

1^η Εργαστηριακή Άσκηση

<u>Ονοματεπώνυμο</u>	<u>ΑΜ</u>
Αντωνιάδης Παναγιώτης	03115009
Ποιητάρης Χριστόφορος	03115201

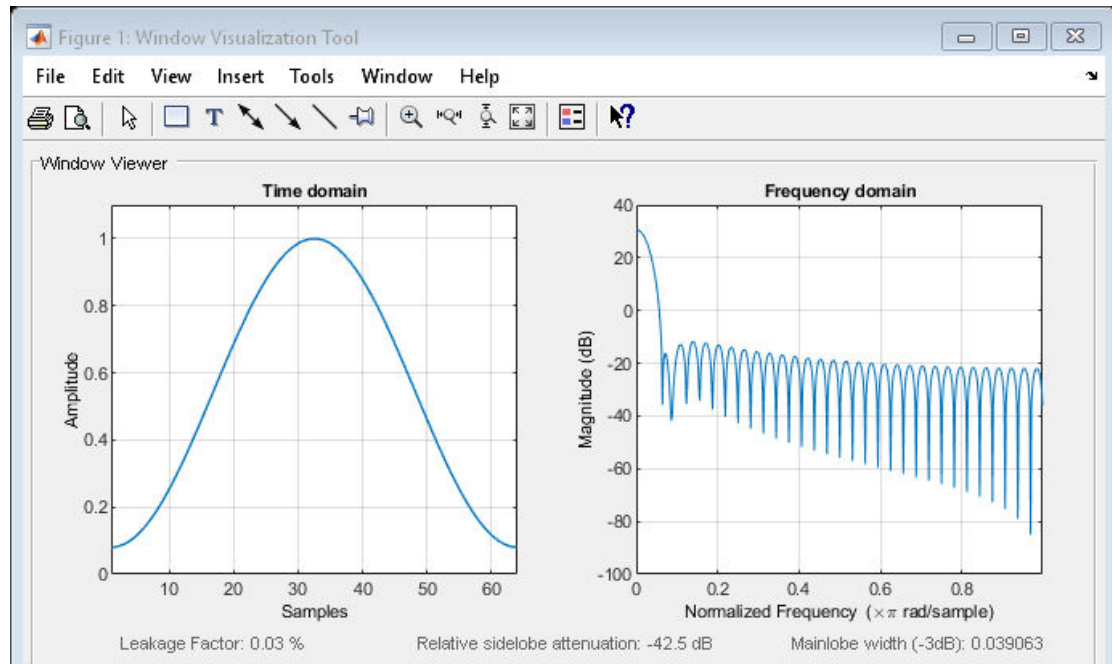
Θέμα: Εισαγωγή στην Ψηφιακή Επεξεργασία Σημάτων με MATLAB και Εφαρμογές σε ακουστικά σήματα

Μέρος 1: Χαρακτηριστικά Βραχέος χρόνου Σημάτων Φωνής και Μουσικής (Ενέργεια και Ρυθμός Εναλλαγής Προσήμου)

Σκοπός του πρώτου μέρους είναι μία εισαγωγή στα χαρακτηριστικά βραχέος χρόνου και συγκεκριμένα στην ενέργεια και τον ρυθμό εναλλαγής προσήμου. Στα σήματα φωνής, λοιπόν, και γενικά σε μη στατικά σήματα, ο υπολογισμός των διάφορων μεγεθών (ενέργεια, DFT κλπ) σε όλο το μήκος τους δεν μας διευκολύνει στη μελέτη και την επεξεργασία, καθώς γνωρίζουμε ότι συνεχώς η ένταση, η ενέργεια και η συχνότητά τους αλλάζουν. Για τον λόγο αυτό, χωρίζουμε το σήμα μας σε παράθυρα (επικαλυπτόμενα ή μη) μέσα στα οποία θεωρούνται και μελετάμε τα χαρακτηριστικά τους ανά παράθυρο. Στην περίπτωση μας, ένα σήμα φωνής μπορεί να θεωρηθεί ως στατικό όταν το κάθε παράθυρο είναι της τάξης του 10-30ms.

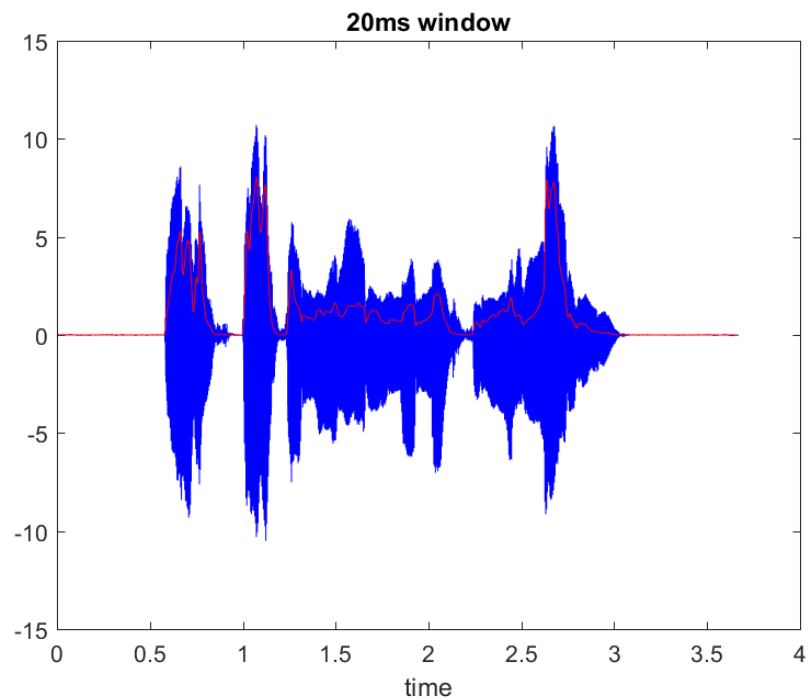
1.1. Αρχικά, θα χρησιμοποιήσουμε το σήμα φωνής που μας δίνεται και βρίσκεται στο αρχείο **speech_utterance.wav**. Το παράθυρο που θα χρησιμοποιηθεί είναι το Hamming παράθυρο. Θα υπολογίσουμε τα χαρακτηριστικά βραχέος χρόνου (ενέργεια και ρυθμό εναλλαγής προσήμου) για διάφορες τιμές μήκους παραθύρου με σταθερό βήμα ολίσθησης ίσο με 1 σημείο. Η επιλογή, τώρα, του συγκεκριμένου παραθύρου έγκειται στο γεγονός ότι μειώνει την παραμόρφωση που δημιουργείται στα άκρα του παραθυροποιημένου σήματος και εξομαλύνει το αποτέλεσμα. Αυτό οφείλεται στο γεγονός ότι οι τιμές στα άκρα του είναι κοντά στο μηδέν και όσο πλησιάζουν στο κέντρο του αυξάνονται. Παρακάτω, βλέπουμε την μορφή στο χρόνο και στην συχνότητα ενός hamming παραθύρου (Εικόνα 1). Ακριβώς από κάτω παρουσιάζονται οι γραφικές παραστάσεις που προέκυψαν, πρώτα για την ενέργεια και στη συνέχεια για το ρυθμό εναλλαγής προσήμου. Ο κώδικας MATLAB που χρησιμοποιήθηκε για την δημιουργία των γραφικών παραστάσεων βρίσκεται στον φάκελο source και η λειτουργία του εξηγείται με αναλυτικά σχόλια σε κάθε αρχείο .m.

Digital Signal Processing



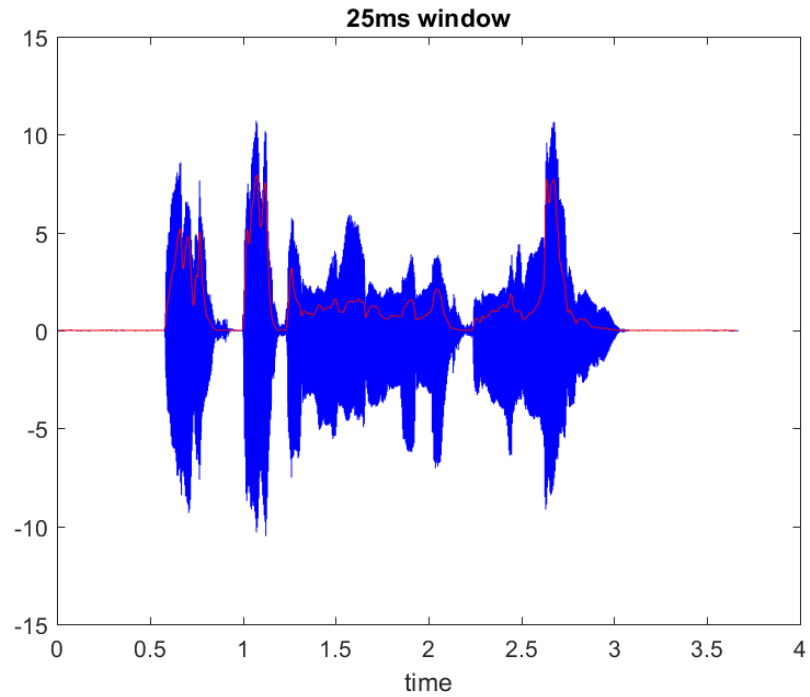
Εικόνα 1: Hamming window of 65 points in the time domain (left) and in the frequency domain (right).

Γραφικές παραστάσεις για το short time energy:

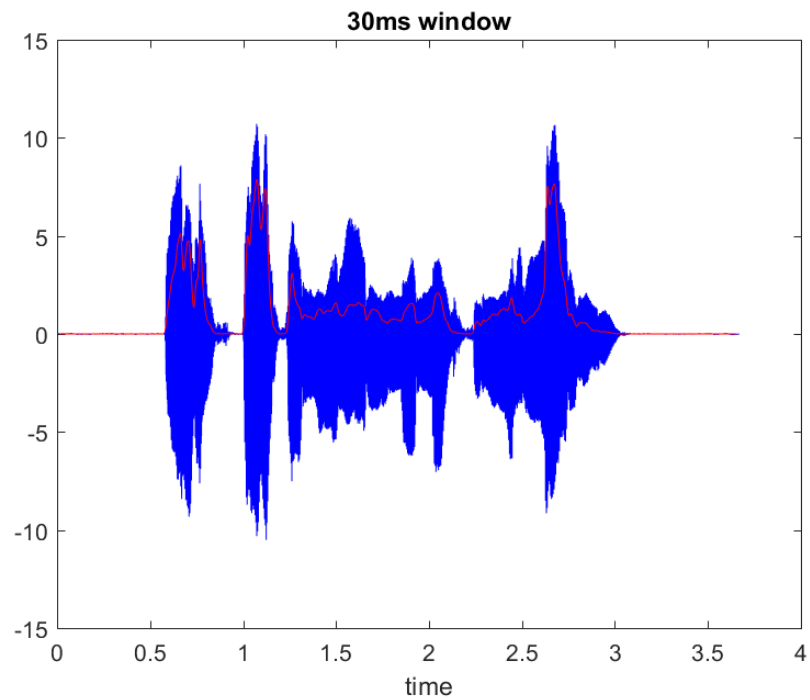


Εικόνα 2: Audio signal (blue) and short time energy with window of length 20ms (red).

Digital Signal Processing

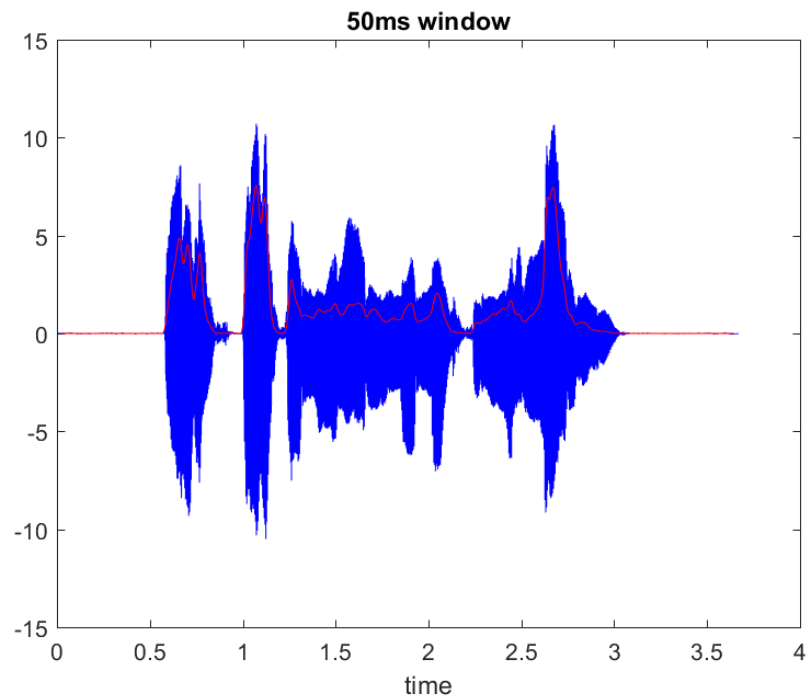


Εικόνα 3: Audio signal (blue) and short time energy with window of length 25ms (red).

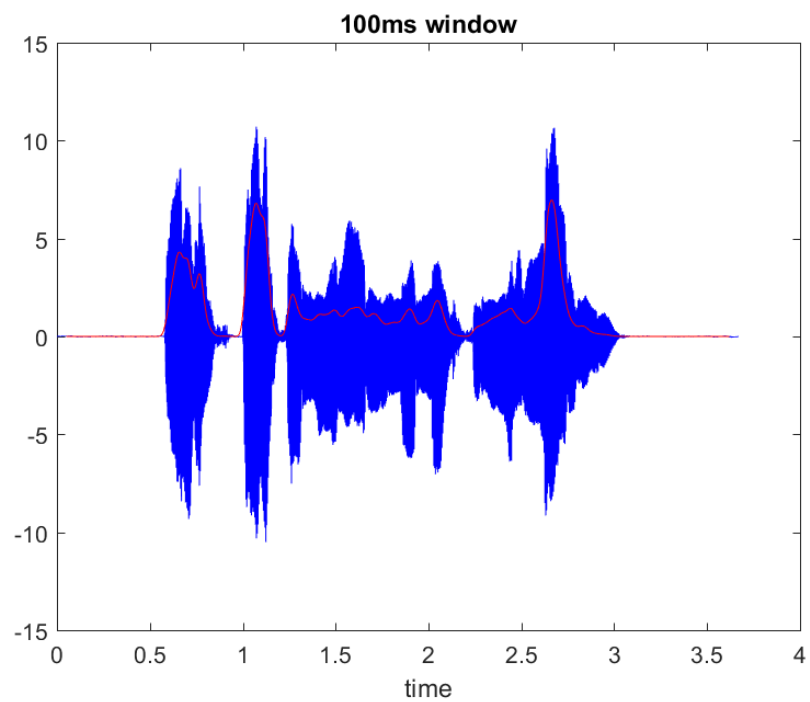


Εικόνα 4: Audio signal (blue) and short time energy with window of length 30ms (red).

