

[Εξηγήστε 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. Δημιουργήστε 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 $d_4[n]$ και $d_6[n]$ και δημιουργήστε τις γραφικές παραστάσεις των $|D_4[k]|$ και $|D_6[k]|$.
- 1.3. Δημιουργήστε 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_touchtones.mat” από το συμπληρωματικό υλικό της άσκησης “dsp19_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 τεχνηριωμένες.

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:** Υπολογίζει 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 Ημιτονοειδών και Ανίχνευση Απότομων Μεταβάσεων με τον Μετ/σμό Fourier Βραχέος Χρόνου (STFT) και τον Μετ/σμό Wavelets (διακριτοποιημένο DT-CWT)

Ο STFT είναι ο πιο διαδεδομένος μετ/σμός για την μελέτη του συχνοτικού περιεχομένου σημάτων σε μικρά χρονικά διαστήματα με υπολογισμό του Μετ/σμού Fourier Διακριτού Χρόνου (DTFT) σε καθένα από αυτά. Αναλυτικά:

$$\text{STFT}(\tau, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n - \tau]e^{-j\omega n}, \text{ όπου } \tau \in Z.$$

Στην MATLAB υλοποίηση του STFT χρησιμοποιείται ο DFT (και όχι ο DTFT), οπότε η συχνότητα ω δειγματοληπτείται στα $\omega_k = 2\pi k/N$, όπου N το μήκος του DFT μεγαλύτερο ή ίσο από το μήκος L του παραθύρου. Το μήκος L του παραθύρου $w[n]$ που θα χρησιμοποιηθεί καθορίζει την σχέση μεταξύ της διακριτικής ικανότητας στη συχνότητα και της ανάλυσης στο χρόνο. Μικρό παράθυρο πετυχαίνει καλή ανάλυση στο χρόνο με όμως χειρότερη διακριτική ικανότητα στην συχνότητα, ενώ αντίστροφα, μεγάλο παράθυρο στο χρόνο πετυχαίνει καλή διακριτική ικανότητα στο πεδίο συχνοτήτων χάνοντας σε ανάλυση στον χρόνο. Ένα από τα χαρακτηριστικά του STFT είναι ότι το μήκος L του παραθύρου w είναι σταθερό και επιλέγεται εξαρχής. Συνήθεις τιμές για εφαρμογές narrow-band επεξεργασίας φωνής είναι $L=30\text{-}50\text{msec}$ (για wide-band $\approx 5\text{ msec}$).

Ένας εναλλακτικός τρόπος αναπαράστασης των σημάτων είναι με τον Μετ/σμό των Wavelets. Στην περίπτωση σημάτων συνεχούς χρόνου ορίζεται ως:

$$WT(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t - \tau}{s}\right) dt,$$

όπου $\psi(t)$ είναι η βασική-μητρική συνάρτηση η οποία μπορεί να μετατοπιστεί κατά τ και να μεγεθυνθεί ή σμικρυνθεί κατά s . Στην περίπτωση διακριτών σημάτων, μπορεί να εφαρμοστεί η διακριτοποιημένη μορφή του DT-CWT (Discrete-time Continuous Wavelet Transform) (ρουτίνες `cwt`, `cwtft` στην MATLAB), στην οποία η μετατόπιση τ και η κλίμακα s παίρνουν τιμές πολλαπλάσιες της περιόδου δειγματοληψίας. Στην περίπτωση των Wavelets δεν υπάρχει ο περιορισμός του σταθερού μήκους παραθύρου όπως συνέβαινε στον STFT.

2.1. Έστω ότι δειγματοληπτούμε με $F_s = 1000\text{Hz}$ το σήμα $x(t)$ στο διάστημα $[0, 2]$ sec:

$$x(t) = 2\cos(2\pi 70t) + 3\sin(2\pi 140t) + 0.15v(t)$$

όπου $v(t)$ λευκός Gaussian θόρυβος μηδενικής μέσης τιμής (υλοποιείται με την συνάρτηση **randn** της MATLAB). Ισχύει ότι $x[n] = x(nT_s)$.

(α) Υπολογίστε το σήμα $x[n]$ και κάνετε την γραφική του παράσταση με χρήση της ρουτίνας **plot**.

(β) Υπολογίστε τον STFT με χρήση της ρουτίνας **spectrogram**. Χρησιμοποιήστε μήκος παραθύρου ίσο με 0.04sec και επικάλυψη ίση με 0.02sec . Αναπαραστήστε το πλάτος $|STFT(\tau, f)|$ με τις κατάλληλες τιμές στους άξονες του χρόνου και των συχνοτήτων, με χρήση της συνάρτησης **surf**.

(γ) Υπολογίστε τον DT-CWT με χρήση της ρουτίνας **cwtft**, επιλέγοντας το “Morlet” wavelet. Για τον υπολογισμό των κλιμάκων s χρησιμοποιήστε την συνάρτηση **wavescales**¹ (βρίσκεται στο συμπληρωματικό υλικό). Αναπαραστήστε το πλάτος $|WT(\tau, f)|$ με τις κατάλληλες τιμές στους άξονες του χρόνου και των συχνοτήτων, με χρήση της συνάρτησης **surf**.

(δ) Τί παρατηρείτε;

¹Η κλήση της συνάρτησης `[s, f] = wavescales('morl', F_s)` επιστρέφει το διάνυσμα κλιμάκων s και τις αντίστοιχες ψευδο-συχνότητες f για καλύτερη σύγκριση του WT με τον STFT.

2.2. Έστω ότι δειγματοληπτούμε με $F_s = 1000\text{Hz}$ το σήμα $x(t)$ στο διάστημα $[0, 2]$ sec:

$$x(t) = 1.7 \cos(2\pi 90t) + 0.15v(t) + 1.7\delta(t - 0.625) + 1.7\delta(t - 0.800))$$

όπου $v(t)$ λευκός θόρυβος μηδενικής μέσης τιμής όμοια με πριν. Ισχύει ότι $x[n] = x(nT_s)$.

Στις χρονικές στιγμές 625msec και 800msec το σήμα είχε απότομες μεταβολές, τις οποίες θέλουμε να εντοπίσουμε.

(α) Υπολογίστε το σήμα $x[n]$ και κάνετε την γραφική του παράσταση με χρήση της ρουτίνας **plot**.

(β) Υπολογίστε τον STFT με χρήση της ρουτίνας **spectrogram**. Χρησιμοποιείστε μήκος παραθύρου ίσο με 0.04sec και επικάλυψη ίση με 0.02sec. Αναπαραστήστε το πλάτος $|STFT(\tau, f)|$ με τις κατάλληλες τιμές στους άξονες του χρόνου και των συχνοτήτων, με χρήση της συνάρτησης **contour**.

(γ) Υπολογίστε τον DT-CWT με χρήση της ρουτίνας **cwtft**, επιλέγοντας το “Morlet” wavelet. Για τον υπολογισμό των κλιμάκων s χρησιμοποιείστε την συνάρτηση **wavescales**. Αναπαραστήστε το πλάτος $|WT(\tau, f)|$ με τις κατάλληλες τιμές στους άξονες του χρόνου και των συχνοτήτων, με χρήση της συνάρτησης **contour**.

(δ) Τί παρατηρείτε;

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 ‘Αλλά και για να επέμβει χρειάζεται συγκεκριμένη καταγγελία’ που περιέχεται στο αρχείο “speech_utterance.wav” (συχνότητα δειγματοληψίας: 16 kHz) του συμπληρωματικού υλικού “dsp19_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 **audioread**). 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.