



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΔΕΥΤΕΡΗ ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΝΑΦΟΡΑ
ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΣΗΜΑΤΟΣ

Κωνσταντίνος Κονιαβίτης
03119182

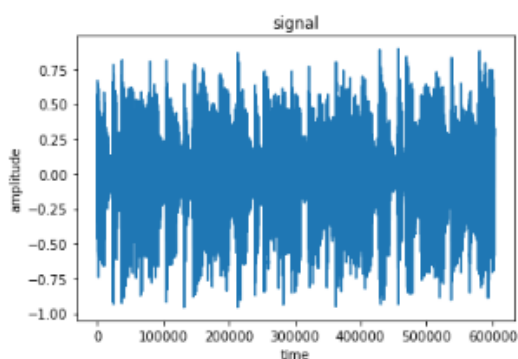
Αναστασία Χριστίνα Λίβα
03119029

Περιεχόμενα

Μέρος 1	2
Μέρος 2	7

Μέρος 1

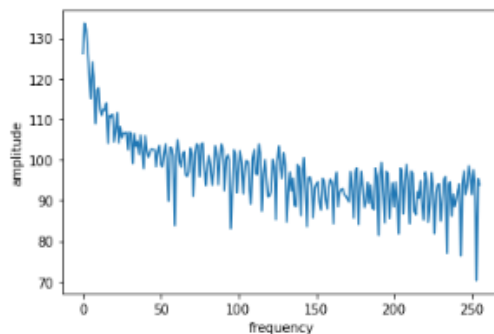
1.0 Χρησιμοποιώντας την εντολή `read` από το πακέτο `soundfile` διαβάζουμε το σήμα `"music.wav"` κι έπειτα το κανονικοποιούμε σε όλο του το μήκος διαιρώντας τα δείγματά του με την απόλυτη τιμή του ίδιου σήματος έτσι ώστε καθόλη τη διάρκειά του οι τιμές του να κυμαίνονται στο διάστημα $[-1, 1]$. Στη συνέχεια χρησιμοποιώντας παράθυρο Hamming παραθυροποιούμε το σήμα ανά 512 δείγματα και πραγματοποιούμε ανάλυση για κάθε πλαίσιο.



Σχήμα 1: Απεικόνιση μουσικού σήματος πριν την κανονικοποίηση

1.1 Αφού υπολογίσουμε τον DFT κάθε παραθυροποιημένου σήματος βρίσκουμε το φάσμα ισχύος $P(k)$ σύμφωνα με τον ακόλουθο τύπο:

$$P(k) = PN + 10 \log_{10} \left| \sum_{n=0}^{N-1} w(n)x(n)e^{-j\frac{2\pi kn}{N}} \right|^2, \quad 0 \leq k \leq \frac{N}{2}$$

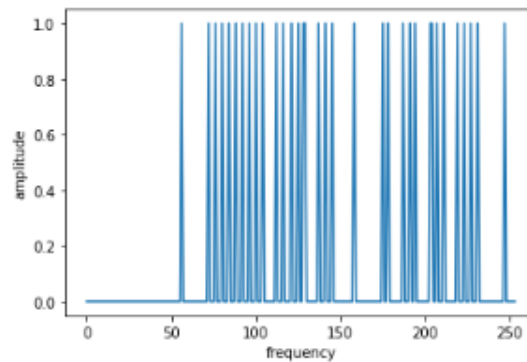


Σχήμα 2: Φάσμα Ισχύος

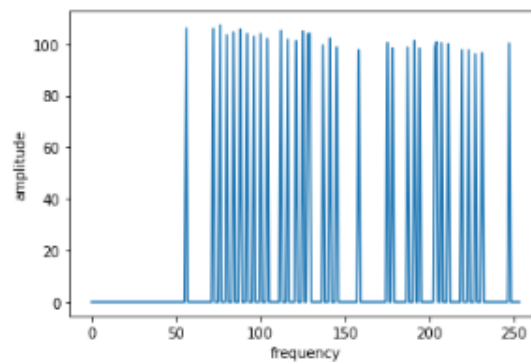
Παραθέτουμε ενδεικτικά το φάσμα ισχύος για ένα από τα παραθυροποιημένα σήματα. Το φάσμα ισχύος δίνει την ισχύ κάθε συχνότητας, δηλαδή την "δύναμη" της κάθε συχνότητας.