Digital Signal Processing

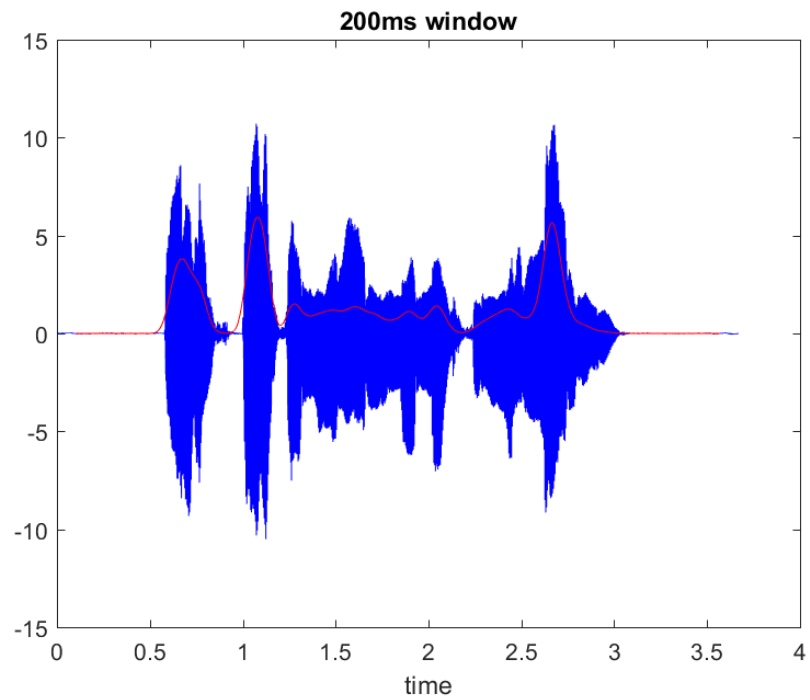


Εικόνα 5: Audio signal (blue) and short time energy with window of length 50ms (red).

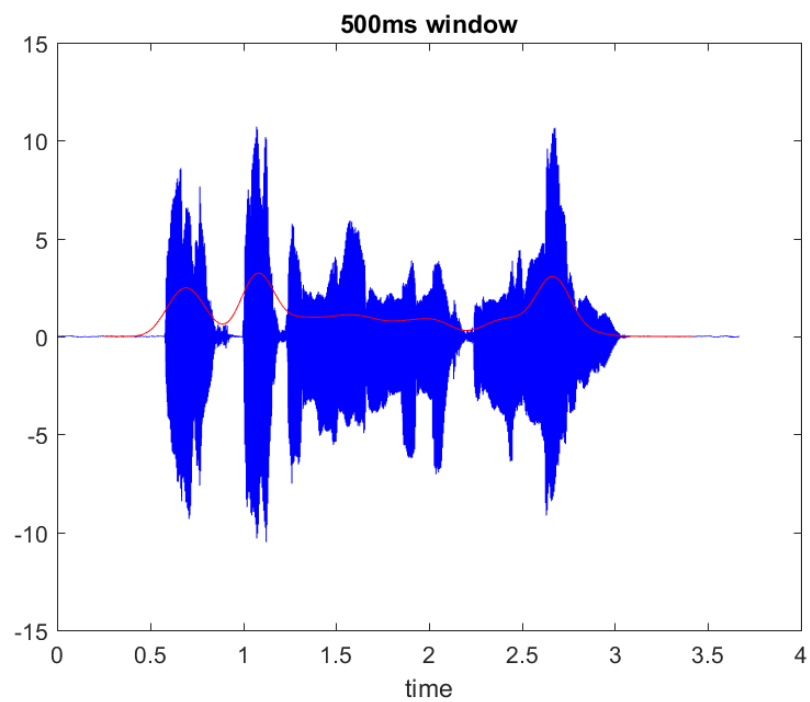


Εικόνα 6: Audio signal (blue) and short time energy with window of length 100ms (red).

Digital Signal Processing

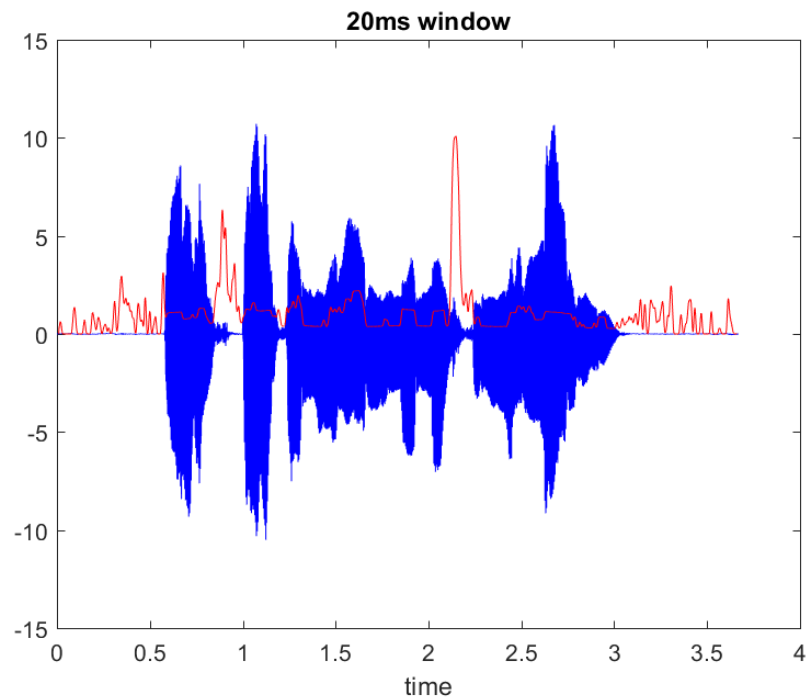


Εικόνα 7: Audio signal (blue) and short time energy with window of length 200ms (red).

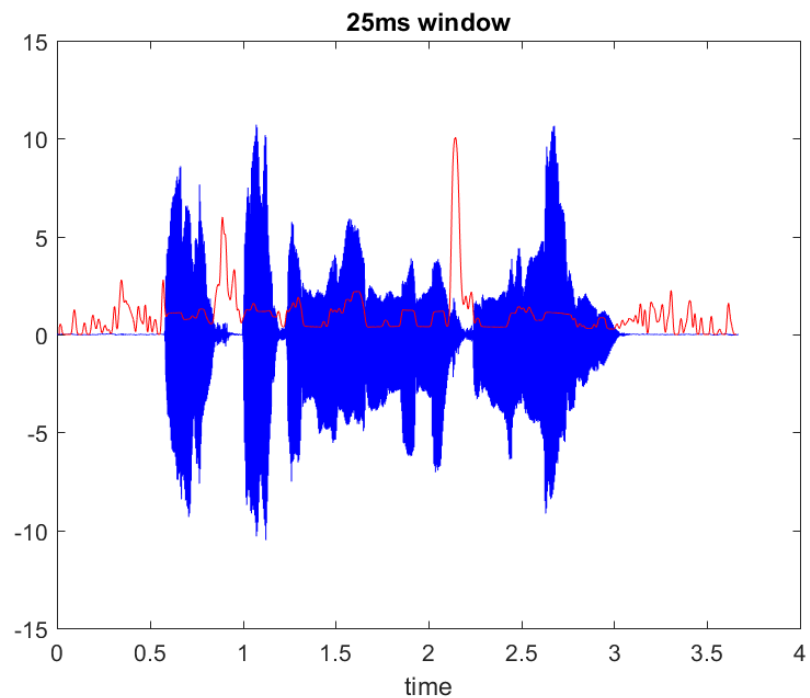


Εικόνα 8: Audio signal (blue) and short time energy with window of length 500ms (red).

Γραφικές παραστάσεις για το zero crossing rate

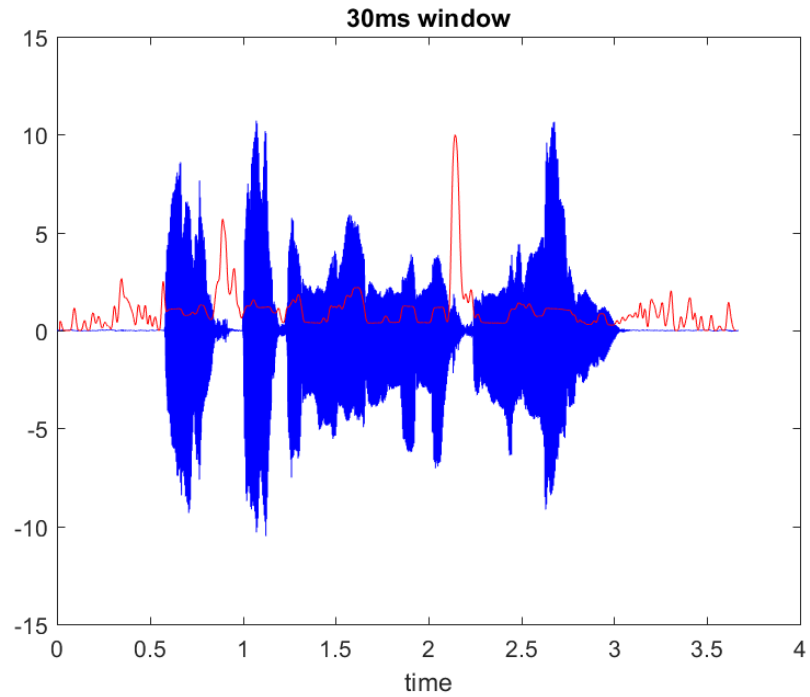


Εικόνα 9: Audio signal (blue) and zero cross rate with window of length 20ms (red).

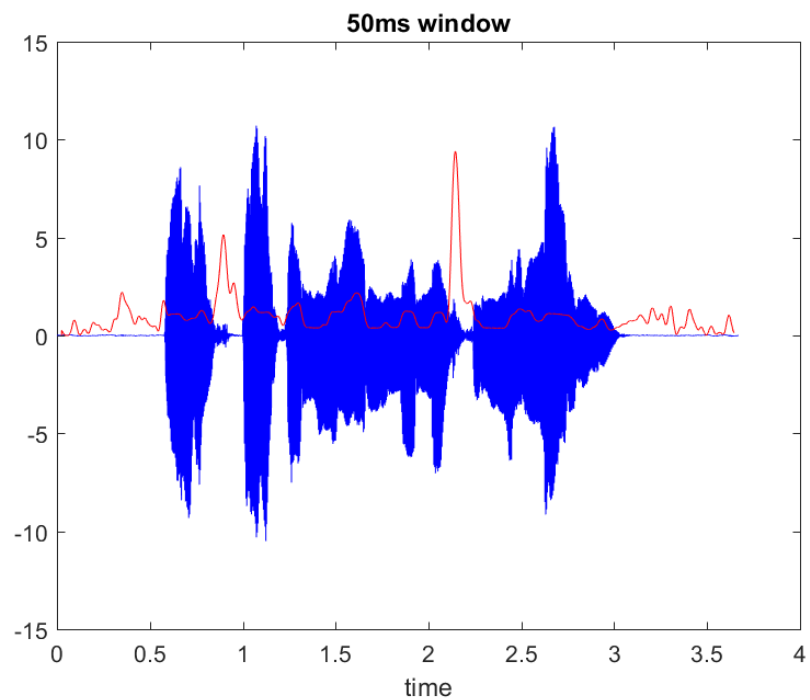


Εικόνα 10: Audio signal (blue) and zero cross rate with window of length 25ms (red).

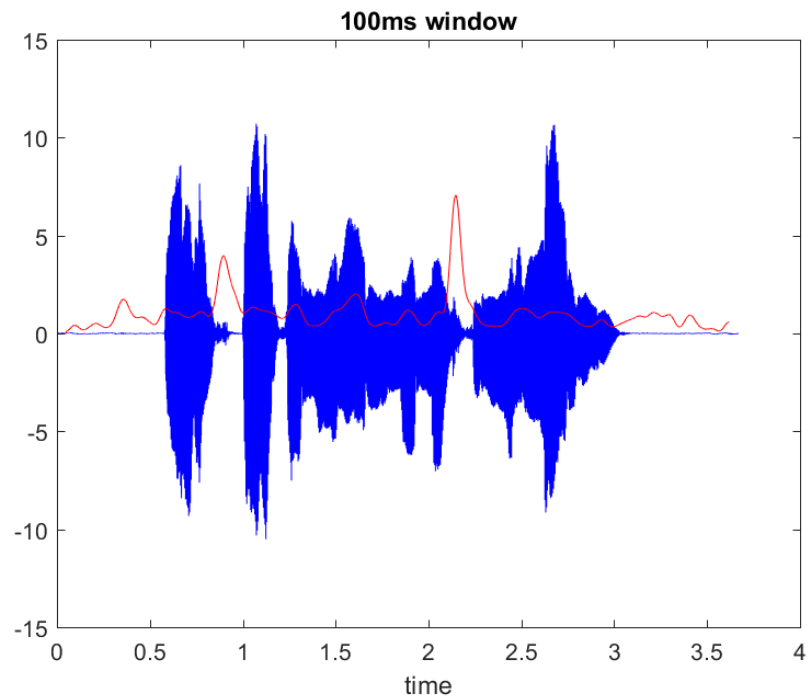
Digital Signal Processing



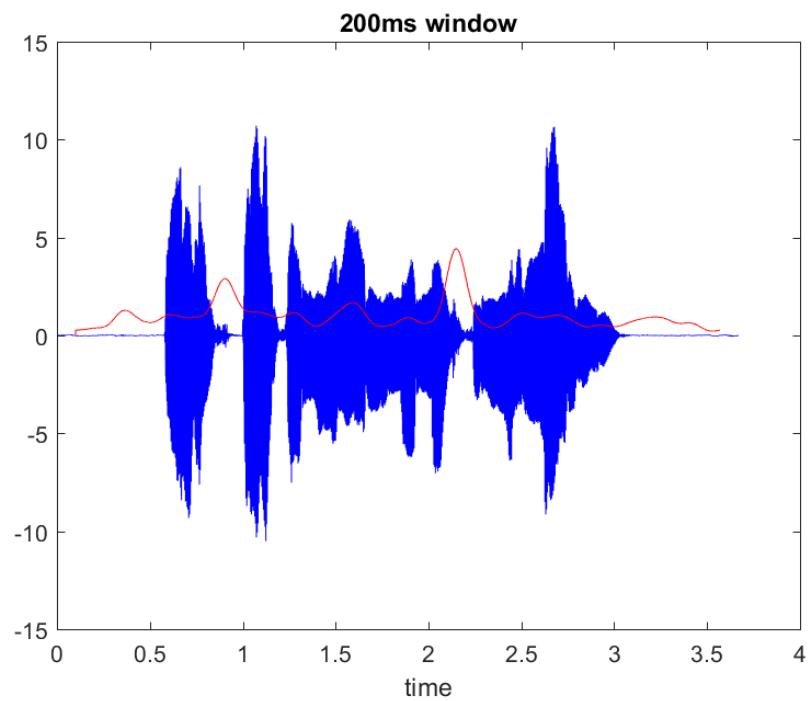
Εικόνα 11: Audio signal (blue) and zero cross rate with window of length 30ms (red).



