



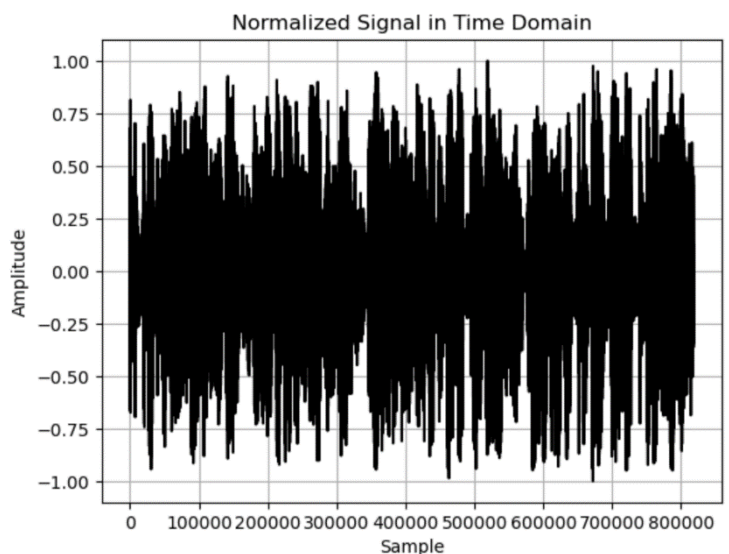
Ψηφιακή Επεξεργασία Σήματος

2^η Εργαστηριακή Άσκηση

Θέμα: Κωδικοποίηση σημάτων Μουσικής βάσει του
ψυχοακουστικού μοντέλου (*Perceptual Audio Coding*)

Μέρος 1^ο : Ψυχοακουστικό Μοντέλο 1

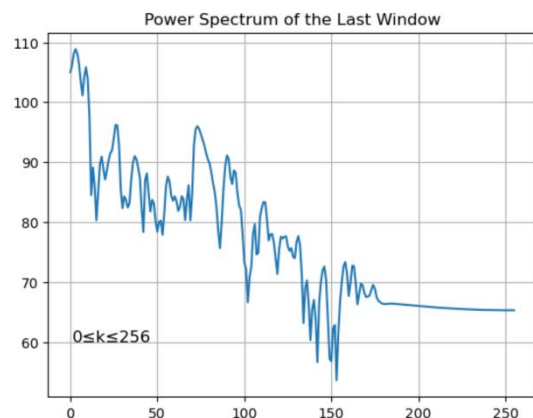
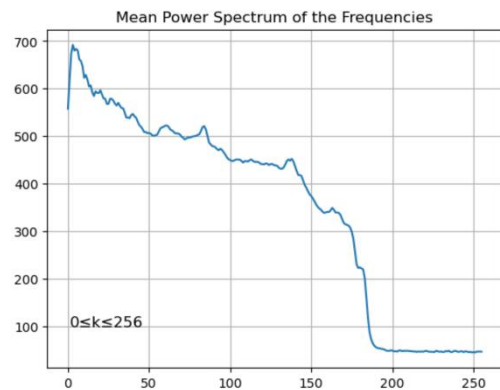
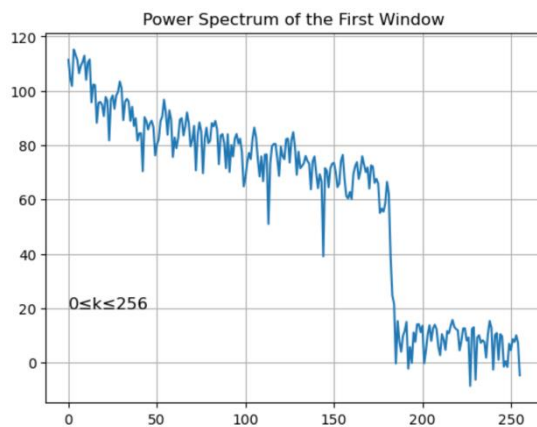
Βήμα 1.0: Μετατρέπουμε το σήμα από stereo format σε mono και το κανονικοποιούμε έτσι ώστε καθ' όλη τη διάρκειά του να λαμβάνει τιμές στο διάστημα $[-1,1]$.