1.2 Εντοπίζουμε ανά critical band τοπικά μέγιστα τα οποία είναι μεγαλύτερα από τις γειτονικές τους συχνότητες τουλάχιστον κατά 7 dB. Οι συχνότητες που αντιστοιχούν στα τοπικά μέγιστα είναι ισχυρότερες, οπότε υπερκαλύπτουν τις γειτονικές τους.

Έτσι, υπολογίζουμε τη συνάρτηση $S_T(k)$, η οποία έχει τιμές 1 εάν στη θέση k υπάρχει τοπικό μέγιστο και 0 σε κάθε άλλη περίπτωση. Στη συνέχεια υπολογίζουμε την ισχύ των τοπικών μεγίστων $P_{TM}(k)$ παραθέτουμε ενδεικτικά τα διαγράμματα $S_T(k)$ και $P_{TM}(k)$ για ένα από τα παραθυροποιημένα σήματα.

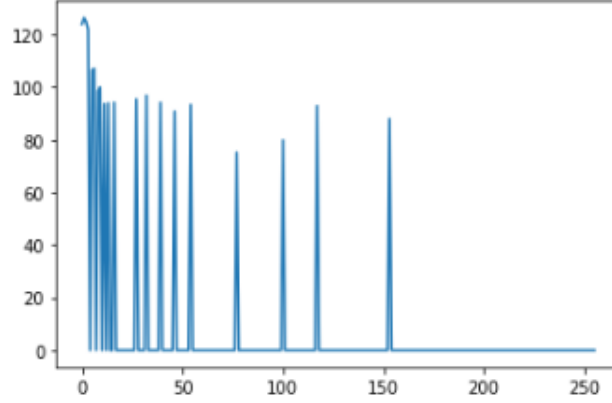
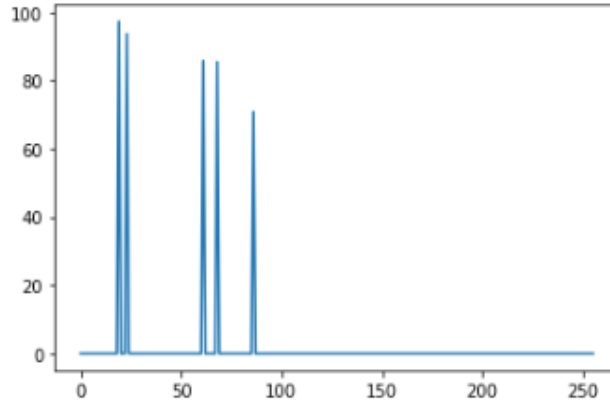


Σχήμα 3: $S_T(k)$



Σχήμα 4: $P_{TM}(k)$

1.3 Φορτώνουμε τις τονικές μάσκες P_{TMc} και τις μάσκες θορύβου P_{NMc} και αναστρέφουμε τους δύο πίνακες.

Σχήμα 5: P_NMc Σχήμα 6: P_TMc

1.4 Υπολογίζουμε τα δύο κατώφλια κάλυψης σύμφωνα με τις εξισώσεις

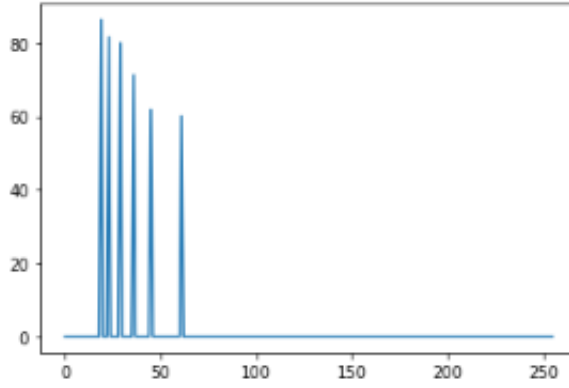
$$T_{TM}(i, j) = P_{TM}(j) - 0,275B(j) + SF(i, j) - 6,025$$

