Εθνικο Μετσοβιο Πολυτεχνειο

Σχολη Ηλεκτρολογων Μηχανικων και Μηχανικων Υπολογιστων

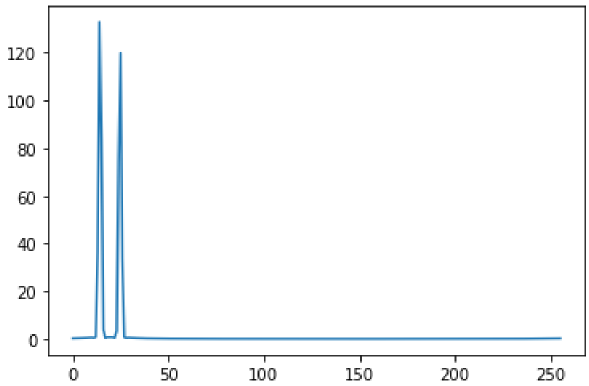
Πρωτη Εργαστηριακη Αναφορα
Ψηφιακη Επεξεργασια Σηματος

Κωνσταντίνος Κονιαβίτης
03119182
Αναστασία Χριστίνα Λίβα
03119029

Περιεχόμενα

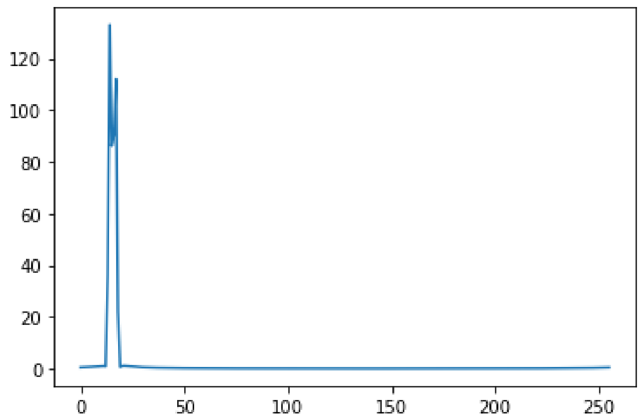
Μέρος 1

1.1 Έστω σήμα $y[n] = w[n](x_1[n] + x_2[n]) = w[n](e^{j(w_1 n+\varphi_1)} + 0.9e^{j(w_2 n+\varphi_2)})$ με μήκος $L = 256$. Υπολογίζουμε τον DFT 256 δειγμάτων ελαττώνοντας σταδιακά τη μεγαλύτερη συχνότητα, δηλαδή την w_2.

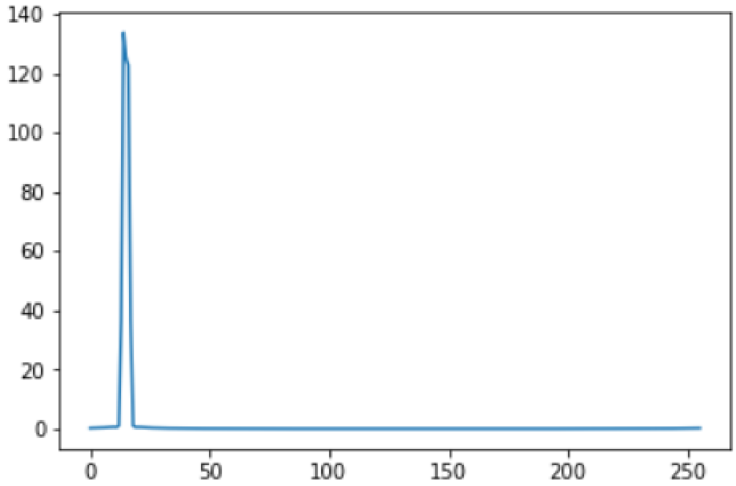


DFT σήματος για $w_1 = \frac{\pi}{9}$ και $w_2 = \frac{\pi}{5}$

Καθώς η συχνότητα w_2 πλησιάζει την w_1 παρατηρούμε ότι τα δύο peaks πλησιάζουν. Όταν η συχνότητα w_2 έχει τιμή ίση με 0.40831853071795843, δηλαδή η διαφορά είναι $\Delta w = |w_1 - w_2| = 0.05925268031909253$ βλέπουμε σχετικά καθαρά τα δύο peaks, ενώ όταν $w_2 = 0.3883185307179584$, οπότε $\Delta w = 0.039252680319092514$ τα peaks μόλις που διακρίνονται. Παρόλο που $\Delta w \neq 0$ δεν είναι δυνατή η διάκριση των δύο κορυφών. Αυτό συμβαίνει γιατί ο DFT ως δειγματοληψία του DTFT δεν τον αποδίδει πλήρως, καθώς λαμβάνει πεπερασμένο αριθμό δειγμάτων.