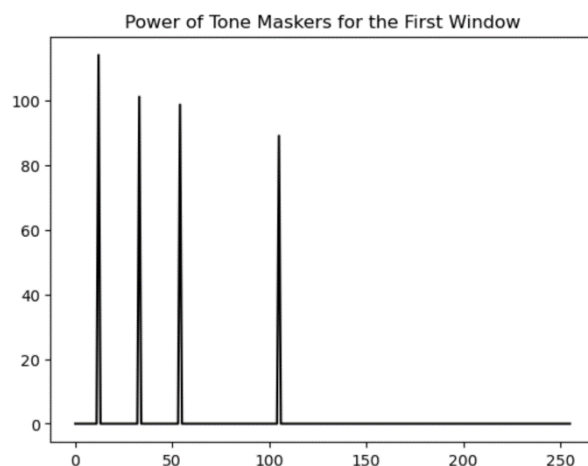
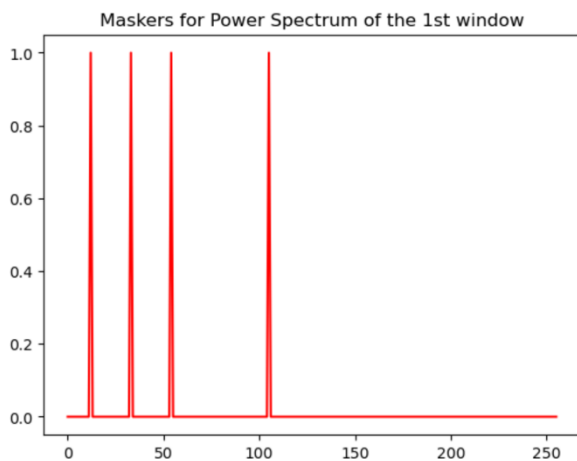
Βήμα 1.1: Παραθυρώνουμε το σήμα χρησιμοποιώντας παράθυρα hanning μήκους 512 δειγμάτων και για κάθε παράθυρο υπολογίζουμε το 512 σημείων φάσμα ισχύος $P(K)$. Σημειώνεται ότι επειδή το τελευταίο παράθυρο του σήματος είχε λιγότερα από 512 δείγματα προσθέσαμε τον απαραίτητο αριθμό μηδενικών δειγμάτων σε αυτό.



Ενδεικτικά απεικονίζουμε τα μονόπλευρα φάσματα ισχύος ($0 \leq k \leq 256$) για το πρώτο και το τελευταίο από τα 1598 παράθυρα, καθώς επίσης και το “μέσο φάσμα ισχύος” όλων των παραθύρων.

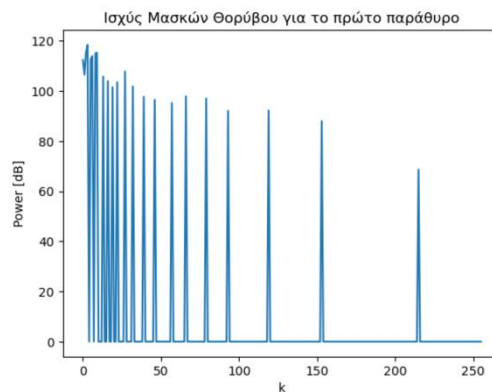


Βήμα 1.2: Με τη χρήση των κατάλληλων σχέσεων υλοποιούμε τη συνάρτηση $St(k)$ για κάθε ένα από τα 1598 παράθυρα. Η συνάρτηση αυτή επιστρέφει 1 αν στη θέση k υπάρχει τονική μάσκα και 0 αν δεν υπάρχει. Ακολούθως βρίσκουμε την ισχύ της κάθε τονικής μάσκας. Τα αποτελέσματα για το 1^ο παράθυρο παρουσιάζονται πιο κάτω:

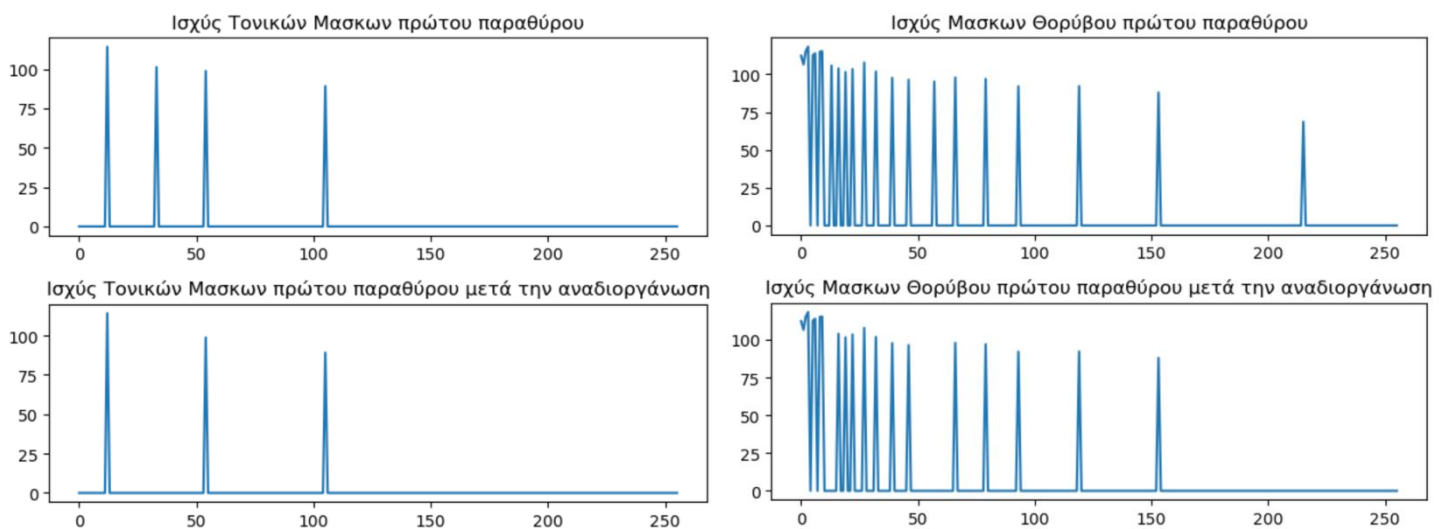




Επιπλέον, με τη χρήση του δοσμένου πίνακα P_{NM} απεικονίζουμε και την ισχύ των μασκών θορύβου για το πρώτο παράθυρο.



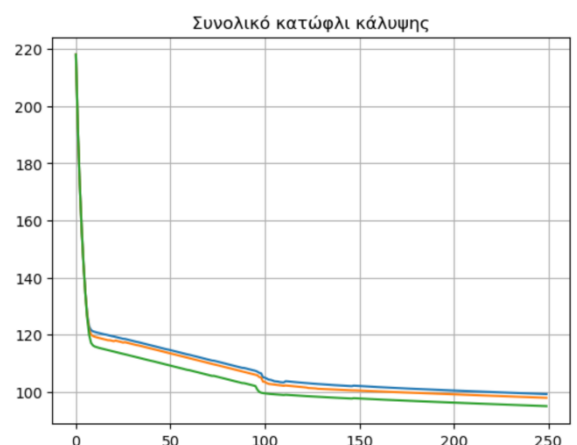
Βήμα 1.3: Με τη χρήση των πινάκων P_{TM} και P_{NM} βρίσκουμε την ισχύ των τονικών μασκών και των μασκών θορύβου μετά την αναδιοργάνωση. Και πάλι παρουσιάζονται τα αποτελέσματα για το 1^ο παράθυρο:



Βήματα 1.4 & 1.5: Αρχικά υπολογίζουμε για κάθε ένα από τα 1598 παράθυρα τα δύο διαφορετικά κατώφλια κάλυψης $T_{TM}(i,j)$ και $T_{NM}(i,j)$. Οι πίνακες αυτοί περιέχουν το ποσοστό κάλυψης στο σημείο i από τη μάσκα (τονική ο πίνακας T_{TM} και θορύβου ο πίνακας T_{NM}) στο σημείο j .

Χρησιμοποιώντας τους υπολογισμούς μας για αυτούς τους 2 πίνακες αλλά και μια συνάρτηση υπολογισμού του ATH (Absolute Threshold of Hearing) που έχουμε υλοποιήσει επιτυγχάνουμε τον υπολογισμό του συνολικού κατωφλίου κάλυψης T_g (Global Masking Threshold).

