

[Εξηγείστε περιεκτικά και επαρκώς την εργασία σας. Επιτρέπεται προαιρετικά η συνεργασία εντός ομάδων των 2 ατόμων. Κάθε ομάδα 2 ατόμων υποβάλλει μια κοινή αναφορά που αντιπροσωπεύει μόνο την προσωπική εργασία των μελών της. Αν χρησιμοποιήσετε κάποια άλλη πηγή εκτός του βιβλίου και του εκπαιδευτικού υλικού του μαθήματος, πρέπει να το αναφέρετε. Η παράδοση της αναφοράς και του κώδικα της εργασίας θα γίνει ηλεκτρονικά στο mycourses.ntua.gr και επιπλέον η αναφορά της εργασίας θα παραδίδεται τυπωμένη και προσωπικά στην γραμματεία του εργαστηρίου Ρομποτικής (2.1.12, παλαιό Κτ.Ηλεκ.), ώρες 09.00-14.30.

Θέμα: Εισαγωγή στην Ψηφιακή Επεξεργασία Σημάτων με MATLAB και Εφαρμογές σε Ακουστικά Σήματα

Μέρος 1ο - Σύστημα Εντοπισμού Τηλεφωνικών Τόνων (Telephone Touch – Tones)

Στόχος αυτής της άσκησης είναι να παρουσιάσει πως λειτουργεί το τηλεφωνικό τονικό σύστημα χρησιμοποιώντας σήματα διαφορετικών συχνοτήτων για να εντοπίζει ποιο πλήκτρο έχει πατηθεί. Ο εντοπισμός αυτών των συχνοτήτων μπορεί να γίνει με την χρήση του Διακριτού Μετ/σμού Fourier (DFT) $X[k]$ του τηλεφωνικού σήματος $x[n]$. Με το πάτημα ενός πλήκτρου στο τηλέφωνο ακούγεται ένας ήχος που είναι το άθροισμα 2 ημιτόνων, το υψίσυχο ημίτονο δείχνει την στήλη που ανήκει το πλήκτρο στο touch-pad της τηλεφωνικής συσκευής και το χαμηλόσυχο ημίτονο δείχνει την αντίστοιχη γραμμή του touch-pad σύμφωνα με τον Πίνακα που ακολουθεί. Παραδείγματος χάρη, το πλήκτρο 5 αντιστοιχεί στο σήμα $d_5[n] = \sin(0.5906n) + \sin(1.0247n)$

	Ω_{column}		
Ω_{row}	0.9273	1.0247	1.1328
0.5346	1	2	3
0.5906	4	5	6
0.6535	7	8	9
0.7217		0	

Πίνακας 1: Διακριτές Συχνότητες για Τηλεφωνικούς Τόνους για Συχνότητα Δειγματοληψίας 8192 Hz.

- 1.1. Δημιουργείτε τους 10 διαφορετικούς τόνους σύμφωνα με την παραπάνω εκφώνηση. Κάθε τόνος πρέπει να έχει μήκος 1000 δείγματα. (Υπόδειξη: Μπορείτε να χρησιμοποιήσετε την ρουτίνα **sound()** του MATLAB για να ακούσετε τους ήχους.)
- 1.2. Υπολογίστε τον DFT των σημάτων $d_1[n]$ και $d_8[n]$ και δημιουργείτε τις γραφικές παραστάσεις των $|D_1[k]|$ και $|D_8[k]|$.
- 1.3. Δημιουργείτε και ηχογραφήστε σε αρχείο “tone_sequence.wav” με την **wavwrite()** του MATLAB ένα σήμα με διαδοχικούς τηλεφωνικούς τόνους ‘μεταφράζοντας’ το άθροισμα των αριθμών μητρώου των μελών της κάθε ομάδας σε τονικά σήματα. Κάθε ένα ψηφίο του αθροίσματος πρέπει να διαχωρίζεται από το προηγούμενο με 100 μηδενικά δείγματα. Για παράδειγμα αν οι αριθμοί μητρώου των δύο μελών είναι $AM_1 = 03092432$ και $AM_2 = 03093543$ τότε τα ζητούμενα ψηφία είναι: 0 6 1 8 5 9 7 5 (= 03092432 + 03093543). Αν η εργασία υλοποιείται από ένα άτομο τότε τα ζητούμενα ψηφία είναι ο αριθμός μητρώου.