$$T_{NM}(i, j) = P_{NM}(j) - 0,175B(j) + SF(i, j) - 2,025$$

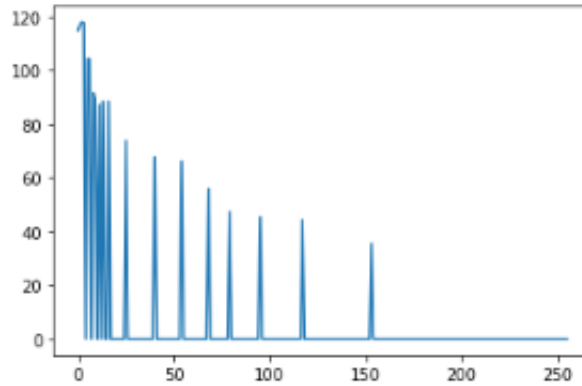
όπου το SF δίνεται από:

$$SF(i, j) = \begin{cases} 17\Delta_b - 0.4P_{TM}(j) + 11, & -3 \leq \Delta_b < -1 \\ (0.4P_{TM}(j) + 6)\Delta_b, & -1 \leq \Delta_b < 0 \\ -17\Delta_b, & 0 \leq \Delta_b < 1 \\ (0.15P_{TM}(j) - 17)\Delta_b - 0.15P_{TM}(j), & 1 \leq \Delta_b < 8 \end{cases}$$

Ακολουθούν τα διαγράμματα των κατωφλιών:



Σχήμα 7: $T_{TM}(k)$

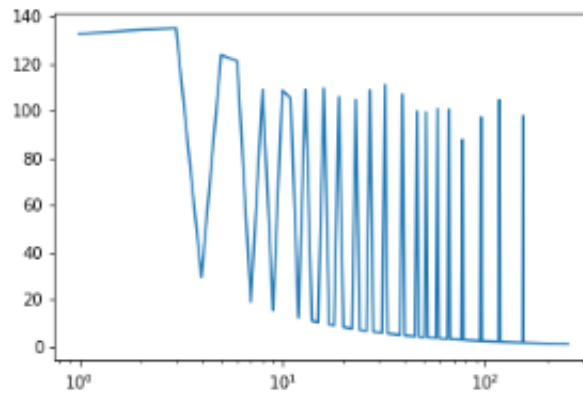


Σχήμα 8: $T_{NM}(k)$

1.5 Χρησιμοποιώντας την ακόλουθη εξίσωση υπολογίζουμε ένα συνολικό κατώφλι κάλυψης συνδυάζοντας το κατώφλι τόνου και το κατώφλι θορύβου που υπολογίσαμε στο προηγούμενο ερώτημα.

$$T_g(i) = 10 \log_{10} \left(10^{0.1T_q(i)} + \sum_{l=1}^L 10^{0.1T_{TM}(i,\ell)} + \sum_{m=1}^M 10^{0.1T_{NM}(i,m)} \right) \text{ dB SPL},$$

Με αυτόν τον τρόπο θα χρησιμοποιήσουμε το κατώφλι ώστε να μην κωδικοποιήσουμε συχνότητες που δεν είναι αντιληπτές από το ανθρώπινο αυτί καθώς επίσης και θορυβώδη δείγματα. Ακολουθεί απεικόνιση του κατωφλίου.