Δίπλα φαίνονται σχηματικά τα συνολικά κατώφλια για τα παράθυρα 2 (μπλε), 500 (πορτοκαλι) και 1000 (πράσινο).

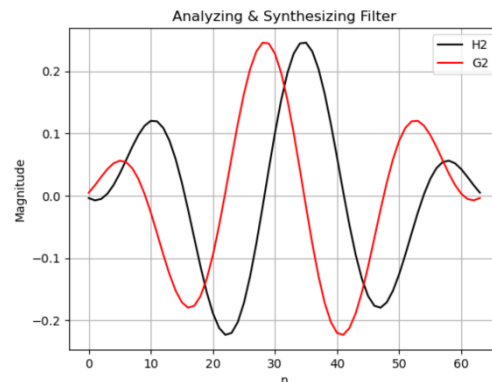




Μέρος 2^ο : Χρονο-Συχνотική Ανάλυση με Συστοιχία Ζωνοπερατών Φίλτρων

Βήμα 2.0: Ορίζουμε τα 32 φίλτρα ανάλυσης και σύνθεσης και αποθηκεύουμε τις κρουστικές τους αποκρίσεις σε 2 πίνακες 32x64.

Στο διπλανό σχήμα φαίνονται οι κρουστικές αποκρίσεις για τα φίλτρα $h_2[n]$ και $g_2[n]$.



Page | 4

Βήματα 2.1, 2.2 & 2.3: Όλα τα βήματα στην επεξεργασία και ανακατασκευή του σήματος (φιλτράρισμα με φίλτρα ανάλυσης, decimation, κβάντιση, interpolation και φιλτράρισμα με φίλτρα σύνθεσης) γίνονται εντός του ίδιου βρόχου (for loop).

Αρχικά κάθε παράθυρο $s[n]$ μήκους 512 δειγμάτων εισέρχεται στη συστοιχία ζωνοπερατών φίλτρων και πραγματοποιούνται 32 συνελίξεις (μια με κάθε φίλτρο). Το αποτέλεσμα $u[n]$ που προκύπτει από κάθε συνέλιξη έχει μήκος 575 δείγματα ενώ μετά τον αποδεκατισμό του κατά παράγοντα $M=32$ η ακολουθία $y[n]$ έχει μήκος 18.

Το επόμενο στάδιο είναι ο υπολογισμός των διαφόρων παραμέτρων που απαιτεί ο προσαρμοζόμενος κβαντιστής και ακολούθως η κβάντιση καθεαυτή.

Από εδώ και στο εξής ο δείκτης 1 χρησιμοποιείται για να υποδείξει σήμα που έχει κβαντιστεί με τον προσαρμοζόμενο κβαντιστή και ο δείκτης 2 για σήμα που έχει κβαντιστεί με τον μη-προσαρμοζόμενο κβαντιστή.