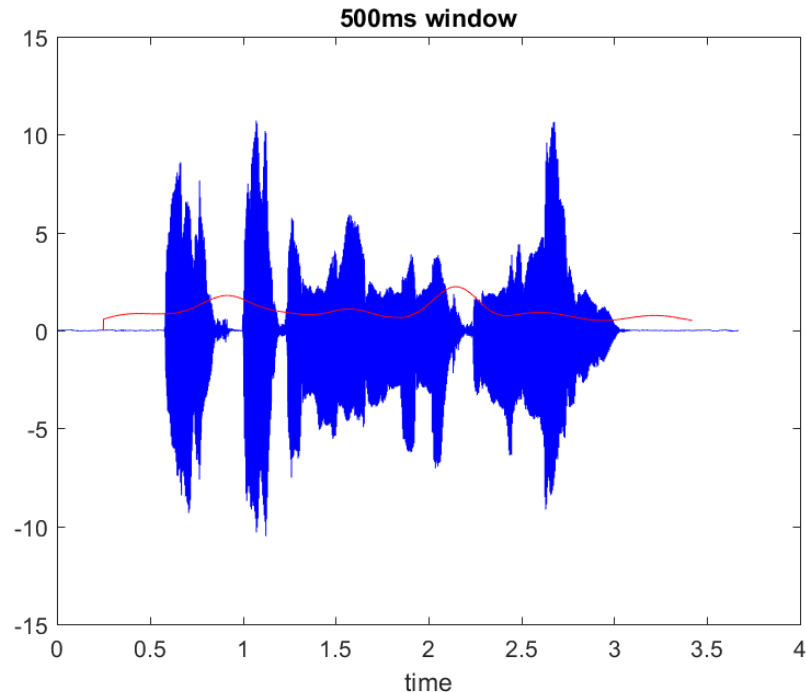
Εικόνα 12: Audio signal (blue) and zero cross rate with window of length 50ms (red).



Εικόνα 13: Audio signal (blue) and zero cross rate with window of length 100ms (red).



Εικόνα 14: Audio signal (blue) and zero cross rate with window of length 200ms (red).



Εικόνα 15: Audio signal (blue) and zero cross rate with window of length 500ms (red).

Παρατηρήσεις:

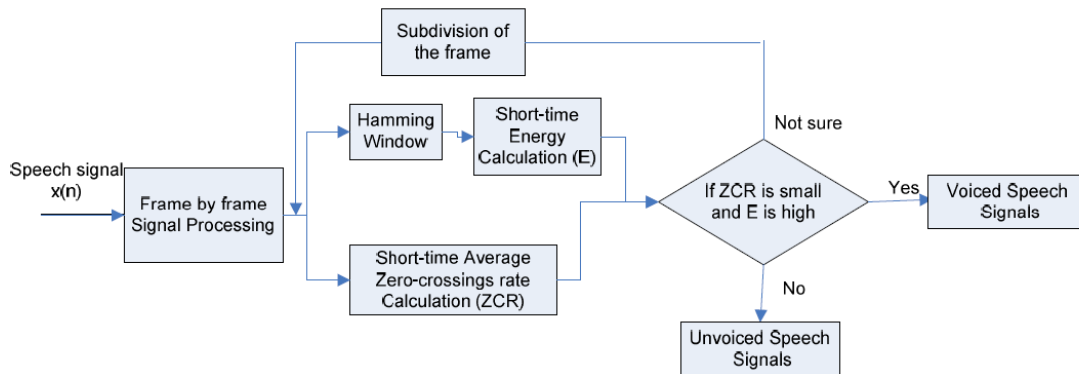
- ✓ Για την καλύτερη κατανόηση και ερμηνεία των γραφικών παραστάσεων έχει γίνει normalization των δεδομένων, το σήμα μας και τα μεγέθη βραχέος χρόνου, δηλαδή, έχουν διαιρεθεί με την μέση τιμή τους. Έτσι, μπορούμε να τα συγκρίνουμε καλύτερα.
- ✓ Τα μεγέθη μας είναι, προφανώς, όλα διακριτά. Ωστόσο, αντί της **scatter** χρησιμοποιήθηκε η **plot** η οποία ενώνει τα διακριτά σημεία και γι' αυτό τα μεγέθη φαίνονται σαν συνεχή.
- ✓ Ο ρυθμός εναλλαγής προσήμου είναι ουσιαστικά ένας εύκολος και γρήγορος τρόπος να μετρήσουμε την συχνότητα του σήματος σε κάθε παράθυρο.

Συμπεράσματα:

- ✓ Καθώς, λοιπόν, αυξάνεται το μήκος του παραθύρου, η ενέργεια βραχέος χρόνου “ακολουθεί” όλο και λιγότερο το αρχικό μας σήμα, με αποτέλεσμα από τα 50-100ms και μετά να μην μας δείχνει μια ρεαλιστική απεικόνιση του σήματος. Αυτό σημαίνει ότι τέτοια παράθυρα θα μας οδηγήσουν εν τέλει σε λάθος συμπεράσματα για το αρχικό μας σήμα. Παράλληλα, η ενέργεια εξομαλύνεται και γίνεται πιο λεία χωρίς πολλές κυματώσεις. Αντίθετα, σε μικρότερα μήκη παραθύρου, η ενέργεια συμβαδίζει με την συμπεριφορά του σήματος έχοντας έτσι και πολλές κυματώσεις.
- ✓ Όσων αφορά τον ρυθμό εναλλαγής προσήμου, ισχύουν αντίστοιχα πράγματα. Σε μεγάλα, δηλαδή, παράθυρα δεν συμβαδίζει με το σήμα μας και είναι σχετικά ομαλός και σε μικρότερα, μας δίνει μία ακριβής εικόνα για το τι συμβαίνει στο σήμα μας.

Digital Signal Processing

- ✓ Συνολικά, η αύξηση του μεγέθους του παραθύρου άνω των 30ms μας δημιουργεί προβλήματα διότι η ανθρώπινη ομιλία έχει εναλλαγές ανά 10 με 20 ms.
- ✓ **Διαχωρισμός φωνής από σιωπή και έμφωνων από άφωνους ήχους:**



Στο παραπάνω διάγραμμα φαίνεται ένας από τους πολλούς τρόπους που υπάρχουν για να διαχωρίσεις τα διαστήματα έμφωνων και άφωνων ήχων χρησιμοποιώντας τα δύο παραπάνω μεγέθη. Αντίστοιχα μπορούμε να διαχωρίσουμε την φωνή από την σιωπή. Ουσιαστικά, όλες οι μέθοδοι βασίζονται στις τρεις παρακάτω βασικές παρατηρήσεις:

1. Τα κομμάτια που περιέχουν έμφωνο ήχο έχουν υψηλή ενέργεια εξαιτίας της περιοδικότητάς του και χαμηλό ρυθμό εναλλαγής προσήμου.
2. Τα κομμάτια που περιέχουν άφωνο ήχο έχουν σημαντικά χαμηλότερη ενέργεια λόγω της μη περιοδικότητες. Αντίθετα, έχουν υψηλό ρυθμό εναλλαγής προσήμου.
3. Τα κομμάτια που περιέχουν σιωπή έχουν ελάχιστη ενέργεια βραχέος χρόνου και πάρα πολύ μεγάλο ρυθμό εναλλαγής προσήμου.

Έτσι, θέτοντας τιμές κατωφλιού για την ενέργεια και τον ρυθμό εναλλαγής προσήμου του κάθε παραθύρου, μπορούμε να διαχωρίσουμε έμφωνους από άφωνους ήχους και από την σιωπή.

Χρήση του praat για μελέτη των παραπάνω:

Με την χρήση του praat απομονώσαμε τρία κομμάτια του σήματος, ένα με έμφωνο ήχο (/aa/) ένα με άφωνο (/f/) και ένα με σιωπή. Στον παρακάτω πίνακα συνοψίζονται τα αποτελέσματά μας:

Sound	Letter	Start Time	End Time	Short Time Energy	Zero Crossing Rate
voiced	/aa/	1.62	1.64	0.7015	47.026
unvoiced	/f/	0.88	0.9	0.0015	123.04
silence	-	0.94	0.94	0.00073	75.76

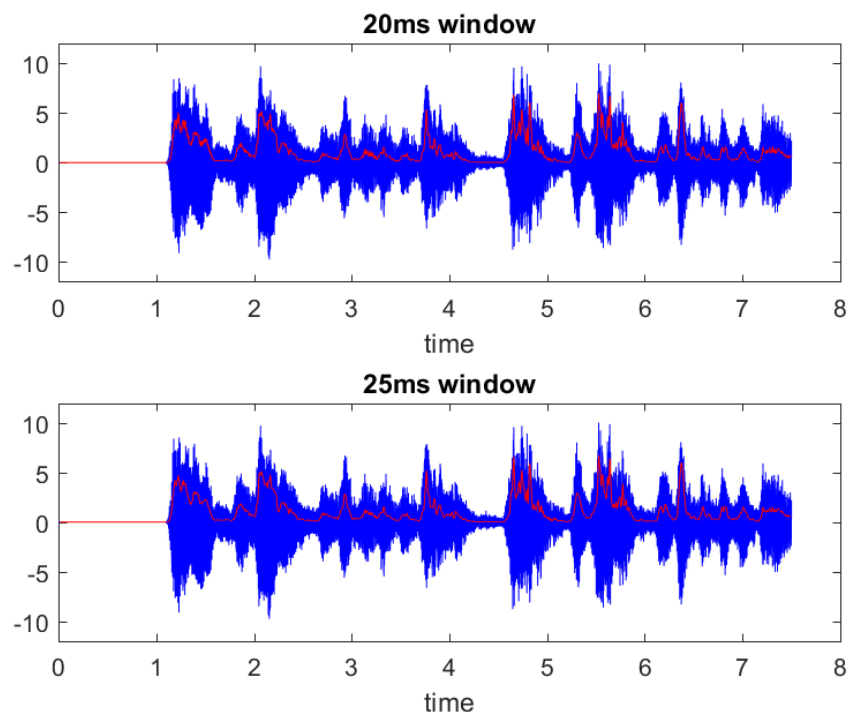
Digital Signal Processing

Ο αλγόριθμος που χρησιμοποιήθηκε για να προκύψουν όλα αυτά φαίνεται σε μορφή ψευδοκώδικα παρακάτω:

```
[start, stop] = read_time; %get the start and stop time in seconds  
start_point = start/Ts;  
stop_point = stop/Ts;  
center_point = start_point + (stop_point - start_point)/2; %get the center sample  
energy = Energy(center_point); %Short Time Energy of this window  
zcr = Zeros(center_point); %Zero Crossing Rate of this window
```

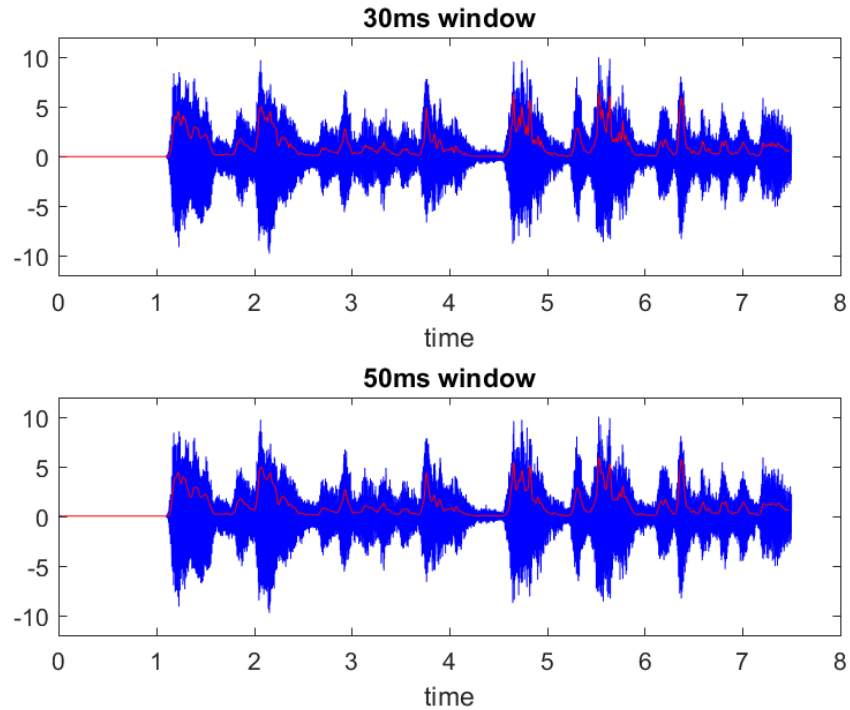
- 1.2.** Σε αυτό το ερώτημα επαναλαμβάνεται η ίδια διαδικασία με το σήμα μουσικής “music.wav”. Ακολουθούν οι αντίστοιχες γραφικές παραστάσεις ομαδοποιημένες αυτήν την φορά γιατί τα συμπεράσματα από αυτές είναι παρόμοια. Η μόνη διαφορά είναι ότι στο zero crossing rate δεν έχει γίνει normalization γιατί οι τιμές που προέκυπταν ήταν πολύ μικρές.