Σχήμα 9: $T_g(k)$

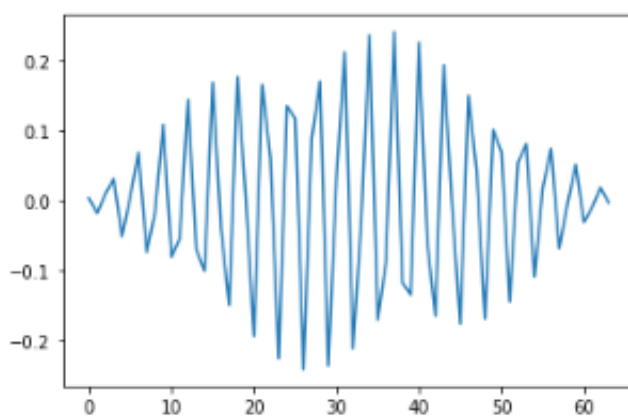
Μερος 2

2.0 Δημιουργούμε τα φίλτρα σύνθεσης $h_k(n)$ και $g_k(n)$, τα οποία παραμετροποιούνται από την παράμετρο k σύμφωνα με την ακόλουθη εξίσωση.

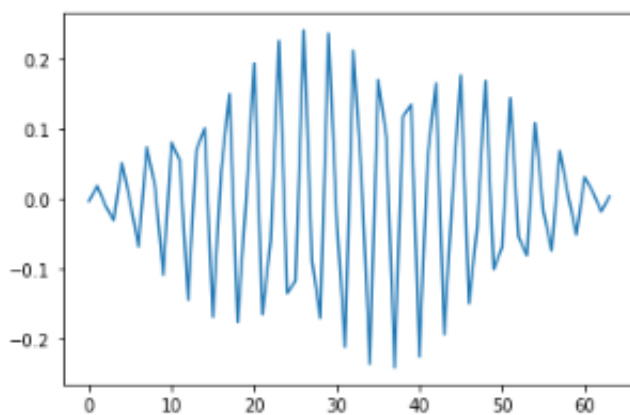
$$h_k(n) = \sin \left[\left(n + \frac{1}{2} \right) \frac{\pi}{2M} \right] \sqrt{\frac{2}{M}} \cos \left[\frac{(2n + M + 1)(2k + 1)\pi}{4M} \right]$$

$$g_k(n) = h_k(2n - 1 - N)$$

Ακολουθούν ενδεικτικά οι απεικονίσεις των φίλτρων για $k = 20$



Σχήμα 10: Φίλτρο h για $k = 20$

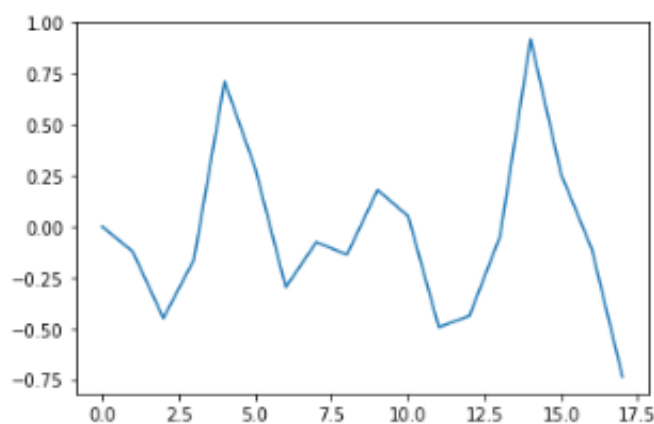


Σχήμα 11: Φίλτρο g για $k = 20$

2.1 Πραγματοποιούμε συνέλιξη του κάθε παραθυροποιημένου σήματος με καθένα από τα 32 φίλτρα h_k και στη συνέχεια κάθε μία συνέλιξη υποδειγματολειτουργείται κατά παράγοντα $M = 32$ άρα το τελικό μας σήμα τελικά αποτελείται από μόνο $\frac{512+64}{32} = 18$ δείγματα. Αυτό το αποτέλεσμα είναι λογικό καθώς το μήκος μιας διακριτής συνέλιξης ισούται με το άθροισμα των διαρκειών των δύο συνελισσόμενων σημάτων + 1.