Αμφότερα τα σήματα αυτά περνάνε από το στάδιο του interpolation κατά ακέραιο συντελεστή $M=32$ και στη συνέχεια από τα φίλτρα σύνθεσης $g_k[n]$. Οι 2 ακολουθίες $Shat1$ και $Shat2$ που προκύπτουν στην έξοδο της συστοιχίας έχουν μήκος 639.

```
for wn in range ((len(x)//len(w))): #1598 Loops for 1598 windows

    Shat_i1=np.zeros(18*M+L-1)
    Shat_i2=np.zeros(18*M+L-1)
    s = w*x[wn*512:(wn+1)*512]

    for i in range (len(hk)): #32 Loops for 32 filters

        u = np.convolve(s,hk[i])#Bhma 2.1 (i)
        y=downsampling(u)# Bhma 2.1(ii)

        lowerlimitfreq = (2*i)*fs/(2*M)
        upperlimitfreq = (2*i+2)*fs/(2*M)
        min_Tgi = min_Tgi_funct(upperlimitfreq, lowerlimitfreq,Tg[wn],Tg_expanded[wn])

        bk1 = int(np.log2(R/min_Tgi)-1)

        y1=y
        y2=y
        ymax=max(y)
        ymin=min(y)
        step1 = (ymax-ymin)/(2*bk1)
        #quantization
        for z in range (18):
            val = y[z]
            y1[z] = quantize(val,ymin,step1)
            y2[z] = quantize(val,-1,step2)

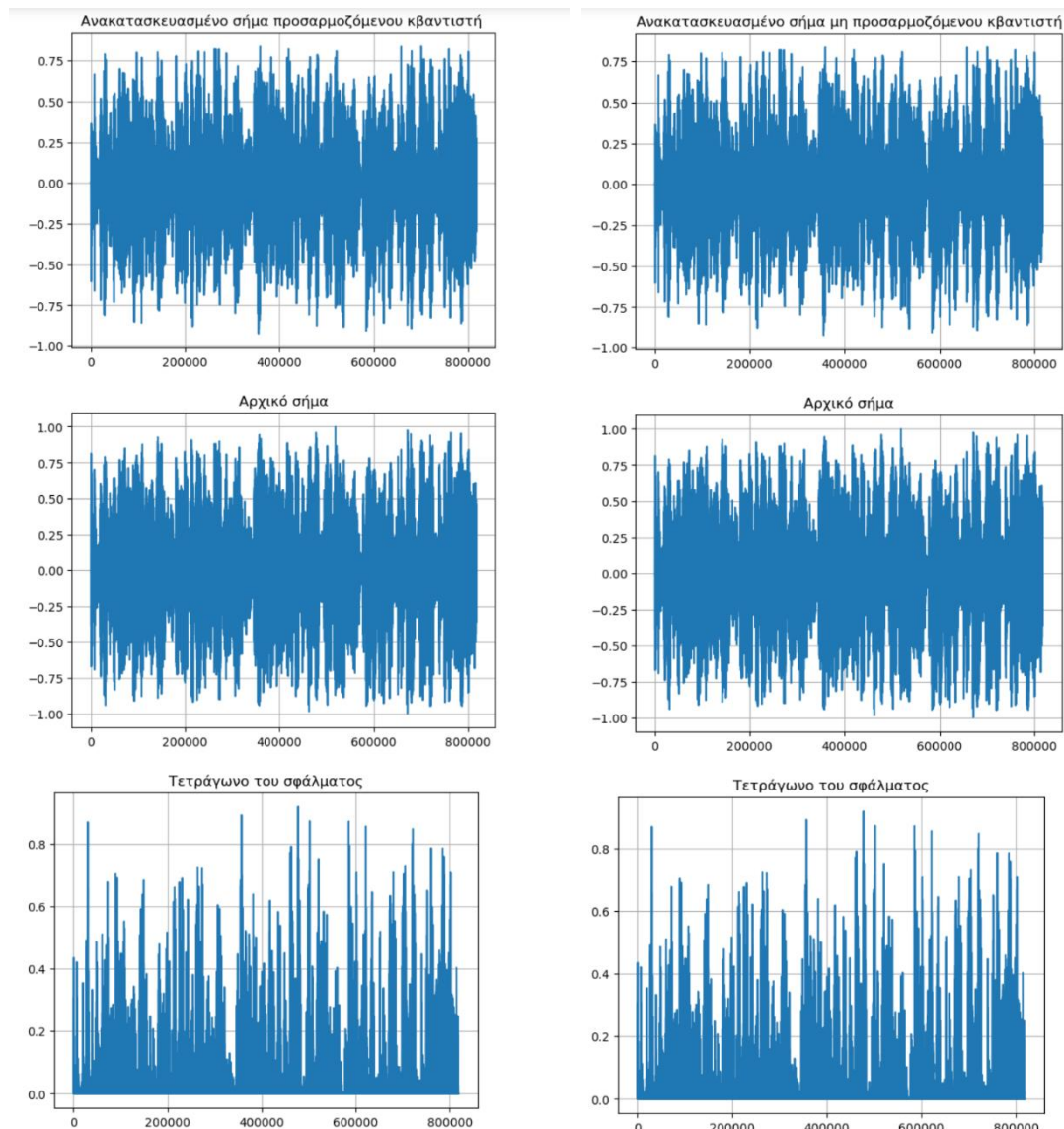
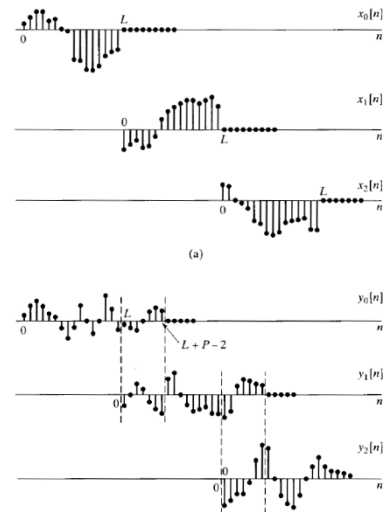
        #interpolation:
        yhat2=interpolation(y2)
        yhat1=interpolation(y1)

        #convolution with G(z):
        Shat_i_j1=np.convolve(yhat1,gk[i])
        Shat_i_j2=np.convolve(yhat2,gk[i])
        Shat_i1=Shat_i1+Shat_i_j1
        Shat_i2=Shat_i2+Shat_i_j2
    Shat1.append(Shat_i1)
    Shat2.append(Shat_i2)
```