Γραφικές παραστάσεις για το short time energy

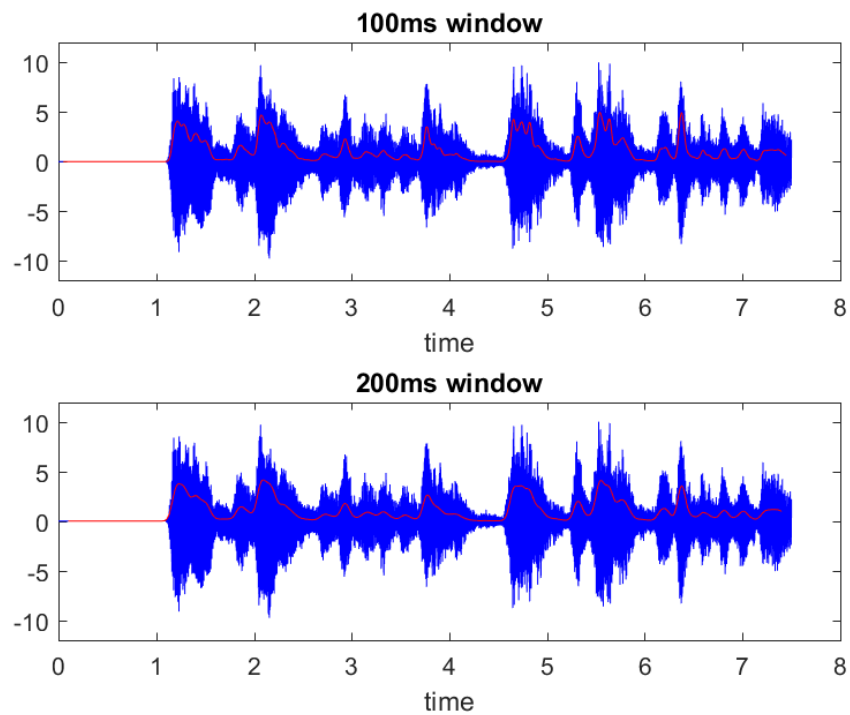


Εικόνα 16: Audio signal (blue) and short time energy with window length 20ms (up) and 25ms (down)

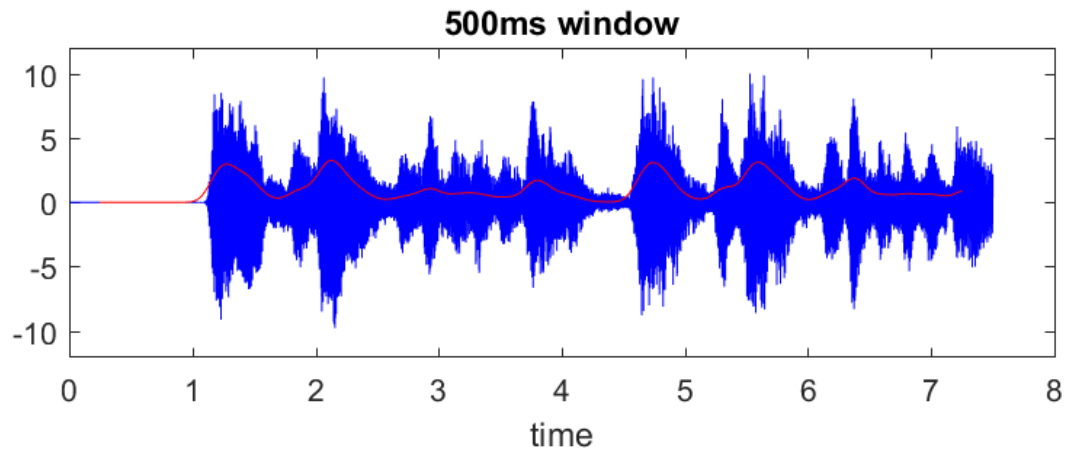
Digital Signal Processing



Εικόνα 17: Audio signal (blue) and short time energy with window length 30ms (up) and 50ms (down)

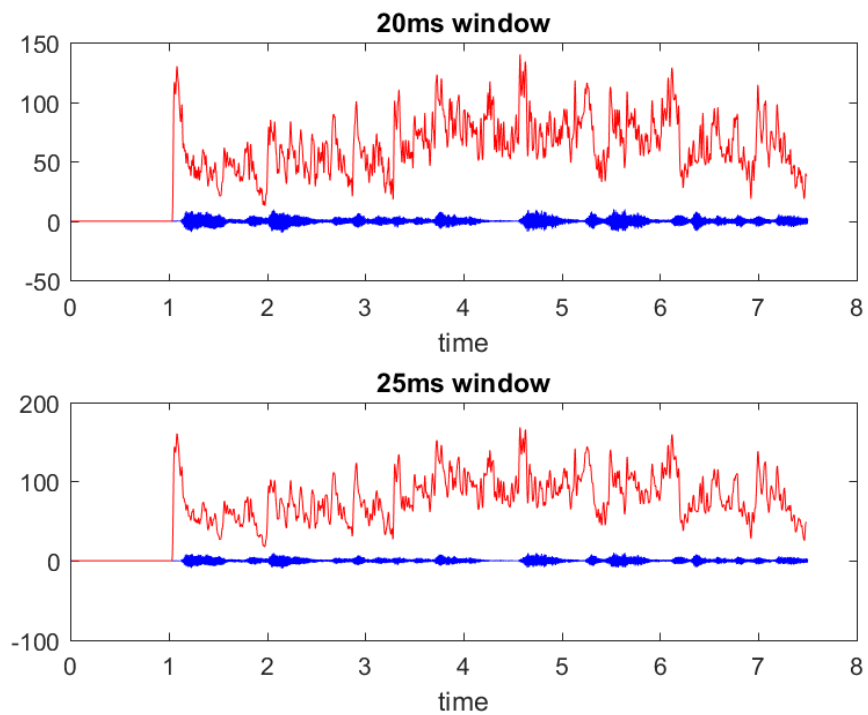


Εικόνα 18: Audio signal (blue) and short time energy with window length 100ms (up) and 200ms (down)



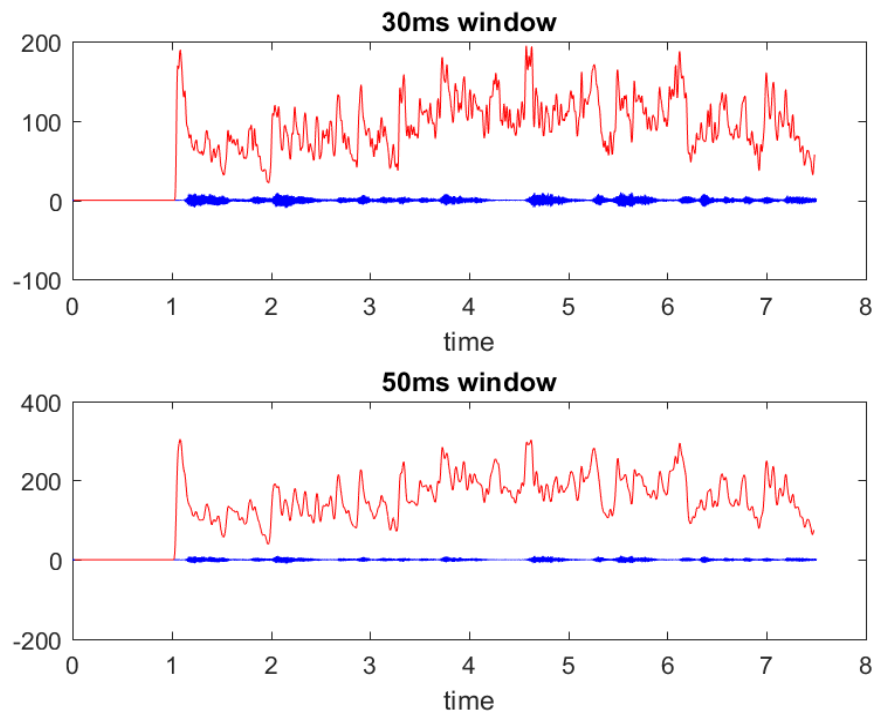
Εικόνα 19: Audio signal (blue) and short time energy with window length 500ms.

Γραφικές παραστάσεις για το zero crossing rate

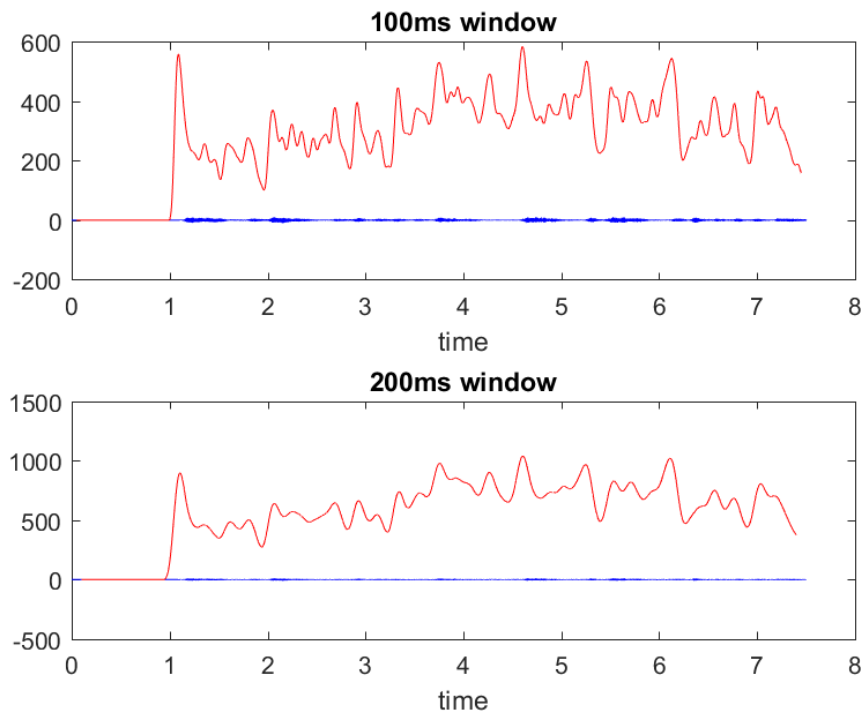


Εικόνα 20: Audio signal (blue) and zero cross rate with window length 20ms (up) and 25ms (down)

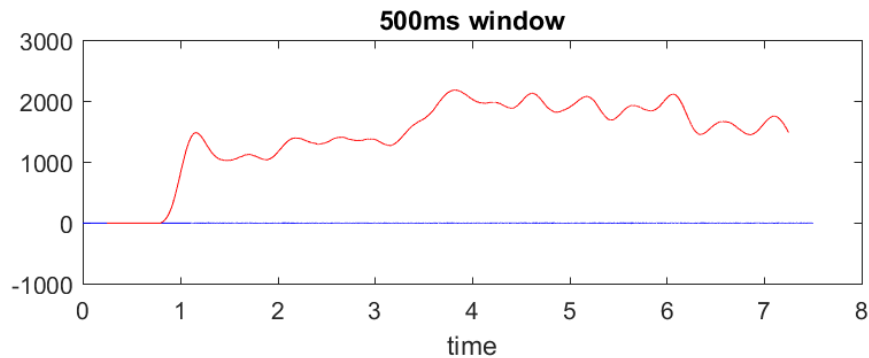
Digital Signal Processing



Εικόνα 21: Audio signal (blue) and zero cross rate with window length 30ms (up) and 50ms (down)



Εικόνα 22. Audio signal (blue) and zero cross rate with window length 100ms (up) and 200ms (down)



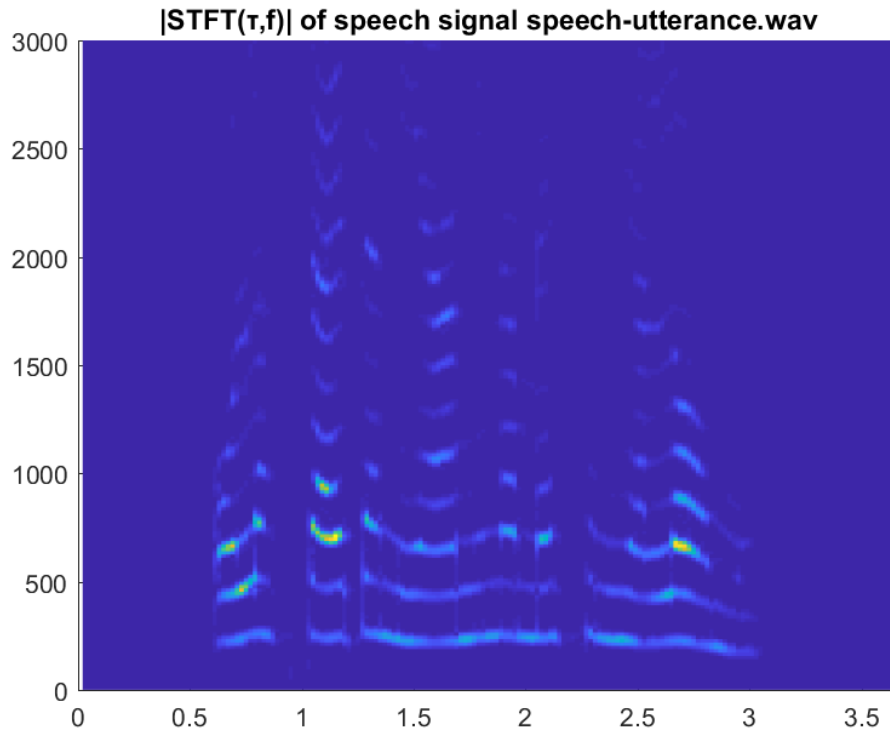
Εικόνα 23: Audio signal (blue) and zero cross rate with window length 500ms.

✚ Μέρος 2^ο: Ανάλυση και Σύνθεση Σήματος με τον Μετ/σμο Fourier Βραχέος Χρόνου (STFT).

Σκοπός του δεύτερου μέρους είναι η υλοποίηση του Short Time Fourier Transform (STFT) για την μελέτη του συχνοτικού περιεχομένου ενός φωνητικού σήματος με το πέρασμα του χρόνου. Επίσης, θα υλοποιηθεί και ο αντίστροφος STFT μέσω του οποίου μπορεί κανείς (υπό κατάλληλες συνθήκες) να ανακατασκευάσει πλήρως το αρχικό σήμα.