- 1.4. Χρησιμοποιώντας την συνάρτηση **fft()** και κατάλληλα χρονικά παράθυρα, (i.) τετραγωνικά, και (ii.) Hamming, μήκους $N = 1000$, υπολογίστε τον Μετ/σμό Fourier των παραθυροποιημένων σημάτων, όπως προκύπτουν από το προηγούμενο σήμα. (Υπόδειξη: Δημιουργήστε τόσα παραθυροποιημένα σήματα όσοι και οι τόνοι που περιέχονται στο αρχικό σήμα).
- 1.5. Υπολογίστε μια λίστα από δείκτες k και τις αντίστοιχες συχνότητες που θεωρείτε ότι βρίσκονται εγγύτερα στις touch-tone συχνότητες.
- 1.6. Δημιουργήστε μια συνάρτηση με το όνομα **ttdecode**, που θα δέχεται σαν όρισμα εισόδου ένα τονικό σήμα (όπως περιγράφηκε στο **Ερωτ. 1.3**) και επιστρέφει ένα διάνυσμα με τα αντίστοιχα ψηφία. Για παράδειγμα αν το σήμα εισόδου **signIn** περιέχει τους τόνους για το νούμερο 210 – 3434120, η έξοδος της συνάρτησης **Vector** θα είναι $Vector = 2\ 1\ 0\ 3\ 4\ 3\ 4\ 1\ 2\ 0$. Επιβεβαιώστε την ορθή λειτουργία της ρουτίνας θέτοντας σαν είσοδο, το σήμα του **Ερωτ. 1.3**.

Υπόδειξη: Η συνάρτηση θα πρέπει πρώτα να υπολογίζει την ενέργεια κάθε ενός από τους τόνους του σήματος εισόδου με την χρήση της ρουτίνας **fft()**. Έπειτα να εντοπίζει ποιες είναι εκείνες οι συχνότητες που έχουν την μεγαλύτερη ενέργεια και να τις αντιστοιχίζει στις αρχικές συχνότητες του Πίνακα 1. Με βάση αυτές τις αντιστοιχίσεις, εντοπίστε ποιο είναι το αντίστοιχο ψηφίο. Υπενθυμίζεται ότι η ενέργεια E_k του σήματος γύρω από τη συχνότητα με δείκτη k ισούται με $E_k = |X[k]|^2$.

- 1.7. Με χρήση της εντολής **load()** του MATLAB φορτώστε το αρχείο “my_toughtones.mat” από το συμπληρωματικό υλικό της άσκησης “dsp14_lab1_Data.zip” που βρίσκεται στο mycourses. Στο work space θα πρέπει να υπάρχουν 2 ξεχωριστά σήματα, το *easySig* και *hardSig*. Προσδιορίστε τα ψηφία στα οποία αντιστοιχούν οι τόνοι των 2 σημάτων με την χρήση της ρουτίνας **ttdecode()**.

Σημειώνεται ότι σε κάθε περίπτωση οι απαντήσεις σας πρέπει να συνοδεύονται με τις σχετικές γραφικές παραστάσεις και σχόλια ώστε να είναι όσο το δυνατό τεκμηριωμένες.

Χρήσιμες Συναρτήσεις του MATLAB:

- **help:** Όλες οι συναρτήσεις του MATLAB έχουν αναλυτική επεξήγηση που μπορεί να εμφανιστεί με την εντολή `>> help <function>` στο MATLAB command prompt.
- **abs:** Υπόλογίζει το πλάτος ενός διανύσματος. Στην περίπτωση ενός μιγαδικού σήματος X , το πλάτος του δίνεται από την εντολή `abs(X)`
- **find:** Η MATLAB συνάρτηση **find()** επιστρέφει τους δείκτες ενός διανύσματος που ικανοποιούν μια δεδομένη συνθήκη. Για παράδειγμα η εντολή:

$$>> index = find(abs(x1) > 1)$$

επιστρέφει τους δείκτες του διανύσματος $x1$ που έχουν πλάτος μεγαλύτερο από την μονάδα.