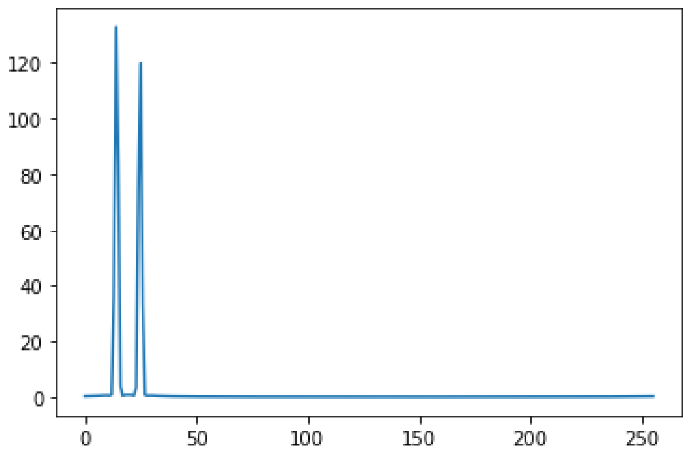


Στο διάγραμμα αυτό έχουμε $w_2 = 0.40831853071795843$, δηλαδή η διαφορά είναι $\Delta w = 0.05925268031909253$. Τα δύο peaks φαίνονται σχετικά καθαρά.

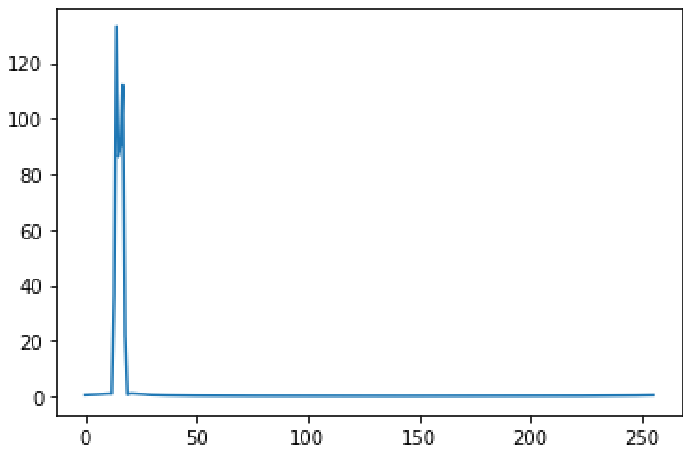


Στο διάγραμμα αυτό έχουμε $w_2 = 0.3883185307179584$, οπότε $\Delta w = 0.039252680319092514$. Οι δύο κορυφές μόλις που διακρίνονται.

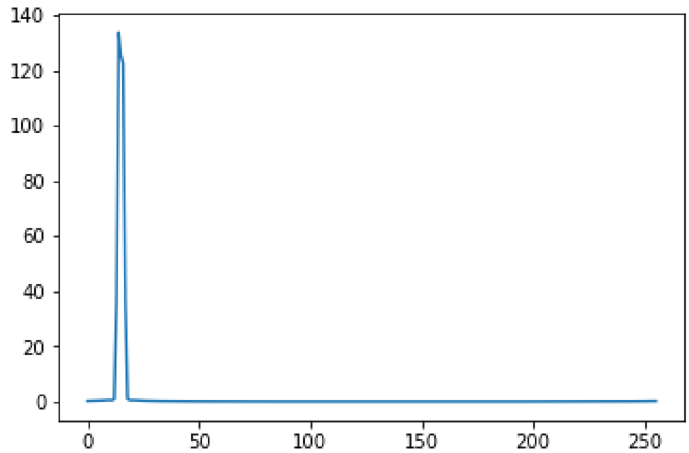
1.2 Αυξάνοντας το μήκος N του DFT παρατηρούμε ότι τα peaks προσομοιάζουν όλο και πιο πολύ τη συνάρτηση Dirac, πράγμα αναμενόμενο καθώς λαμβάνουμε περισσότερα δείγματα του DTFT. Παρότι λαμβάνουμε περισσότερα δείγματα του DTFT η διακριτική μας ικανότητα όσον αφορά τις συχνότητες δεν έχει βελτιωθεί, όπως βλέπουμε στα ακόλουθα διαγράμματα