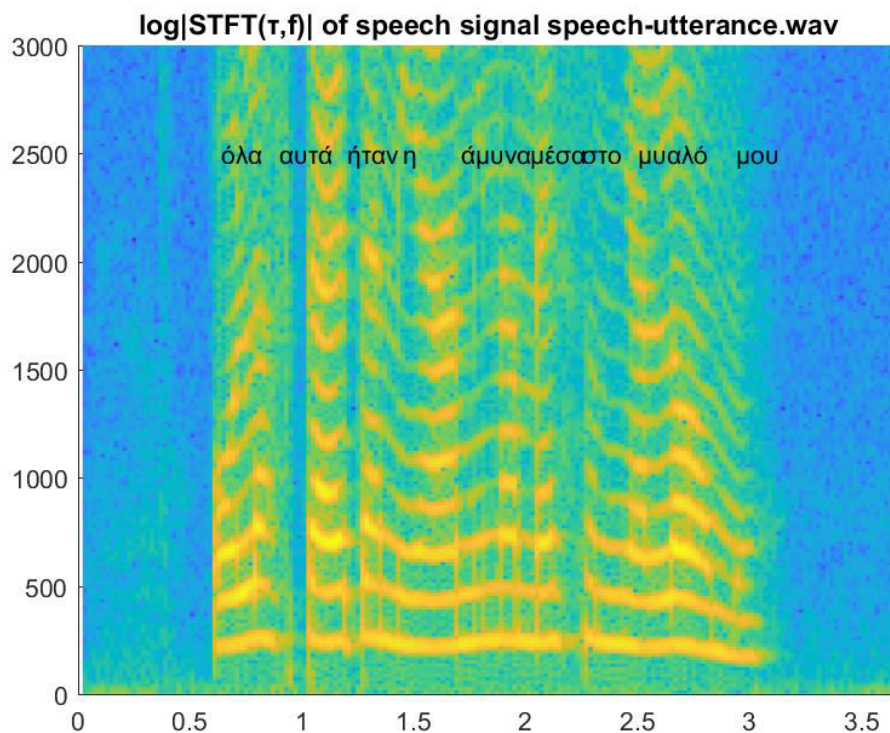
2.1. Υλοποίηση στο MATLAB.

2.2. Το πλάτος $|\text{STFT}(\tau, f)|$ συναρτήσει του χρόνου τ και της συχνότητας f φαίνεται παρακάτω:



Εικόνα 24

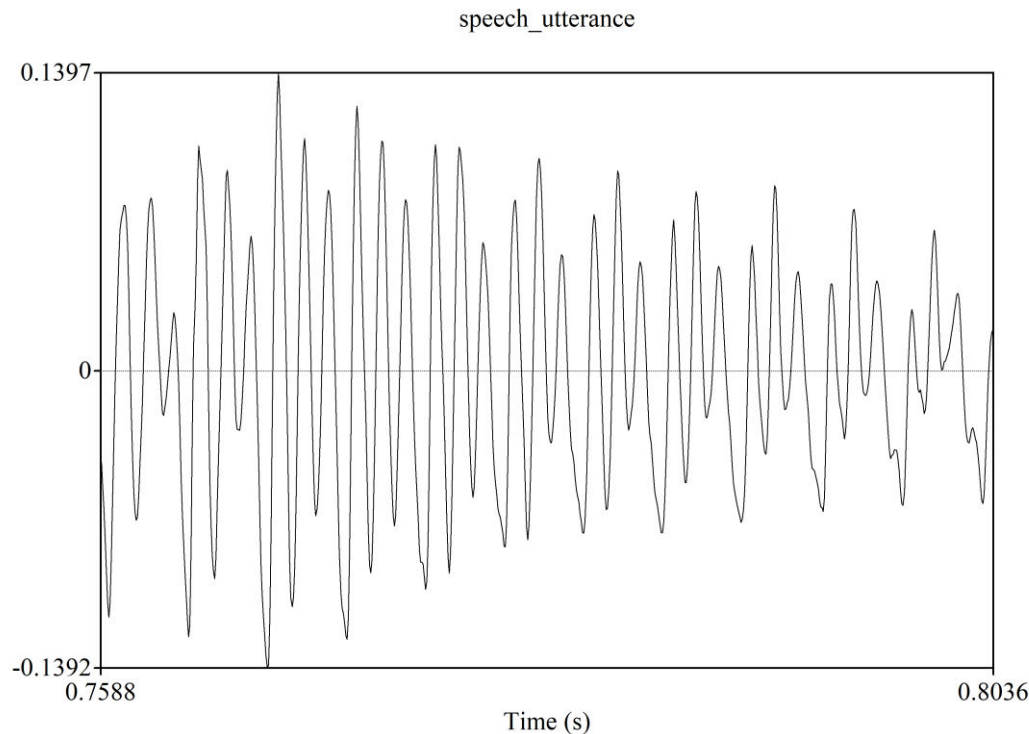
Επειδή λοιπόν οι εναλλαγές στο πλάτος δεν φαίνονται καλά, βολεύει περισσότερο η αναπαράσταση του $\log|STFT(\tau, f)|$ όπου φαίνεται παρακάτω.



Εικόνα 25.

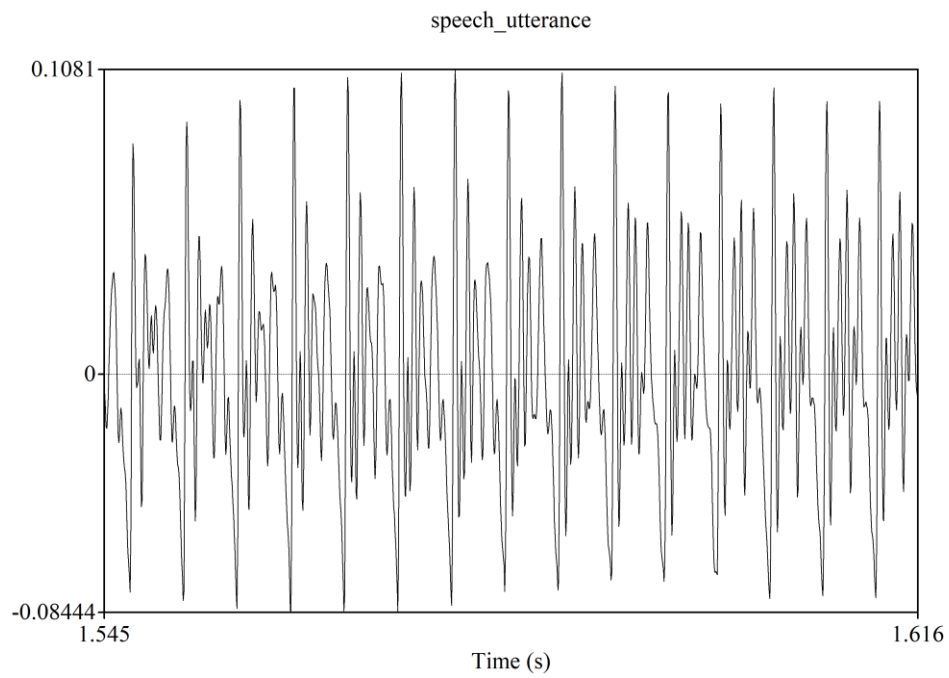
Στο τελευταίο διάγραμμα έχουν αντιστοιχιστεί στο πάνω μέρος και οι λέξεις που ακούγονται κάθε χρονική στιγμή. Στη συνέχεια θα απομονώσουμε δύο χρονικά τμήματα που να αντιστοιχούν στο φωνήεν /α/ και δύο στο φωνήεν /ο/. Αυτό που θέλουμε να διαπιστώσουμε είναι ότι στο spectrogram ενός σήματος φωνής τα φωνήεντα είναι αυτά που έχουν μεγάλη ένταση γιατί έχουν περιοδικότητα σε συγκεκριμένες συχνότητες. Κάθε φωνήεν, μάλιστα, έχει και διαφορετική συχνότητα με αποτέλεσμα να μπορούμε να τα διακρίνουμε κάθε φορά. Για την απομόνωση των χρονικών τμημάτων χρησιμοποιήθηκε αρχικά το εργαλείο praat για να υπολογιστούν πειραματικά ποια είναι τα τμήματα που θέλουμε. Στη συνέχεια μέσω του MATLAB δημιουργήσαμε τα spectrograms μόνο για τα συγκεκριμένα χρονικά τμήματα. Όλα αυτά παρουσιάζονται στο παρακάτω πίνακα και στις αντίστοιχες εικόνες.

Word	Letter	Start time	End time
όλα	/α/	0.76	0.8
άμυνα	/α/	1.5	1.6
όλα	/ο/	0.59	0.66
μυαλό	/ο/	2.63	2.71

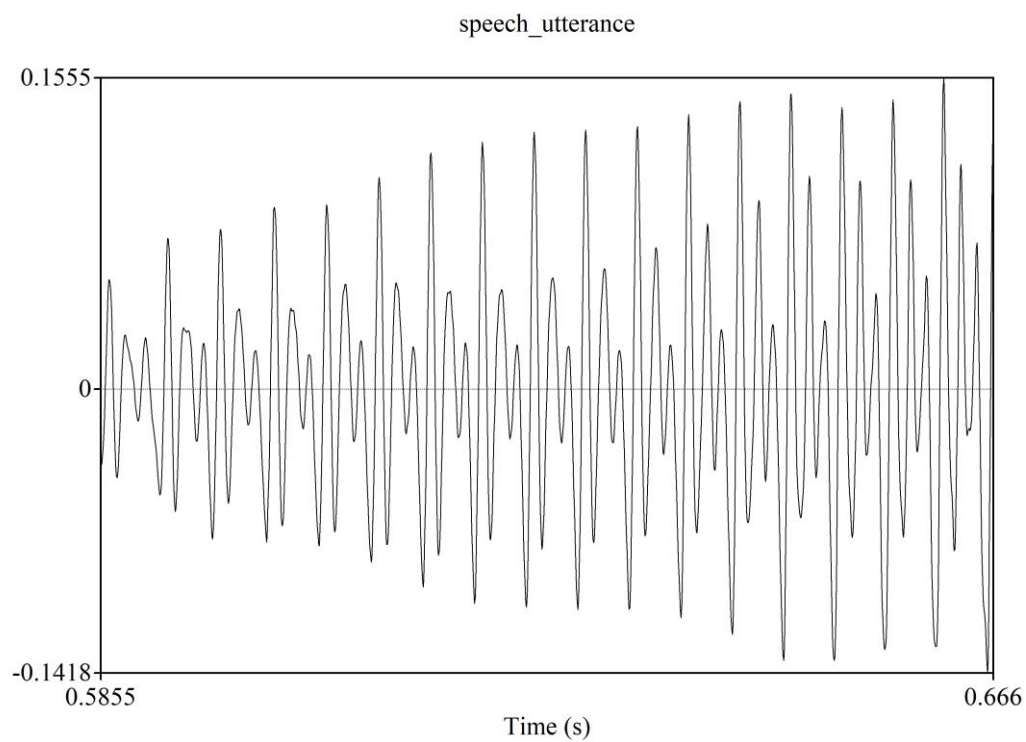


Εικόνα 26: Η κυματομορφή του /α/ στη λέξη 'όλα'.

Digital Signal Processing

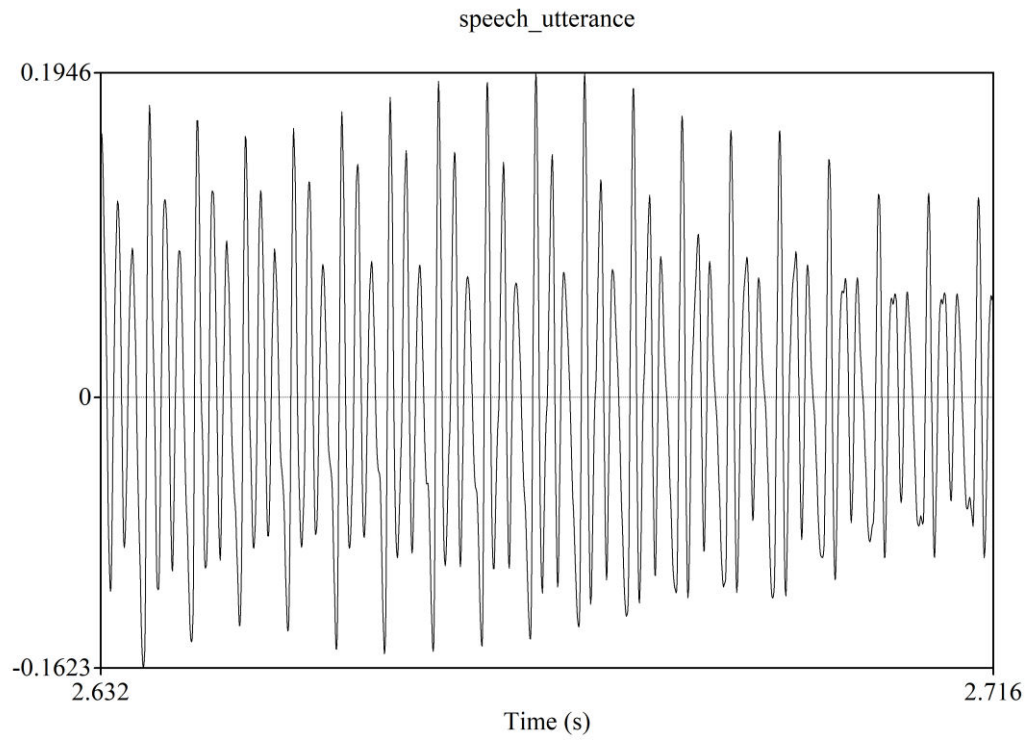


Εικόνα 27: Η κυματομορφή του /α/ στη λέξη 'άμυνα'.