Οι 1598 ακολουθίες με μήκος 639 που προκύπτουν προφανώς έχουν μεγαλύτερο μήκος από τα αρχικά πλαίσια ανάλυσης των 512 δειγμάτων. Για τον λόγο αυτό πρέπει να προστεθούν μεταξύ τους με τη μέθοδο Overlap-Add (μια οπτικοποίηση της μεθόδου μπορεί να φανεί στο διπλανό σχήμα). Πιο συγκεκριμένα, τα 127 τελευταία στοιχεία της ακολουθίας ενός παραθύρου θα προστίθονται με τα 127 πρώτα στοιχεία της ακολουθίας του αμέσως επόμενου παραθύρου.

Τα αποτελέσματα της ανακατασκευής φαίνονται παρακάτω:





Για να κάνουμε τη σύγκριση των 2 σημάτων (του αρχικού και του ανακατασκευασμένου) λάβαμε υπόψη την χρονική καθυστέρηση που εισάγουν τα φίλτρα σύνθεσης. Η καθυστέρηση αυτή ισούται με 64, όσο δηλαδή και το μήκος των κρουστικών αποκρίσεων των φίλτρων σύνθεσης.

Για τον ομοιόμορφο προσαρμοζόμενο κβαντιστή το Μέσο Τετραγωνικό Σφάλμα υπολογίστηκε ως:

Mean Square Error = 0.019026. Το ποσοστό συμπίεσης είναι: $rate = 1 - \frac{6620160}{16 \cdot 818176} = 0.49429$

Page | 6

Για τον μη-προσαρμοζόμενο κβαντιστή το Μέσο Τετραγωνικό Σφάλμα υπολογίστηκε ως: Mean

Square Error = 0.019026. Το ποσοστό συμπίεσης είναι: $rate = 1 - \frac{6545408}{16 \cdot 818176} = 0.500$

Σημείωση: Είναι φανερό ότι τα αποτελέσματα στην περίπτωση του προσαρμοζόμενου κβαντιστή δεν είναι τα αναμενόμενα. Θα περιμέναμε για παράδειγμα ότι το μέσο τετραγωνικό σφάλμα θα ήταν μικρότερο απ' ό,τι είναι στον μη-προσαρμοζόμενο κβαντιστή αλλά παραδόξως τα 2 σφάλματα είναι ίσα.

Κατά πάσα πιθανότητα έχει γίνει από πλευράς μας κάποιο προγραμματιστικό λάθος, όσον αφορά την κβάντιση με τον προσαρμοζόμενο κβαντιστή, το οποίο όμως δεν καταφέραμε να εντοπίσουμε παρά τις προσπάθειές μας. Εντούτοις συνεχίσαμε την επεξεργασία και τους υπολογισμούς των σημάτων (εννοώντας ότι δεν περιοριστήκαμε στην υλοποίηση μόνο του ενός κβαντιστή) διότι πιστεύουμε ότι η μεθοδολογία μας είναι ορθή κατά το μεγαλύτερο μέρος της.