Σε αυτό το διάγραμμα βλέπουμε τον DFT δίχως να έχουν αλλάξει οι συχνότητες

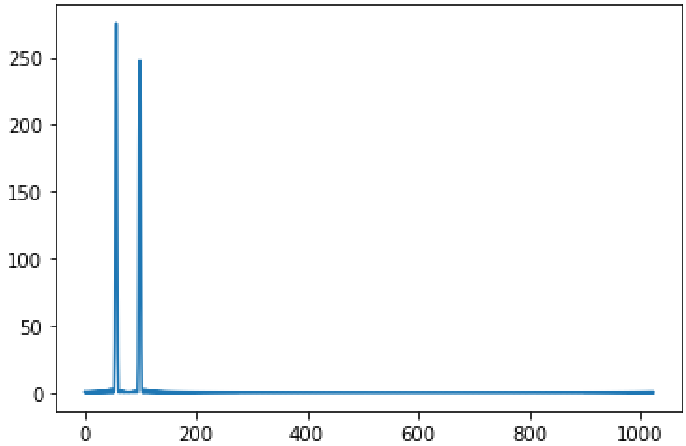


Σε αυτό το διάγραμμα έχουμε $w_2 = 0.40831853071795843$, δηλαδή η διαφορά είναι $\Delta w = 0.05925268031909253$. Τα δύο peaks φαίνονται σχετικά καθαρά.

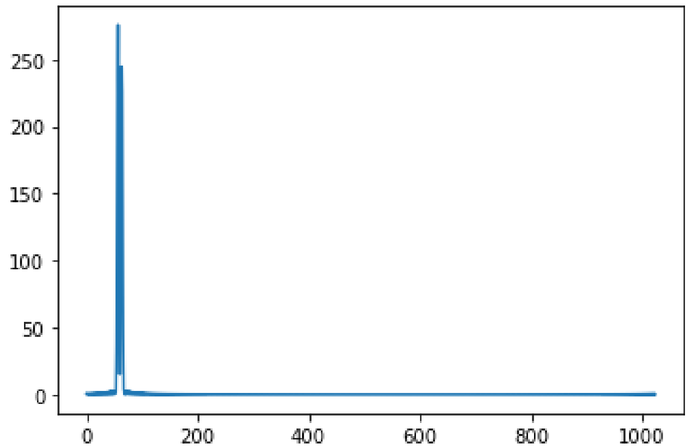


Στο από πάνω διάγραμμα έχουμε $w_2 = 0.3883185307179584$, οπότε $\Delta w = 0.039252680319092514$. Οι δύο κορυφές μόλις που διακρίνονται.

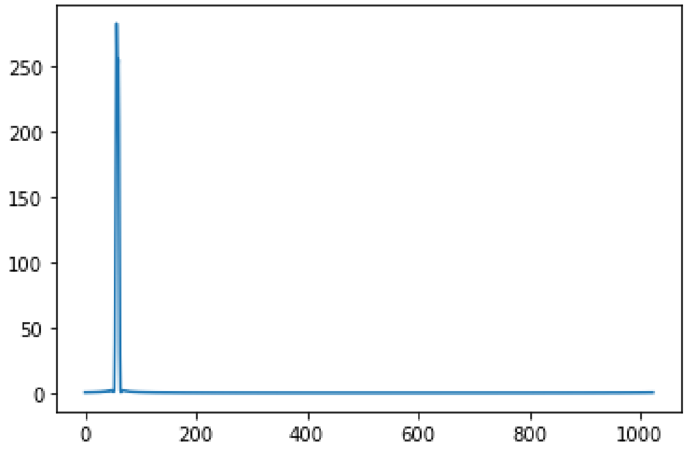
1.3 Ακολουθεί ο DFT του σήματος για N=1024 και L=512:



Για τη συχνότητα που παρατηρήσαμε οριακή διάκριση πριν (για N=256, L=256), τώρα (για N=1024, L=512) παρατηρούμε ξεκάθαρη διάκριση. Αυτό συμβαίνει γιατί έχουμε πολλά περισσότερα δείγματα στον DFT. Ακολουθεί το αντίστοιχο διάγραμμα:

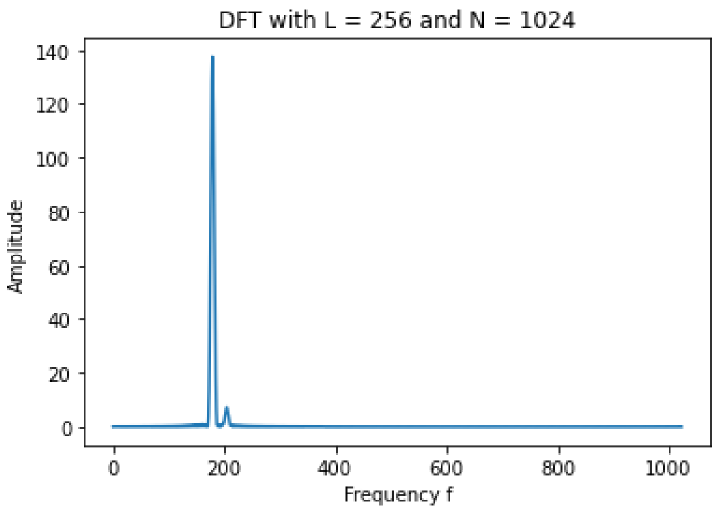


Όταν $w_2 = 0.3883185307179584$, άρα $\Delta w = 0.039252680319092514$, οι κορυφές μόλις διακρίνονται, ενώ για $w_2 = 0.3683185307179584$, άρα $\Delta w = 0.019252680319092497$, οι δύο κορυφές συμπίπτουν. Παρατηρούμε επίσης ότι το νέο φασματικό όριο είναι μικρότερο από ότι προηγουμένως, πράγμα αναμενόμενο καθώς λαμβάνουμε περισσότερα δείγματα του σήματος. Ακολουθεί το αντίστοιχο διάγραμμα:

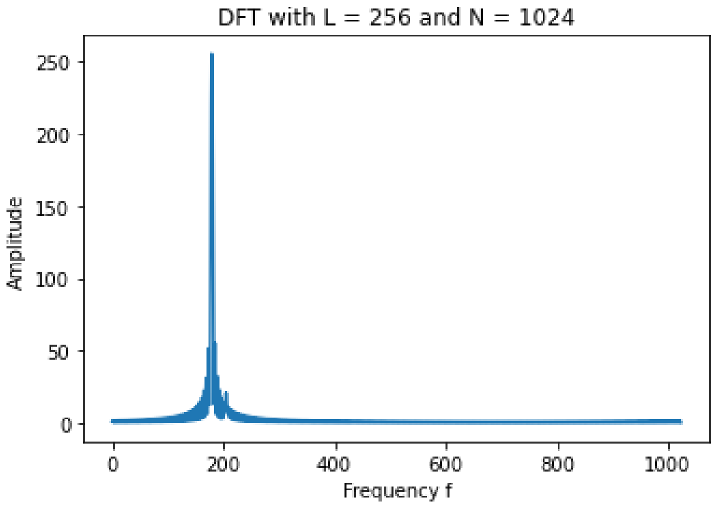


1.4 Έστω σήμα $y[n] = w[n](x_1[n] + x_2[n]) = w[n](e^{j(0.35\pi n+\varphi_1)} + 0.05e^{j(0.4\pi n+\varphi_2)})$ Παρατηρούμε πως χρησιμοποιώντας το παράθυρο Hamming έχουμε ευκρινέστερη διάκριση συχνοτήτων στις κορυφές από ότι χρησιμοποιώντας το τετραγωνικό παράθυρο. Το φαινόμενο αυτό οφείλεται στο γεγονός ότι στο τετραγωνικό παράθυρο έχουμε απότομη πτώση σε dB με αποτέλεσμα την απώλεια πληροφορίας. Το πλάτος της δεύτερης συχνότητας είναι κατά 100 φορές μικρότερο, και κατά συνέπεια δεν είναι δυνατή η διάκριση της δεύτερης κορυφής.