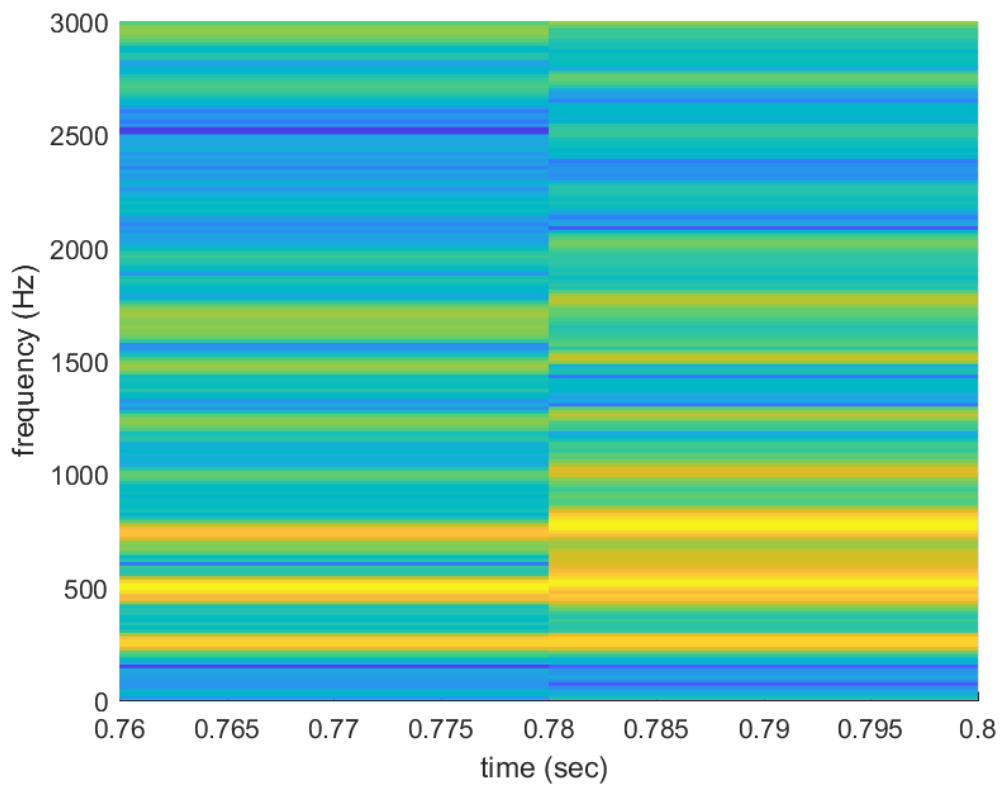


Εικόνα 28: Η κυματομορφή του /ο/ στη λέξη 'όλα'.

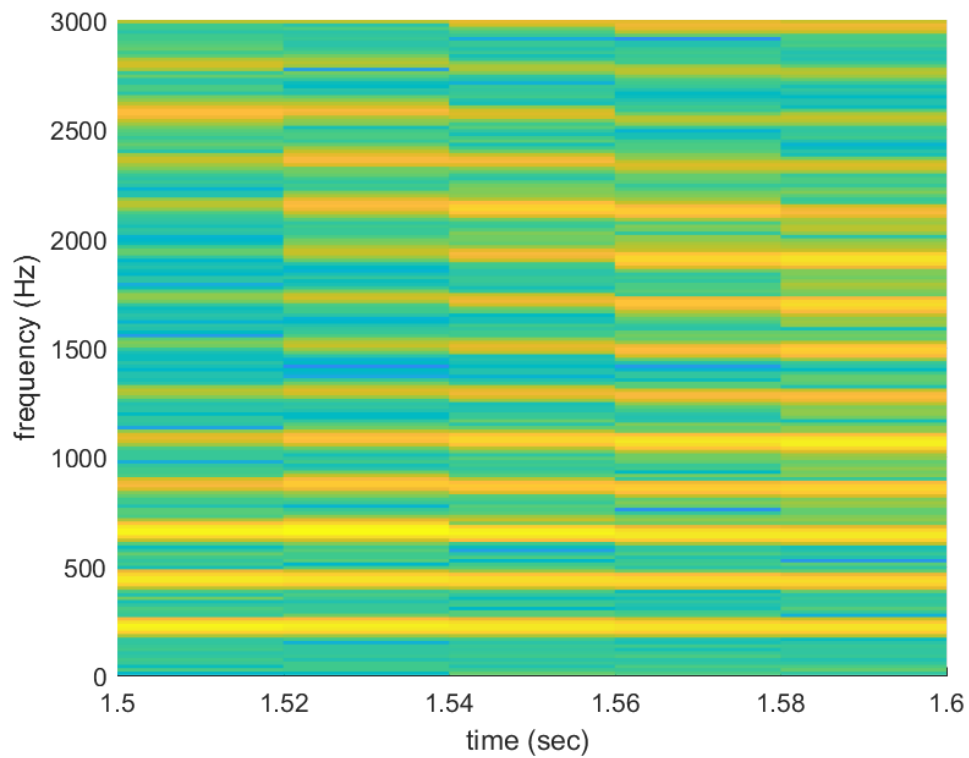
Digital Signal Processing



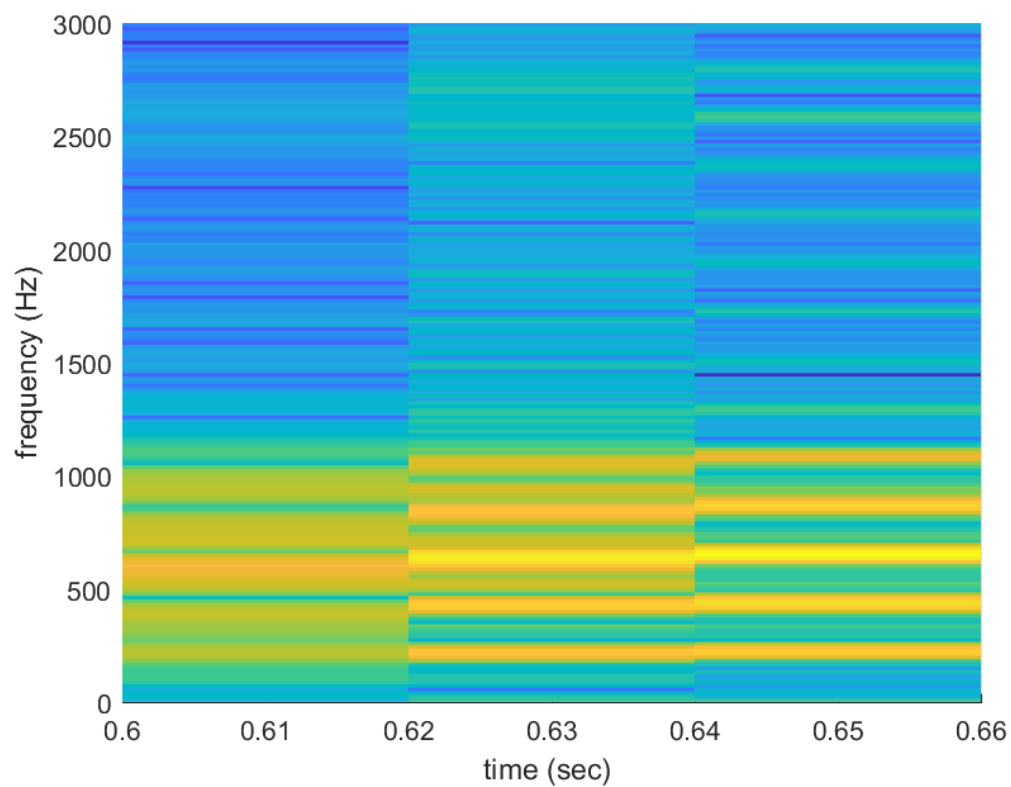
Εικόνα 29: Η κυματομορφή του /ο/ στη λέξη 'μυαλό'.



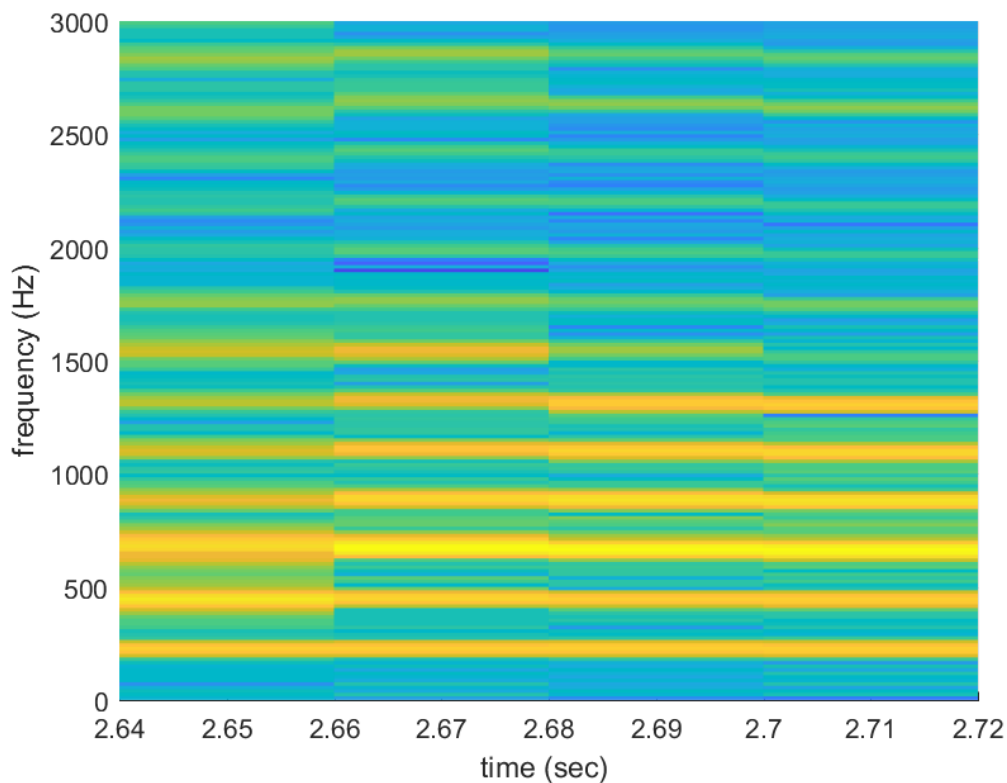
Εικόνα 30: Spectrogram του γράμματος /α/ στην λέξη 'όλα'.



Εικόνα 31: Spectrogram του γράμματος /α/ στην λέξη 'άμυνα'.



Εικόνα 32: Spectrogram του γράμματος /ο/ στην λέξη 'όλα'.



Εικόνα 33: Spectrogram του γράμματος /ο/ στην λέξη 'μυαλό'.

Παρατηρήσεις:

- ✓ Παρατηρούμε ότι τα φωνήεντα /ο/ και /α/ έχουν έντονες κίτρινες γραμμές που σημαίνει ότι είναι περιοδικά με συγκεκριμένες συχνότητες. Αυτό προκύπτει βέβαια και από τις γραφικές παραστάσεις στο χρόνο αλλά είναι πολύ πιο ξεκάθαρο στο spectrogram.
- ✓ Τα spectrogram τους δηλαδή αποτελούνται από τα λεγόμενα formants, τα οποία για κάθε φωνήεν παίρνουν συγκεκριμένες τιμές και εκφράζουν τις συχνότητες του γράμματος. Στην περίπτωσή μας δεν είναι τόσο εμφανής η διαφορά μεταξύ των δύο φωνηέντων, αλλά υπάρχουν πολλές μελέτες όπου φαίνονται τα formants των φωνηέντων, όπως στην εικόνα παρακάτω:

a			
	Male	Female	Child
F_1	736	873	922
F_2	1466	1699	1811
F_3	2459	2713	3197

o			
	Male	Female	Child
F_1	583	657	730
F_2	1137	1219	1462
F_3	2479	2817	3240

Εικόνα 34: Φωνήεντα /α/ και /ο/ και οι συχνότητές τους.

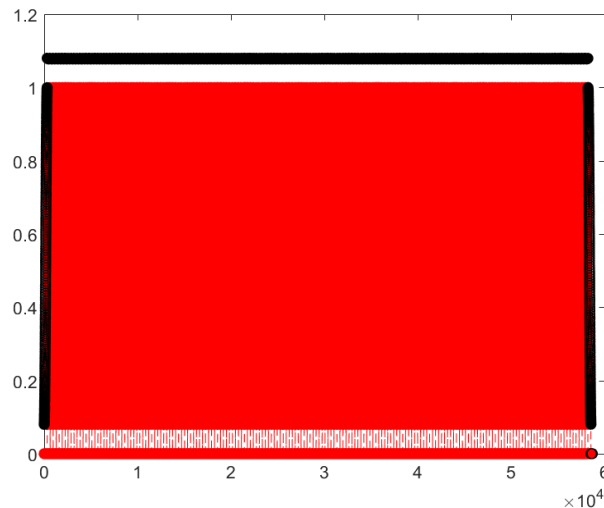
-Οποιοδήποτε υλικό χρησιμοποιήθηκε πέρα από το βιβλίο και το εκπαιδευτικό υλικό του μαθήματος αναφέρεται στην ενότητα Βιβλιογραφία-

2.3. Ανακατασκευή του σήματος μέσω MATLAB. Το σήμα που προέκυψε είναι το αρχείο speech_utterance_rec.wav .

2.4. Σύμφωνα και με το Παράρτημα Α, η πλήρης ανακατασκευή του σήματος είναι εφικτή όταν ισχύει η παρακάτω σχέση:

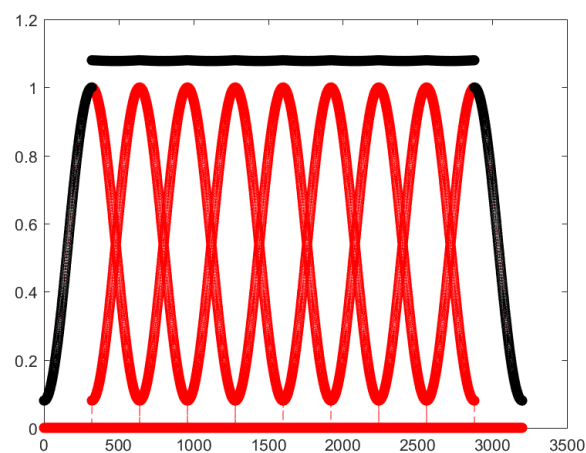
$\sum_{m \in \mathbb{Z}} w[n - mR] = 1, \forall n \in \mathbb{Z}$ όπου $w[n]$ παράθυρο hamming στην περίπτωση μας.

