίσαμε τώρα τον διακριτοποιημένο CWT, αναπαριστούμε το πλάτος του με χρήση της συνάρτησης `pcolormesh`:

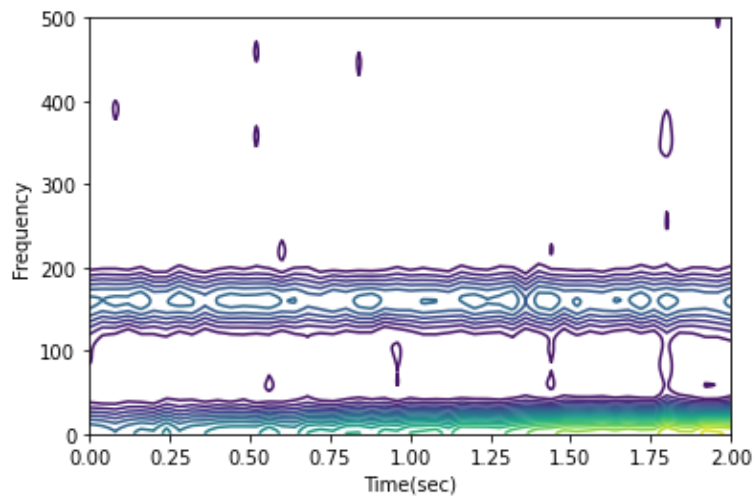


Ίδια διαδικασία για να υπολογίσουμε το δειγματοληπτημένο σήμα $x_2[n]$, όπου στις ζητούμενες θέσεις($n=725$ και $n=900$) έγινε αύξηση της τιμής τους ώστε να διακρίνονται εύκολα:

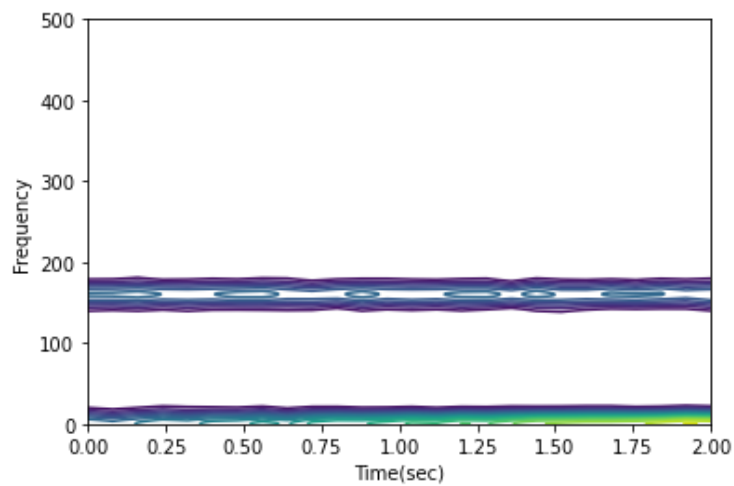


Αφού υπολογίσαμε τον STFT τώρα, αναπαριστούμε το πλάτος του με χρήση της συνάρτησης `contour`.

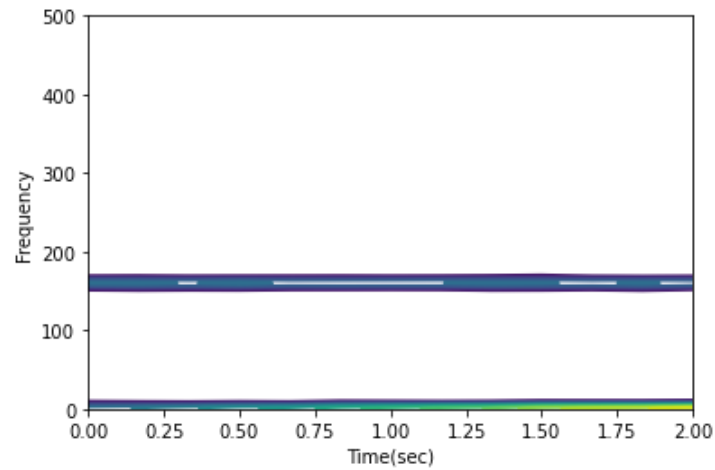
Για μήκος παραθύρου 0.04sec και επικάλυψη 0.02sec:



Για μήκος παραθύρου 0.08sec και επικάλυψη 0.04sec:

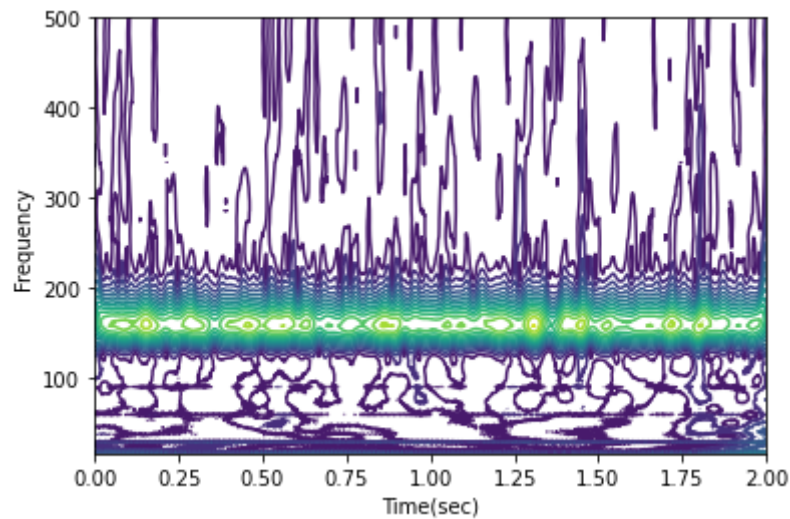


Για μήκος παραθύρου 0.08sec και επικάλυψη 0.04sec:



Ισχύουν και εδώ οι παραπάνω παρατηρήσεις.

Υπολογίζουμε τώρα τον διακριτοποιημένο CWT και αναπαριστούμε το πλάτος του με χρήση της συνάρτησης contour:



Τα όποια artifacts και θόρυβος στις απεικονίσεις προέρχονται από το όρισμα του σήματος.

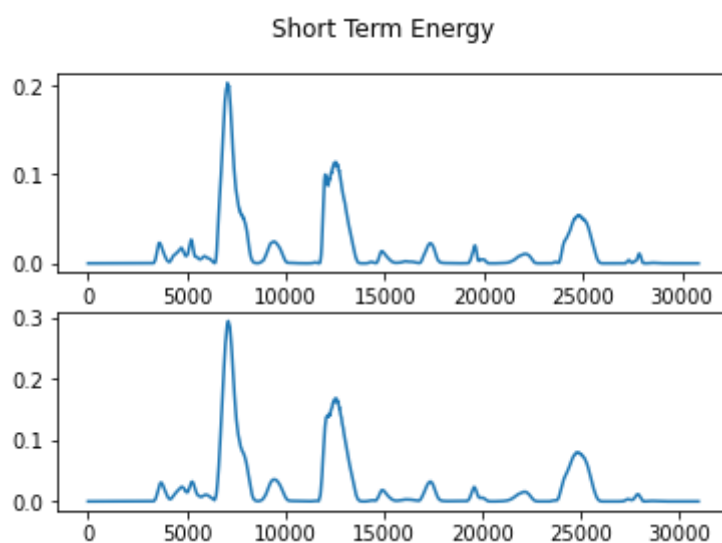
Γενικά ο CWT δίνει καλύτερη πληροφορία για το σήμα και έχει καλύτερη ανάλυση αφού χρησιμοποιεί πολλαπλά παράθυρα(κλίμακες).

Μέρος 3: Χαρακτηριστικά Βραχέος Χρόνου Σημάτων Φωνής και Μουσικής

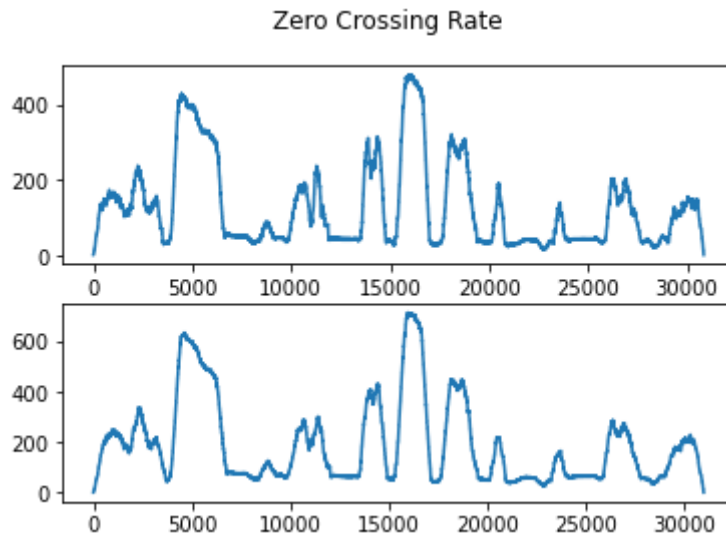
Σε αυτό το μέρος στόχος μας είναι να ξεχωρίσουμε άφωνους ήχους από έμφωνους, χρησιμοποιώντας την ενέργεια βραχέος χρόνου και τον ρυθμό εναλλαγής προσήμου σε δύο διαφορετικά ακουστικά σήματα.