Ακολουθούν τα διαγράμματα των δύο DFT:



Χρήση παραθύρου Hamming

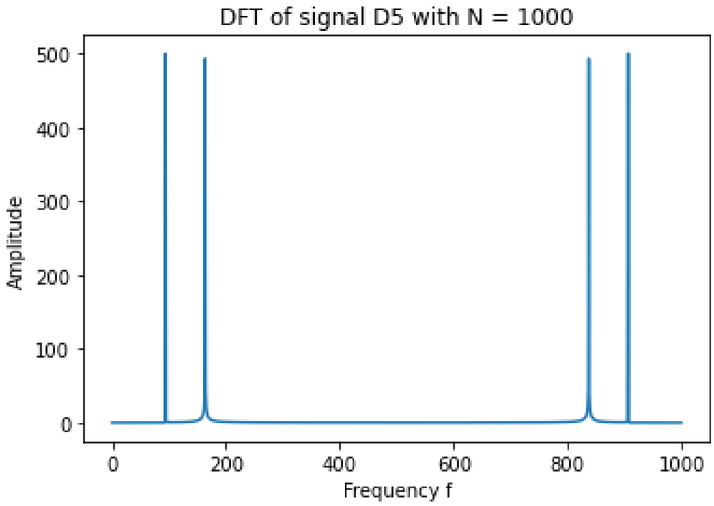


Χρήση τετραγωνικού παραθύρου, όπου φαίνεται ξεκάθαρα ο θόρυβος.

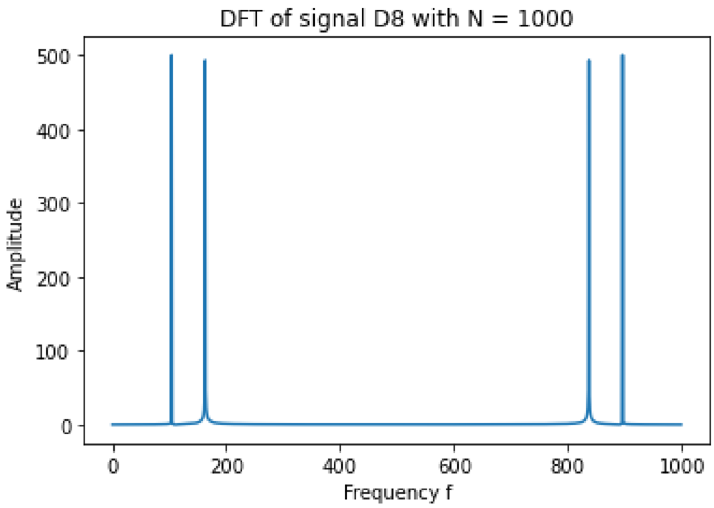
Μέρος 2

2.1 Δημιουργούμε 10 σήματα διαφορετικών συχνοτήτων $d_0, d_1, ..., d_9$ καθένα εκ των οποίων αντιστοιχεί σε ένα πλήκτρο τηλεφώνου.

2.2 Υπολογίζουμε DFT 1000 σημείων στα σήματα d_5 και d_8. Για κάθε σήμα παρατηρούμε τέσσερα peaks, πράγμα αναμενόμενο καθώς κάθε σήμα προκύπτει από άθροισμα δύο ημιτόνων διαφορετικών συχνοτήτων.



DFT πλήκτρου 5



DFT πλήκτρου 8

2.3 Δημιουργούμε αλληλουχία τηλεφωνικών τόνων βάσει του αθροίσματος ψηφίων των αριθμών μητρώων μας (03119029 + 03119182 = 06238211) τοποθετώντας 100 μηδενικά δείγματα ανάμεσα σε κάθε ψηφίο και αποθηκεύουμε το σήμα σε αρχείο με όνομα tone_sequence.wav.

2.4 Απομονώνουμε κάθε υποσήμα που αντιστοιχεί σε ψηφίο από την αλληλουχία που δημιουργήσαμε νωρίτερα και υπολογίζουμε τον DFT καθενός από αυτά χρησιμοποιώντας παράθυρο Hamming και τετραγωνικό παράθυρο. Οι DFT παρατίθενται στον κώδικα.