Μέρος 2ο - Φασματική Ανάλυση και Ανίχνευση Ημιτονοειδών με τον Διακριτό Μετ/σμό Fourier (DFT)

Στην φασματική ανάλυση, η διακριτική ικανότητα ενός συστήματος ορίζεται ως η ικανότητα ανίχνευσης δύο διαφορετικών σημάτων όταν οι συχνότητές τους βρίσκονται αρκετά κοντά. Επιπλέον, ο (DFT) $X[k]$ N -σημείων ενός σήματος $x[n]$ προκύπτει από δειγματοληψία (στο πεδίο της Συχνότητας) του Μετ/σμού Fourier Διακριτού Χρόνου (DTFT) $X(e^{j\omega})$ στις συχνότητες $\omega = \frac{2\pi}{N}k$, όπου $0 \leq k \leq N-1$ και N το μήκος του DFT. Αναλυτικά:

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j\frac{2\pi}{N}nk}, \text{ όπου } 0 \leq k \leq N-1$$

Για την φασματική ανάλυση ενός σήματος εισόδου $y[n]$ συνήθως χρησιμοποιείται ένα παράθυρο $w[\cdot]$ ώστε $y[n] = w[n] \cdot x[n]$. Αυτό το παράθυρο καθορίζει αφενός το μέρος εκείνο του αρχικού σήματος $x[n]$ που θα χρησιμοποιηθεί κατά την ανάλυση και αφετέρου την διακριτική ικανότητα του Μετ/σμού Fourier.

Το παράθυρο Hamming $w_{\text{hamm}}[\cdot]$ ορίζεται ως:

$$w_{\text{hamm}}[n] = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N}\right), \quad 0 \leq n \leq N$$

Το παράθυρο αυτό δίνεται από την συνάρτηση `hamming` στο MATLAB.

Έστω ότι τα δύο αρχικά σήματα ανάλυσης $x_1[n]$ και $x_2[n]$ είναι τα εξής:

$$x_1[n] = A_1 e^{j(\omega_1 n + \phi_1)} \quad \text{και} \quad x_2[n] = A_2 e^{j(\omega_2 n + \phi_2)}$$

όπου $A_1 = 1$ και $A_2 = 0.8$ τα πλάτη των ημιτόνων, $n = 0, \dots, L-1$ και $\Delta\omega = |\omega_1 - \omega_2|$ η παράμετρος υπό εξέταση. Θεωρήστε τυχαίες τις φάσεις με τιμές $\phi_1, \phi_2 \in [0, 2\pi]$. Σαν σήμα ανάλυσης θεωρήστε το άθροισμα $y[n] = w[n] \cdot (x_1[n] + x_2[n])$.

- 2.1. Υλοποιείτε σε MATLAB την συνάρτηση **dft**, η οποία θα δέχεται ως είσοδο ένα διακριτό σήμα $x[n]$ και έναν ακέραιο N και θα υπολογίζει τον DFT μήκους N σημείων του σήματος. Στο MATLAB ο DFT υλοποιείται αποδοτικά από την ρουτίνα **fft**. Συγκρίνετε για επαλήθευση την έξοδο της δικής σας συνάρτησης **dft** με την έξοδο της ρουτίνας **fft**.
- 2.2. Αρχικά, θεωρήστε μήκος σήματος $L = 256$ δείγματα. Υπολογίστε τον DFT μήκους $N = 256$ δειγμάτων και σχεδιάστε το πλάτος του. Μεταβάλλοντας την συχνότητα ω_2 του δεύτερου σήματος x_2 προς την ω_1 , βρείτε πόσο μικρή μπορεί να γίνει η διαφορά $\Delta\omega$ ώστε να ξεχωρίζουν οι δύο κορυφές. Η συχνότητα ω_2 του δεύτερου ημιτόνου τροποποιείται προς την ω_1 με μικρά βήματα. Θεωρήστε αρχικές τιμές των συχνοτήτων $\omega_1 = \pi/9$ και $\omega_2 = \pi/5$. Τί παρατηρείτε;
- 2.3. Επαναλάβετε το παραπάνω πείραμα αλλά ο DFT θα έχει μήκος $N=512$ και 1024 δείγματα (μετά από την διαδικασία του zero-padding). Εξετάστε την ευκρίνεια φάσματος του σήματος και την δυνατότητα φασματικής διάκρισης των δύο ημιτόνων για τα διαφορετικά μήκη του DFT. Τί παρατηρείτε για τις διαφορετικές τιμές του N ; Εξηγήστε τα αποτελέσματά σας.
- 2.4. Για τις συχνότητες στις οποίες παρατηρήσατε οριακή δυνατότητα φασματικής διάκρισης, σχεδιάστε ξανά το πλάτος του DFT, θεωρώντας τώρα μήκος σήματος $L = 512, 1024$ δείγματα. Τί παρατηρείτε τώρα; Βρείτε το όριο φασματικής διάκρισης $\Delta\omega$ για τις νέες τιμές του L , όμοια με το ερώτημα 2.1. Εξηγήστε τα αποτελέσματά σας.