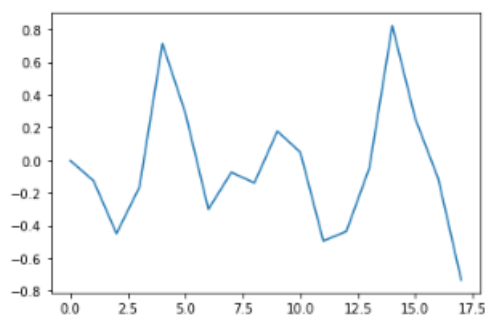
$$y_k(n) = u_k(m \cdot n)$$

Ακολουθεί το διάγραμμα του y_k πριν την κβάντιση



Σχήμα 12: y_k πριν την κβάντιση

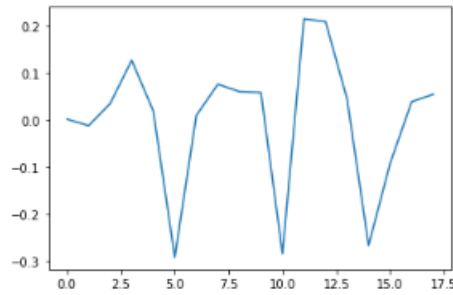
2.2 Ακολουθεί το διάγραμμα του y_k μετά την κβάντιση με μη προσαρμοζόμενο κβαντιστή των 8 bits.



Σχήμα 13: y_k μετά την κβάντιση με μη προσαρμοζόμενο κβαντιστή των 8 bits

Ακολουθεί το διάγραμμα του y_k μετά την κβάντιση με προσαρμοζόμενο κβαντιστή του οποίου τα bits προκύπτουν από την εξίσωση:

$$B_k = \text{int} \left(\log_2 \left(\frac{R}{\min(T_g(i))} \right) - 1 \right)$$



Σχήμα 14: y_k μετά την κβάντιση με προσαρμοζόμενο κβαντιστή

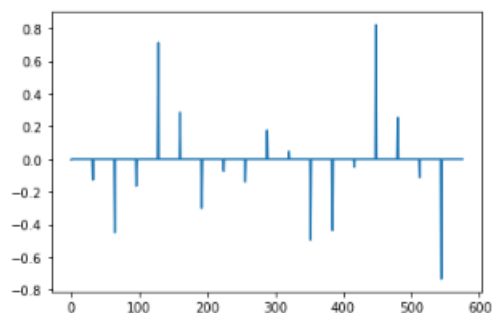
Ο κβαντιστής που χρησιμοποιήθηκε για το παραπάνω διάγραμμα υλοποιήθηκε βάσει του εξής αλγορίθμου.

Αρχικά ορίζουμε μία συνάρτηση y που παίρνει ως παραμέτρους το σήμα προς κβάντιση και τον επιθυμητό ρυθμό από bits. Στη συνέχεια ορίζουμε

$$\Delta = \frac{\max(y) - \min(y)}{L}$$

όπου $L = 2^B$. Βάσει του Δ ορίζονται L ίσα επίπεδα κβαντισμού τα οποία απέχουν μεταξύ τους κατά Δ . Έπειτα για κάθε δείγμα του σήματος βρίσκουμε το κοντινότερό του επίπεδο κβαντισμού και αντικαθιστά το κάθε δείγμα με το επίπεδο κβαντισμού που του αντιστοιχεί. Στο τέλος επιστρέφεται η κβαντισμένη πλέον ακολουθία.

2.3 Πραγματοποιούμε υπερδειγματοληψία του σήματος κατά παράγοντα M με τον εξής τρόπο: Προσθέτουμε M μηδενικά ανάμεσα σε κάθε δείγμα του $y_k(n)$

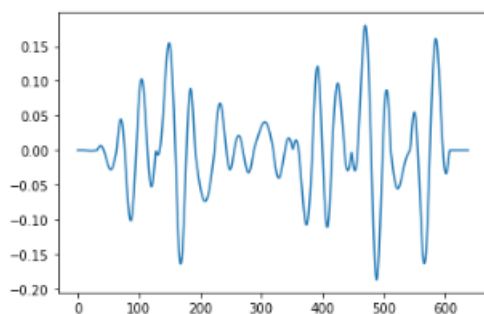
Σχήμα 15: Υπερδειγματολειτουργημένο σήμα $w_n(k)$

Στη συνέχεια προσθέτουμε τα σήματα που προέκυψαν από τις συνελίξεις μεταξύ τους οπότε έχουμε πάλι τον ίδιο αριθμό παραθυρομένων σημάτων.