2.5 Χρησιμοποιώντας τη συνάρτηση argmax που βρίσκεται στη βιβλιοθήκη numpy βρίσκουμε τα αντίστοιχα indexes των touch-tone συχνοτήτων κάθε υποσήματος που αντιστοιχεί σε πλήκτρο αφού πρώτα υπολογίσουμε τον DFT του. Αυτά τα indexes τα αποθηκεύουμε σε έναν πίνακα 20 θέσεων, του οποίου τα δύο πρώτα στοιχεία αντιστοιχούν στο πλήκτρο 0, τα επόμενα δύο στο 1 κ.ο.κ.

Ακολουθεί ο εν λόγω πίνακας:

$$[163, 164, 85, 148, 85, 163, 85, 180, 94, 148, 94, 163, 94, 180, 104, 148, 104, 163, 104, 180]$$

2.6 Δημιουργούμε συνάρτηση που παίρνει ως όρισμα μια αλληλουχία τηλεφωνικών τόνων και επιστρέφει λίστα των πλήκτρων που παράγουν την αλληλουχία αυτή.

Ακολουθήσαμε τα εξής βήματα:

1. Απομονώνουμε κάθε υποσήμα της αλληλουχίας με κατάλληλο χρονικό παράθυρο
2. Υπολογίζουμε DFT 1000 σημείων κάθε υποσήματος
3. Βρίσκουμε τα indexes των touch-tone συχνοτήτων του
4. Συγκρίνουμε τα indexes με τα στοιχεία του πίνακα που υπολογίσαμε στο προηγούμενο ερώτημα
5. Αποφαινόμαστε βάσει της σύγκρισης για το ποιο πλήκτρο αντιστοιχεί σε κάθε υποσήμα

2.7 Με σκοπό να αποκωδικοποιήσουμε τις δοθείσες αλληλουχίες χρησιμοποιήσαμε τη συνάρτηση που δημιουργήσαμε στο προηγούμενο ερώτημα και προέκυψαν τα εξής αποτελέσματα:

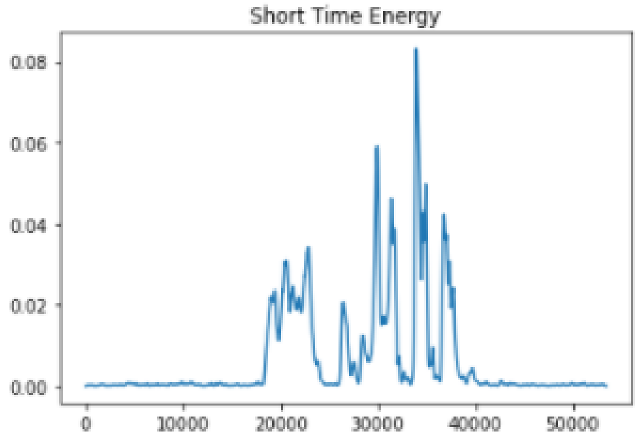
1. Easy Sig: [7, 3, 5, 8, 2, 8, 0, 2]