- 2.5. Θεωρήστε τώρα τα σήματα $x_1[n]$, $x_2[n]$ με συχνότητες $\omega_1 = 0.5\pi$ και $\omega_2 = 0.6\pi$, πλάτη $A_1 = 1$ και $A_2 = 0.1$ και τυχαίες φάσεις. Θεωρήστε ως σήμα ανάλυσης το $y[n] = w[n](x_1[n] + x_2[n])$. Για μήκος σήματος $L = 256$, σχεδιάστε το πλάτος του DFT μήκους $N = 1024$: (i) για τετραγωνικό παράθυρο $w[n] = 1$, $n = 1, \dots, L$ και (ii) για παράθυρο Hamming $w[n] = w_{\text{hamm}}[n]$. Τί παρατηρείτε; Εξηγείστε τα αποτελέσματά σας.

Μέρος 3ο - Χαρακτηριστικά Βραχέος Χρόνου Σημάτων Φωνής και Μουσικής (Ενέργεια και Ρυθμός Εναλλαγής Προσήμου)

Οι μετρήσεις βραχέος χρόνου είναι μετρήσεις που γίνονται σε ένα μετακινούμενο παράθυρο του σήματος. Πιο συγκεκριμένα, η ενέργεια βραχέος χρόνου ορίζεται ως:

$$E_n = \sum_{m=-\infty}^{\infty} [x(m)w(n-m)]^2 \quad (1)$$

όπου w ένα παράθυρο της επιλογής μας, το οποίο συνήθως είναι το Hamming παράθυρο (συνάρτηση `hamming` στο MATLAB). Αντίστοιχα, ο ρυθμός εναλλαγής προσήμου (Zero Crossing Rate) ορίζεται ως:

$$Z_n = \sum_{m=-\infty}^{\infty} |\text{sgn}[x(m)] - \text{sgn}[x(m-1)]|w(n-m) \quad (2)$$

- 3.1. Θεωρήστε το σήμα φωνής της πρότασης ‘Την ίδια τύχη είχαν και άλλες παρόμοιες σκέψεις’ που περιέχεται στο αρχείο “speech_utterance.wav” (συχνότητα δειγματοληψίας: 16 kHz) του συμπληρωματικού υλικού “dsp15_lab1_Data.zip” της άσκησης στο mycourses. Προαιρετικά, ηχογραφήστε στον υπολογιστή σας (π.χ. με χρήση του εργαλείου `praat`) την εκφώνηση μιας πρότασης και διαβάστε το σήμα στο MATLAB (με χρήση της συνάρτησης `wavread`). Ο στόχος είναι να μετρήσετε την ενέργεια βραχέος χρόνου και το ρυθμό εναλλαγής προσήμου. Χρησιμοποιήστε παράθυρο Hamming, μήκους 20-30 ms. Τί παρατηρείτε μεγάλωνοντας το μήκος του παραθύρου; Μπορείτε να χρησιμοποιήσετε αυτές τις μετρήσεις για να διαχωρίσετε φωνή από σιωπή ή έμφωνους (π.χ. /aa/, /ih/) από άφωνους ήχους (π.χ. /f/, /p/); Προαιρετικά, χρησιμοποιήστε το `praat` ως εργαλείο επισκόπησης του σήματος φωνής.

- 3.2. Επαναλάβετε την ίδια διαδικασία για το σήμα μουσικής “music.wav” που επίσης βρίσκεται στο συμπληρωματικό υλικό.

ΠΑΡΑΔΟΤΕΑ (α) Συνοπτική αναφορά που θα απαντάει στα δοθέντα ερωτήματα και θα περιλαμβάνει τις ζητούμενες γραφικές αναπαραστάσεις (β) Ηλεκτρονική παράδοση του κώδικα MATLAB καθώς και της ακολουθίας touch-tone υπο μορφή wav που φτιάξατε στο ερώτημα 1.3 του Μέρους 1.

Matlab Tutorial: Εισαγωγικές σημειώσεις για τη χρήση και τις βασικές εντολές του προγραμματιστικού περιβάλλοντος Matlab θα βρείτε στη διεύθυνση http://cvsp.cs.ntua.gr/courses/dsp/Material/MATLAB_Tutorial_ntua_2010.pdf