Η συνάρτηση ola.m, λοιπόν, ελέγχει αν ισχύει αυτή η σχέση. Δέχεται τρία ορίσματα: το μήκος του παραθύρου, την επικάλυψη, και το μήκος του σήματος. Ως έξοδο τυπώνει όλα τα παράθυρα το ένα μετά το άλλο με κόκκινο χρώμα και το άθροισμά τους με μαύρο (που θα πρέπει να ισούται με 1). Αν την καλέσουμε με τις δικές μας τιμές, δηλαδή μήκος παραθύρου = 640, επικάλυψη = 320 και μέγεθος σήματος = 58720 έχουμε το παρακάτω:



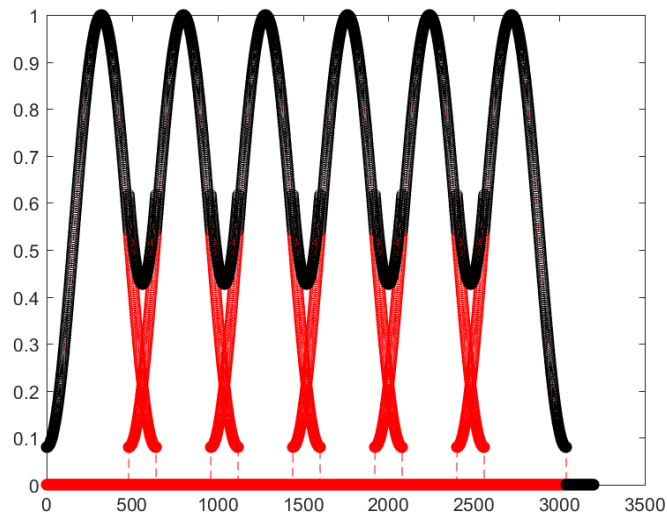
Εικόνα 35

Προφανώς, επειδή τα επικαλυπτόμενα παράθυρα είναι πάρα πολλά βλέπουμε απλά ένα κόκκινο παράθυρο. Η μαύρη γραμμή, ωστόσο, που μας ενδιαφέρει είναι λίγο πάνω από την μονάδα πράγμα που επιβεβαιώνει ότι με αυτές στις συνθήκες το σήμα ανακατασκευάζεται πλήρως. Με μήκος σήματος $= 5 \cdot 640 = 3200$ προκύπτει το παρακάτω που μας επιτρέπει να δούμε και τα επιμέρους παράθυρα (η αλλαγή του μήκους του σήματος δεν επηρεάζει το αποτέλεσμα):



Εικόνα 36

Αν, τώρα, χρησιμοποιηθεί μήκος παραθύρου ίσο με 40 ms και βήμα ανάλυσης 30 ms προκύπτει το ανακατασκευασμένο σήμα `speech_utterance_rec2.wav`, το οποίο αν το ακούσουμε είναι εμφανώς παραμορφωμένο. Μέσω της συνάρτησης `ola.m` με ορίσματα 640, 480, 3200 προκύπτει το παρακάτω που μας δείχνει ότι η προϋπόθεση σωστής ανακατασκευής δεν ισχύει.



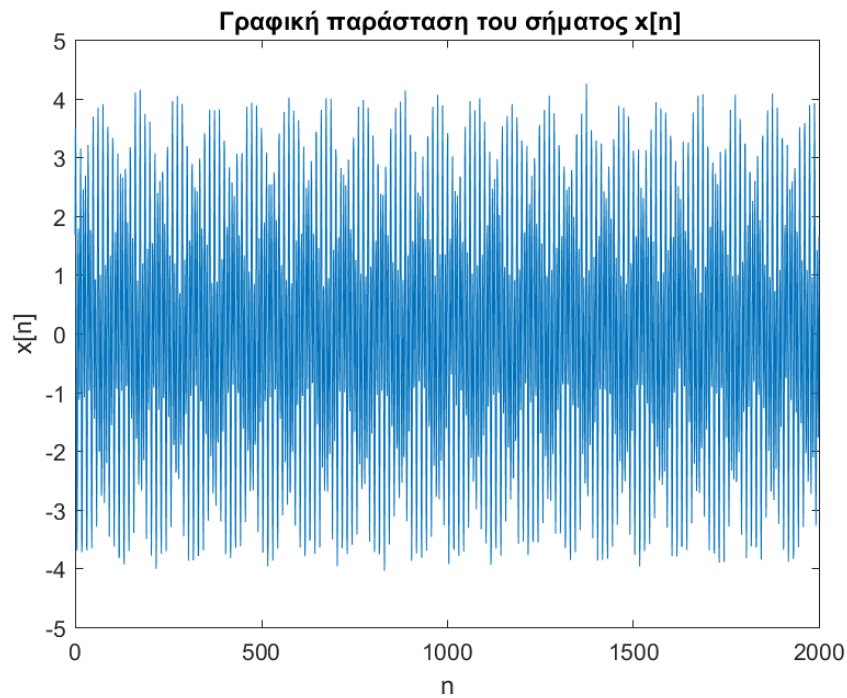
Εικόνα 37

✚ Μέρος 3^ο: Φασματική Ανάλυση Ημιτονοειδών και Ανίχνευση Απότομων Μεταβάσεων με τον Μετ/σμο Fourier Βραχέος Χρόνου (STFT) και τον Μετ/σμο Wavelets (διακριτοποιημένο DT-CWT).

Σκοπός του τρίτου μέρους η μελέτη του μετασχηματισμού με wavelets και η σύγκρισή του με τον STFT στην φασματική ανάλυση ημιτονοειδών σημάτων και στην ανίχνευση απότομων μεταβάσεων.

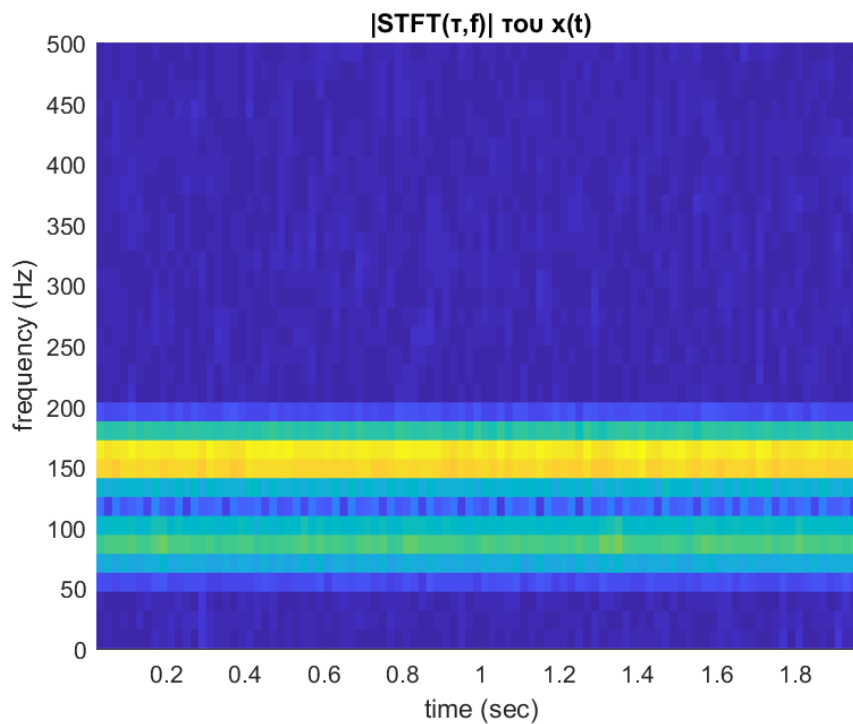
- 3.1. Στο ερώτημα αυτό θα ασχοληθούμε με το σήμα : $x(t) = 1.5 \cos(2\pi 80t) + 2.5 \sin(2\pi 150t) + 0.15v(t)$ στο διάστημα $[0,2]$ με ρυθμό δειγματοληψίας $F_s = 1000$ Hz.

a) Αρχικά παρουσιάζεται η μορφή του παραπάνω σήματος στον χρόνο.



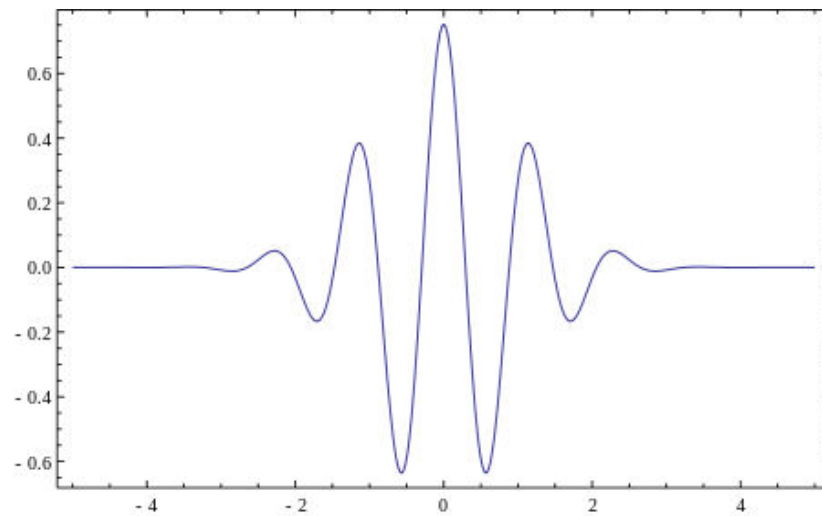
Εικόνα 38: $x(t)$ στο διάστημα $[0,2]$.

- b)** Χρησιμοποιώντας hamming παράθυρο μήκους 0.04sec και επικάλυψη ίση με 0.02sec υπολογίζουμε τον STFT του σήματος, το πλάτος του οποίου φαίνεται παρακάτω.

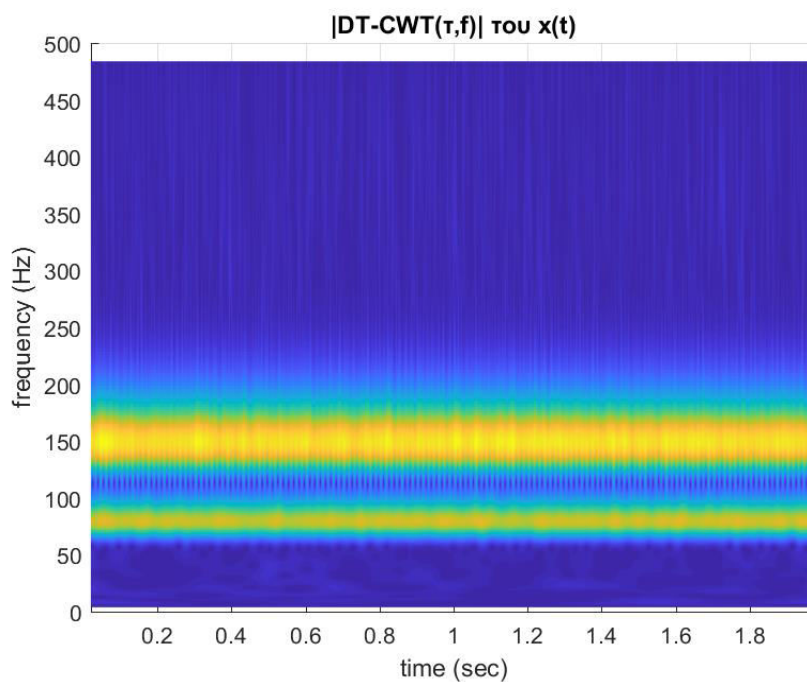


Εικόνα 39: Spectrogram του σήματος $x(t)$.

- c) Χρησιμοποιώντας το Morlet wavelet (Εικόνα 40) υπολογίσαμε τον DT-CWT, το πλάτος του οποίου συναρτήσει του χρόνου τ και της συχνότητας f φαίνεται στην Εικόνα 41.



Εικόνα 40: Πραγματικός μέρος του Morlet wavelet στον χρόνο.



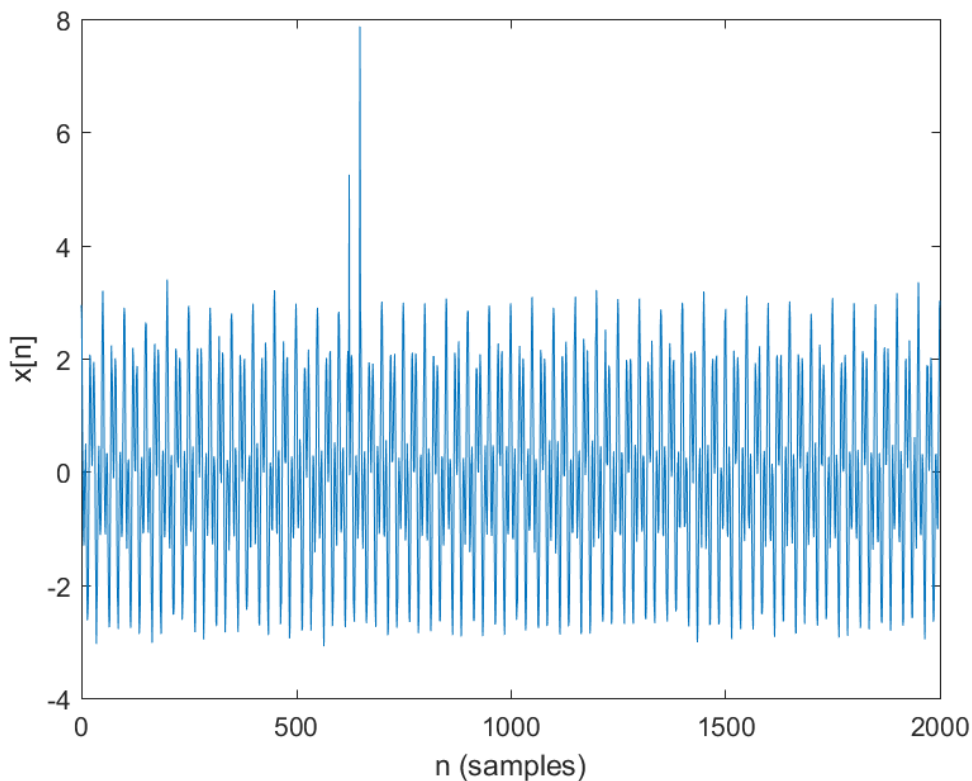
Εικόνα 41: Spectrogram με μετασχηματισμό DT-CWT του σήματος $x(t)$.