2. Hard Sig: [4, 3, 5, 6, 6, 2, 0, 9, 9, 5]

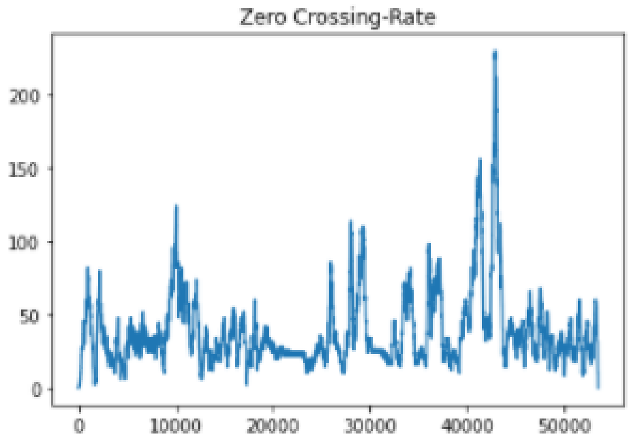
Μέρος 3

3.1 Σε σήμα φωνής μετράμε την ενέργεια βραχέος χρόνου και το ρυθμό εναλλαγής προσήμου χρησιμοποιώντας παράθυρο μήκους $N = 20\ ms$. Βάσει αυτών των μετρήσεων μπορούμε να διαχωρίσουμε έμφωνους και άφωνους ήχους. Συγκεκριμένα, στους έμφωνους ήχους αντιστοιχεί μεγάλη ενέργεια βραχέος χρόνου και μικρός ρυθμός εναλλαγής προσήμου, ενώ αντίθετα σε άφωνους ήχους αντιστοιχεί μικρή ενέργεια βραχέος χρόνου και μεγάλος ρυθμός εναλλαγής προσήμου.

Αυξάνοντας το μήκος του παραθύρου παρατηρούμε ότι η διάκριση έμφωνων και άφωνων ήχων γίνεται όλο και πιο δυσχερής, καθώς έμφωνοι και άφωνοι ήχοι μπλέκονται μεταξύ τους.



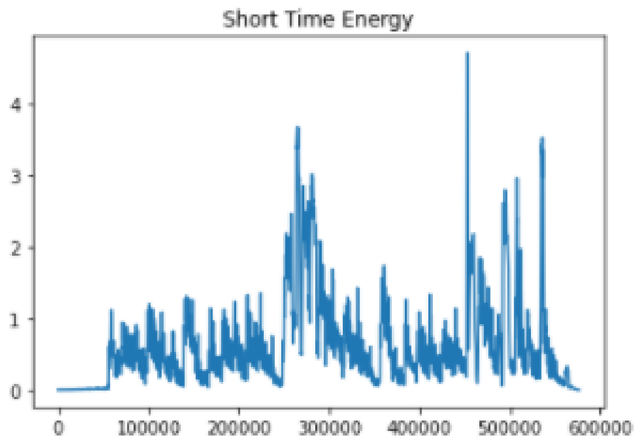
Ενέργεια βραχέος χρόνου



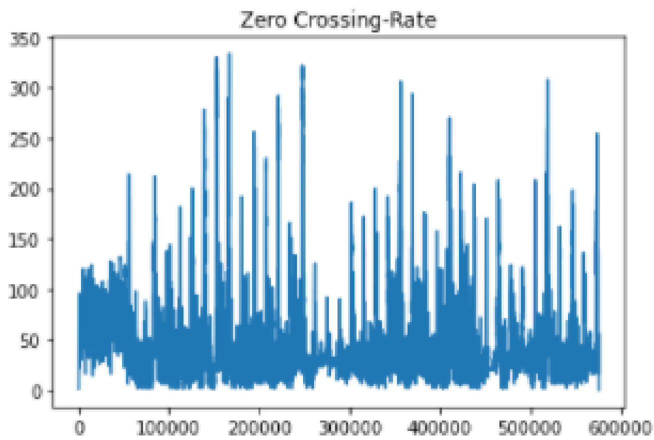
Ρυθμός εναλλαγής προσήμου

3.2 Σε σήμα μουσικής μετράμε την ενέργεια βραχέος χρόνου και το ρυθμό εναλλαγής προσήμου χρησιμοποιώντας παράθυρο μήκους $N = 20\ ms$. Βάσει αυτών των μετρήσεων μπορούμε να διαχωρίσουμε τη σιωπή από τη μουσική. Πιο συγκεκριμένα, αρχικά παρατηρούμε σχετικά μικρή ενέργεια βραχέος χρόνου και μεγάλο ρυθμό εναλλαγής προσήμου καθώς αρχικά έχουμε τα ρυθμικά όργανα σε χαμηλή ένταση, ενώ αργότερα η ενέργεια αυξάνεται και ο ρυθμός εναλλαγής προσήμου ελαττώνεται καθώς ξεκινά να παίζει το μελωδικό όργανο σε μεγαλύτερη ένταση.

Αυξάνοντας το μήκος του παραθύρου παρατηρούμε ότι η διάκριση ρυθμικών και μελωδικών ήχων γίνεται όλο και πιο δυσχερής, καθώς μπλέκονται μεταξύ τους.



Ενέργεια βραχέος χρόνου



Ρυθμός εναλλαγής προσήμου