Πραγματοποιούμε συνέλιξη του κάθε υπερδειγματολειτουργημένου σήματος κατα παράγοντα $M = 32$ με καθένα από τα 32 φίλτρα g_k .

$$y_k(n) = u_k(m \cdot n)$$

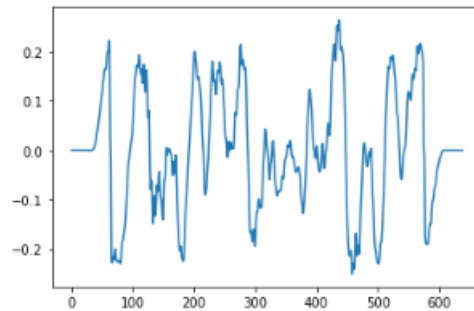
Ακολουθεί απεικόνιση ενός εκ των σημάτων μετά τη συνέλιξη αλλά πριν την πρόσθεση.



Σχήμα 16: Σήμα πριν την πρόσθεση

Ακολουθεί απεικόνιση ενός εκ των σημάτων μετά τη συνέλιξη και την πρόσθεση.

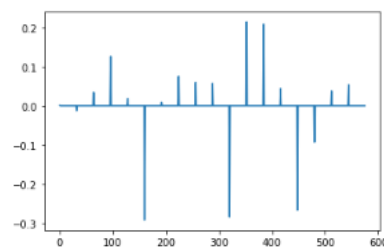
Η τελική ανακατασκευή του σήματος μουσικής \hat{s} γίνεται με εφαρμογή της τεχνικής OverLap - Add, όπως εξηγήθηκε στη θεωρία του Block Convolution, χωρίς επικάλυψη μεταξύ των διαδοχικών πλαισίων ανάλυσης. Στη συνέχεια ακούμε το σήμα που



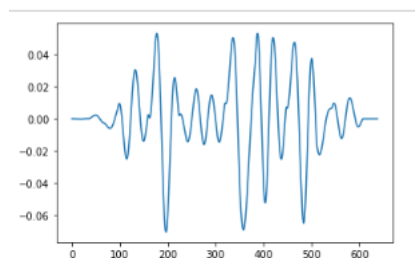
Σχήμα 17: Σήμα μετά την πρόσθεση

βρίσκεται στο υποβληθέν .zip. Σε αυτή την περίπτωση παρατηρούμε πολύ θόρυβο στο ανακατασκευασμένο μουσικό σήμα καθώς ο κβαντιστής που έχει χρησιμοποιηθεί είναι ΜΗ προσαρμοζόμενος.

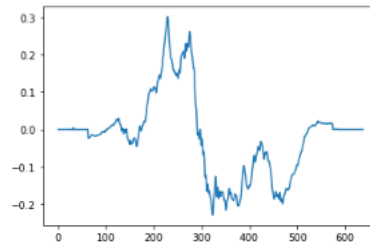
Η ίδια διαδικασία ακολουθείται και για το σήμα που προέκυψε από τον προσαρμοζόμενο κβαντιστή.



Σχήμα 18: Υπερδειγματοληπτημένο Σήμα



Σχήμα 19: Σήμα πριν την πρόσθεση



Σχήμα 20: Σήμα μετά την πρόσθεση

Συμπερασματικά, ακούμε το τελικό σήμα το οποίο έχει ακόμα θόρυβο, ο οποίος ωστόσο είναι λιγότερος από το τελικό σήμα που προκύπτει από το μη προσαρμοζόμενο κβαντιστή.

Το ποσοστό συμπίεσης με τον μη προσαρμοζόμενο κβαντιστή είναι $\frac{8}{16} \cdot 100\% = 50\%$ ενώ με τον προσαρμοζόμενο είναι $\frac{10}{16} \cdot 100\% = 62,5\%$