- d) Αυτό που παρατηρούμε είναι ότι το spectrogram που προέκυψε από τον μετασχηματισμό με wavelets είναι πολύ πιο ομαλό και μας δείχνει με μεγαλύτερη ακρίβεια τι συμβαίνει τόσο στο χρόνο όσο και στη συχνότητα (εδώ μας νοιάζει η συχνότητα κυρίως). Αυτό συμβαίνει γιατί ο STFT έχει σταθερό μήκος παραθύρου κάθε φορά το οποίο ορίζεται εκ των προτέρων. Το μήκος αυτό καθορίζει την σχέση μεταξύ της διακριτικής ικανότητας στη

συχνότητα και της ανάλυσης στο χρόνο. Έτσι, επιλέγοντας μικρό παράθυρο έχουμε πολύ καλή ανάλυση στο χρόνο αλλά χάνουμε πληροφορία στη συχνότητα και, αντίστροφα, για μεγάλο παράθυρο έχουμε καλή διακριτική ικανότητα στο πεδίο των συχνοτήτων χάνοντας σε ακρίβεια στο πεδίο του χρόνου. Στο μετασχηματισμό με wavelets, ωστόσο, δεν υπάρχει τέτοιος περιορισμός με αποτέλεσμα να προσδίδει ένα καλύτερο σε ποιότητα spectrogram.

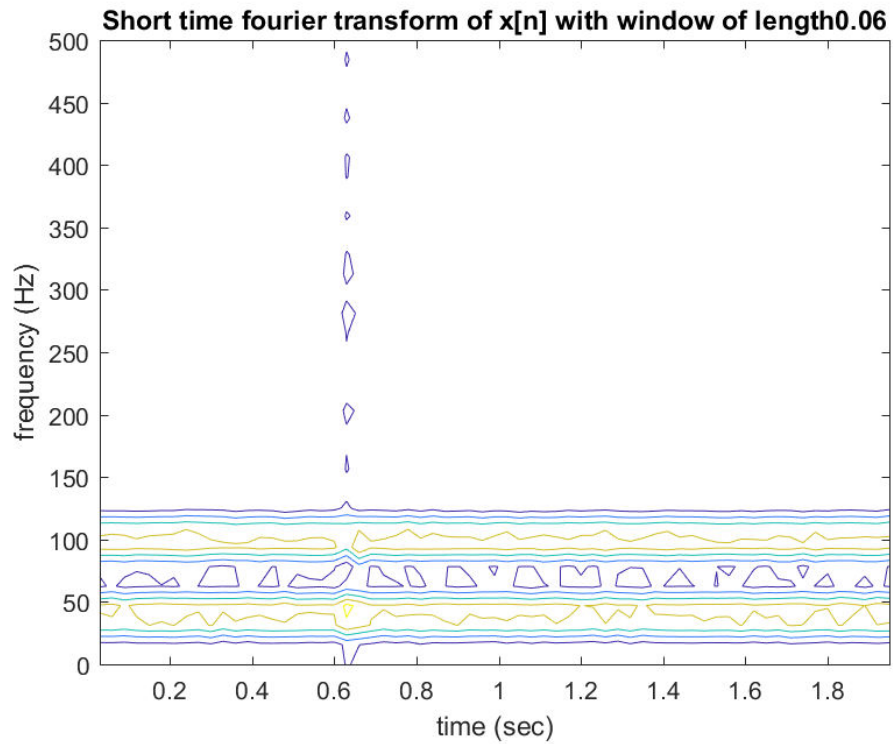
- 3.2.** Στο ερώτημα αυτό θα ασχοληθούμε με το σήμα : $x(t) = 1.5 \cos(2\pi 40t) + 1.5 \sin(2\pi 100t) + 0.15v(t) + 5(\delta(t - 0.625)) + \delta(t - 0.650)$ στο διάστημα $[0,2]$ με ρυθμό δειγματοληψίας $F_s = 1000$ Hz. Ουσιαστικά, προσθέτουμε αυτήν την φορά συναρτήσεις $\delta[n]$ για να μελετήσουμε ποιος μετασχηματισμός ανταποκρίνεται καλύτερα σε απότομες μεταβολές.

- a)** Παρουσιάζεται πρώτα η μορφή του σήματος $x(t)$ στον χρόνο για διάστημα $[0,2]$ σε sec.

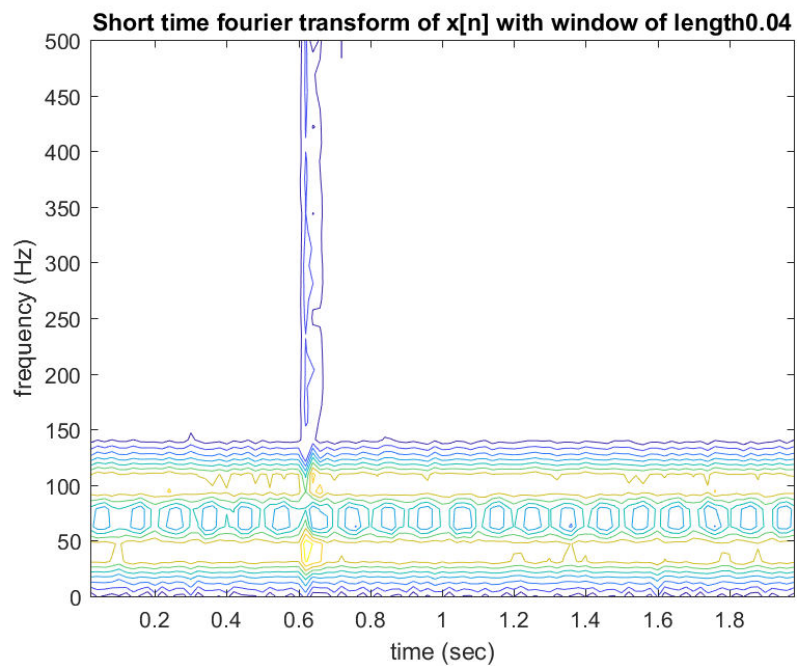


Εικόνα 42: Σήμα $x[n]$ στον χρόνο.

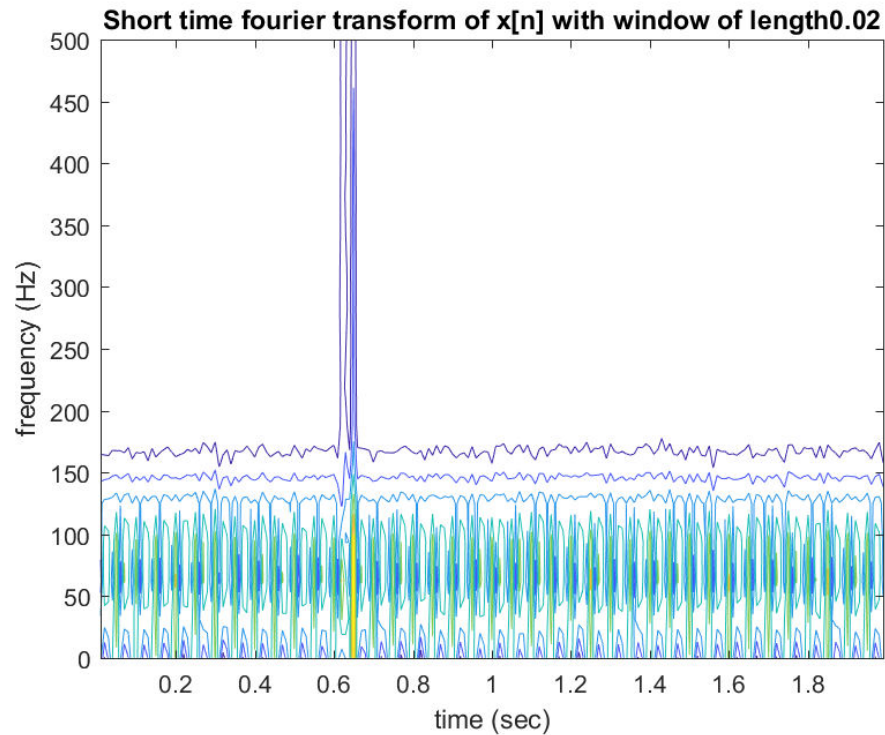
- b)** Τώρα θα υπολογίσουμε τον STFT του παραπάνω σήματος με hamming παράθυρο μεταβλητού μήκους και σταθερής επικάλυψης 50%. Το μήκος ισούται με 60ms, 40ms και 20ms.



Εικόνα 43: Contour του $|STFT|$ του $x[n]$ με μήκος παραθύρου 60ms.

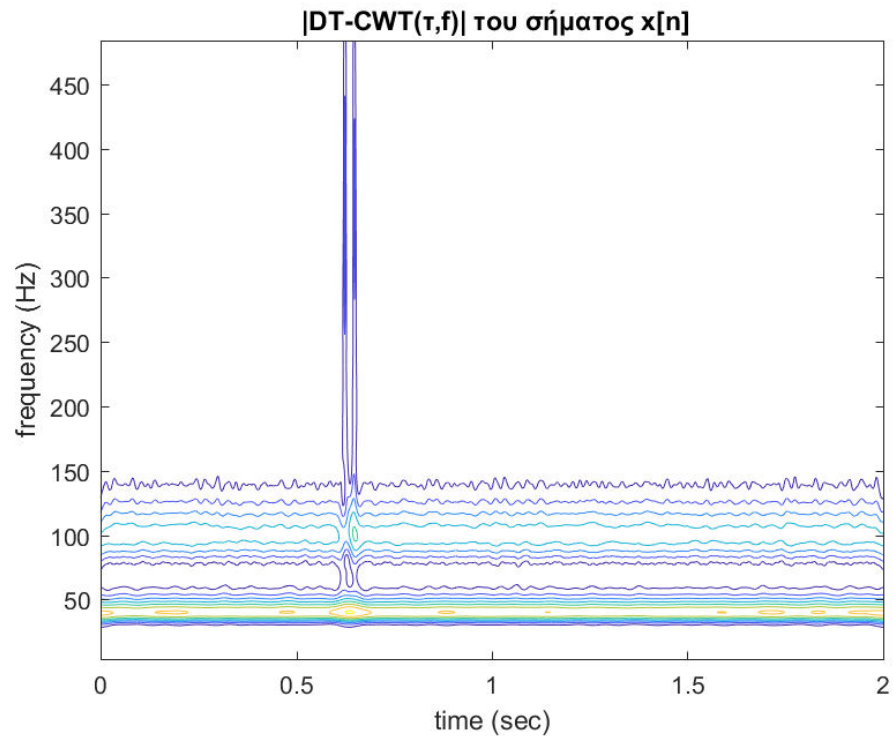


Εικόνα 44: Contour του $|STFT|$ του $x[n]$ με μήκος παραθύρου 40ms.



Εικόνα 45: Contour του $|STFT|$ του $x[n]$ με μήκος παραθύρου 20ms.

- c) Αντίστοιχα υπολογίστηκε και ο μετασχηματισμός με wavelets του $x[n]$, το πλάτος του οποίου φαίνεται στην παρακάτω εικόνα.



Εικόνα 46: Contour του $|DT-CWT|$ του σήματος $x[n]$

d) Στο σήμα της συγκεκριμένης άσκησης έχουμε δύο κυρίαρχα χαρακτηριστικά. Πρώτον, περιέχει δύο ημιτονοειδή σήματα με συχνότητες 50 Hz και 100 Hz και, δεύτερον, παρουσιάζονται στα χρονικά σημεία 0.625 και 0.650 απότομες μεταβολές του πλάτους. Άρα οι παρατηρήσεις μας πρέπει να εστιάσουν σε αυτά τα δύο, δηλαδή κατά πόσο ο κάθε μετασχηματισμός εντοπίζει τις δύο κυρίαρχες συχνότητες και, παράλληλα, και τις δύο απότομες μεταβολές. Συγκεκριμένα, έχουμε:

- ✓ STFT με μήκος παραθύρου 0.06 sec: Διακρίνονται με μεγάλη ευκολία και ευκρίνεια οι δύο κυρίαρχες συχνότητες του σήματος, ωστόσο οι απότομες μεταβολές ίσα-ίσα που φαίνονται με τη μορφή κάποιων κουκίδων.
- ✓ STFT με μήκος παραθύρου 0.04 sec: Διακρίνονται αρκετά εύκολα οι κυρίαρχες συχνότητες, αλλά με όχι τόσο μεγάλη ακρίβεια οι χρονικές στιγμές των απότομων μεταβολών στο χρόνο.
- ✓ STFT με μήκος παραθύρου 0.02 sec: Διακρίνονται με μεγάλη ακρίβεια οι μεταβολές στο χρόνο αλλά οι κυρίαρχες συχνότητες είναι αδύνατον να διακριθούν έστω και στο περίπου.
- ✓ DT-CWT με Morlet wavelet: Διακρίνονται με πολύ μεγάλη ακρίβεια οι απότομες μεταβολές στο χρόνο. Όσων αφορά τις κυρίαρχες συχνότητες, διακρίνονται και αυτές (γαλάζιο χρώμα) με λιγότερη ακρίβεια.

Συμπέρασμα: Όπως ήδη αναφέραμε το μήκος του παραθύρου στον STFT καθορίζει που θα έχουμε καλύτερη ανάλυση στη συχνότητα ή στο χρόνο. Αυτό επιβεβαιώνεται και στην περίπτωση μας. Οι δύο ακραίες τιμές (0.06sec και 0.02 sec) δεν καταφέρνουν σίγουρα τον στόχο μας καθώς η πρώτη (έχοντας καλή ανάλυση μόνο στη συχνότητα) “χάνει” τις απότομες μεταβολές στο χρόνο και η δεύτερη (έχοντας καλή ανάλυση στη χρόνο) δεν αντιλαμβάνεται τις κυρίαρχες συχνότητες του σήματος. Έτσι, η επιλογή μας περιορίζεται ανάμεσα στις υπόλοιπες δύο με τον STFT των 40 ms να διακρίνει με καλύτερη ακρίβεια τις συχνότητες και τον wavelet μετασχηματισμό τις χρονικές στιγμές. Αυτός, όμως, που τα καταφέρνει εξίσου καλά και στα δύο είναι ο δεύτερος. Συνεπώς, ο μετασχηματισμός με wavelets μας επιτρέπει να διακρίνουμε τόσο βασικές συχνότητες ενός σήματος όσο και τις απότομες μεταβολές του στο χρόνο (συνδυάζει ακρίβεια στο χρόνο και την συχνότητα).



Βιβλιογραφία:

- <http://iitg.vlab.co.in/?sub=59&brch=164&sim=857&cnt=1>
- https://www.asee.org/documents/zones/zone1/2008/student/ASEE12008_0044_paper.pdf
- http://www.csd.uoc.gr/~hy578/2018/Project0_Part1.pdf
- https://en.wikipedia.org/wiki/Morlet_wavelet
- <http://inet.vidyasagar.ac.in:8080/jspui/handle/123456789/779>
- <http://www.d.umn.edu/~skatsev/Phys5053/WaveletHelp.pdf>