Το πρώτο πρόκειται για μια πρόταση και το δεύτερο για την εισαγωγή ενός κομματιού μουσικής.

Ενέργεια βραχέος χρόνου για το φωνητικό σήμα με παράθυρα 20&30ms:



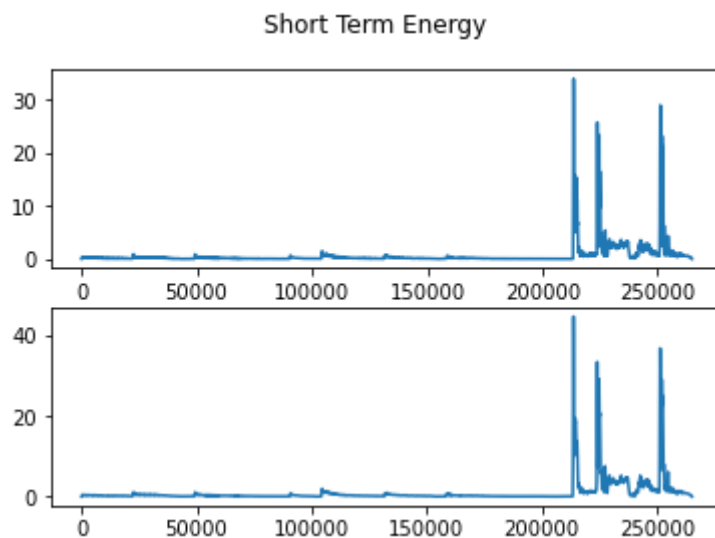
Ρυθμός εναλλαγής προσήμου για το φωνητικό σήμα με παράθυρα 20&30ms:



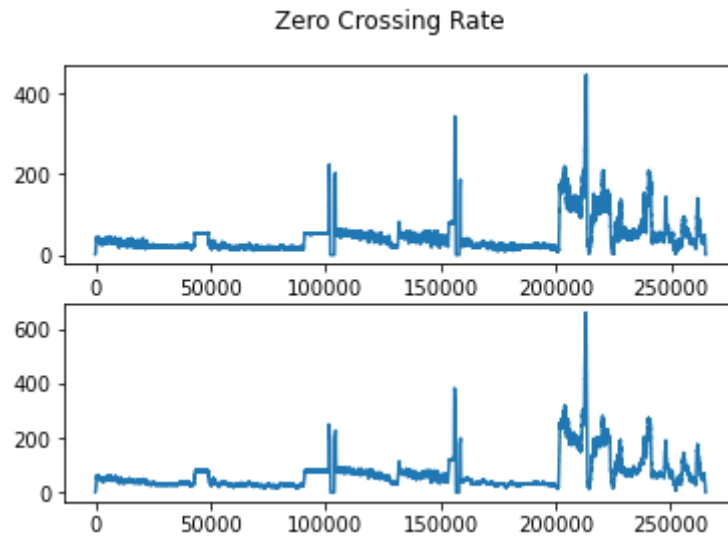
Μεγαλύτερο παράθυρο μας δίνει μια πιο ομαλή συνέλιξη χωρίς να χάνουμε πληροφορία για το φωνητικό μας σήμα.

Με την ενέργεια βραχέος χρόνου μπορούμε να ξεχωρίσουμε άφωνους από έμφωνους ήχους αλλά και την σιωπή. Επίσης από τον ρυθμό εναλλαγής προσήμου ξεχωρίζουμε εύκολα τους άφωνους ήχους, αφού εκεί η τιμή του πλάτους είναι πιο υψηλή σε σχέση με τους έμφωνους

Ενέργεια βραχέος χρόνου για το μουσικό σήμα με παράθυρα 20&30ms:



Ρυθμός εναλλαγής προσήμου για το φωνητικό σήμα με παράθυρα 20&30ms:



Εδώ βλέπουμε πως με την ενέργεια βραχέος χρόνου είναι ξεκάθαρο πότε ξεκινάει ο τραγουδιστής να τραγουδάει. Τα peaks στον ρυθμό εναλλαγής προσήμου αντιστοιχούν στα πιο “μπάσσα” τμήματα του κομματιού.