

Εθνικό Μετσόβιο Πολυτεχνείο



Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Η/Υ

6^ο Εξάμηνο

Ροή Σ

Ψηφιακή Επεξεργασία Σημάτων



1^η Εργαστηριακή Άσκηση στο Matlab

Αθανασίου Νικόλαος

03112074

ΜΕΡΟΣ 1^ο

Σύστημα Εντοπισμού Τηλεφωνικών Τόνων

Σκοπός της άσκησης είναι η μελέτη και η αποκωδικοποίηση του τονικού συστήματος με χρήση πίνακα αντιστοιχίας πλήκτρων και σημάτων με διαφορετικές συχνότητες. Το πάτημα ενός πλήκτρου στο τηλέφωνο παράγει έναν ήχο ως άθροισμα δύο ημιτόνων, ένα υψηλής συχνότητας (στήλη) και ένα χαμηλότερης (γραμμή) του touchpad. Οι συχνότητες που αντιστοιχούν σε κάθε στήλη και γραμμή φαίνονται στον παρακάτω πίνακα.

ω_2	ω_1		
	0.9273	1.0247	1.1328
0.5346	1	2	3
0.5906	4	5	6
0.6535	7	8	9
0.7217	0		

Διακριτές συχνότητες για τηλεφωνικούς τόνους σε συχνότητα δειγματοληψίας 8192Hz.

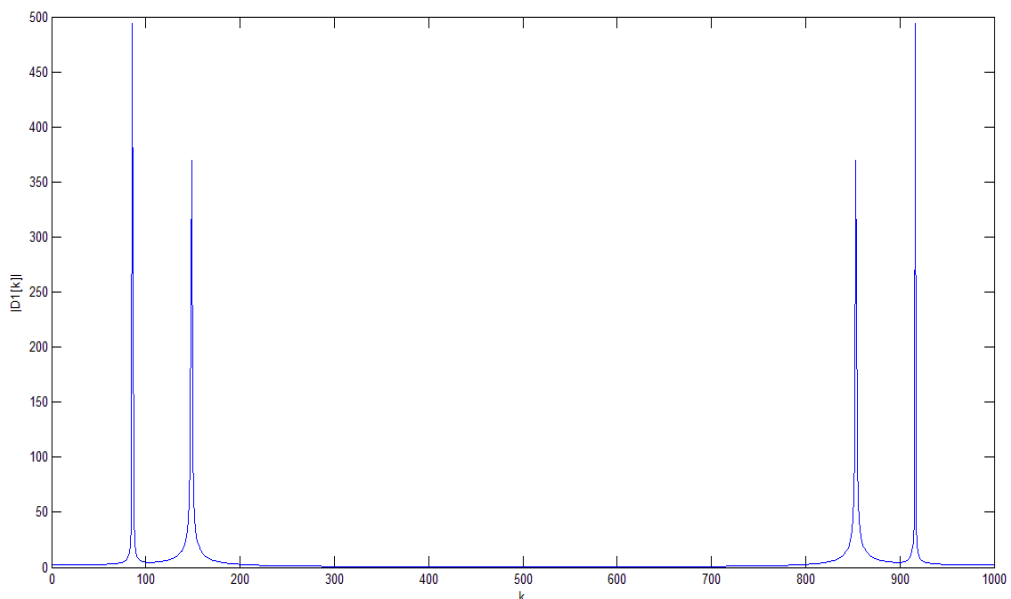
1.1.

Αρχικά δημιουργούνται 10 διαφορετικοί τόνοι ($d_0 - d_9$) σύμφωνα με το παραπάνω πίνακα και τις οδηγίες της άσκησης ως άθροισμα δύο ημιτόνων των δύο αντίστοιχων συχνοτήτων.

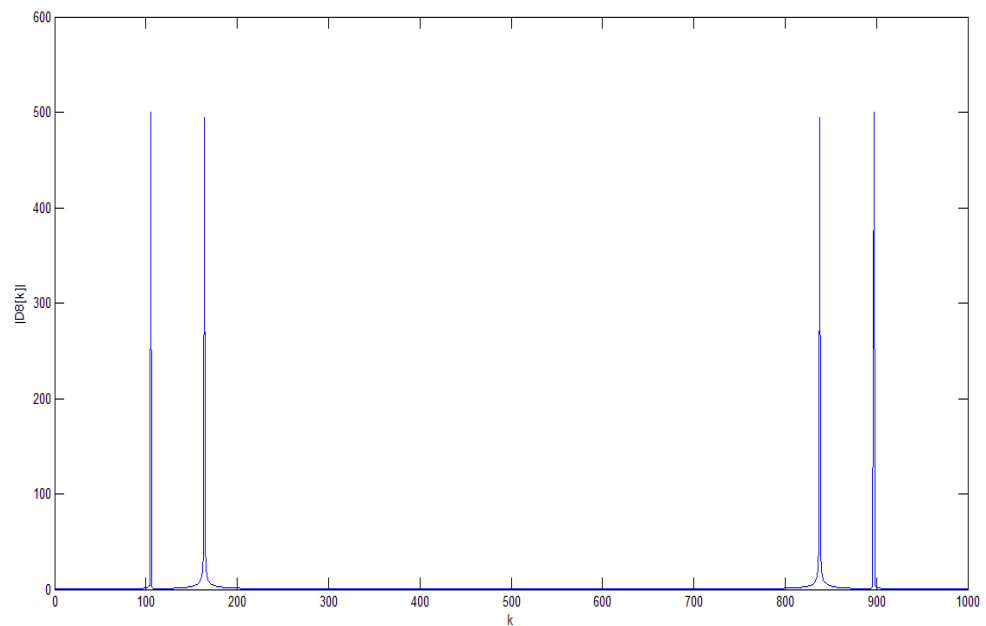
1.2.

Στη συνέχεια υπολογίζεται ο DFT των $d_1[n]$ και $d_8[n]$. Στο Matlab, ο μετασχηματισμός υλοποιείται μέσω του γνωστού αλγορίθμου- συνάρτησης *fft()*. Επίσης δημιουργούνται οι γραφικές παραστάσεις των μέτρων των παραπάνω μετασχηματισμών (σύμφωνα πάντα με την εξίσωση)

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



Πλάτος DFT του $d_1[n]$



Πλάτος DFT του $d_8[n]$

Παρατηρούμε ότι το παρουσιάζονται οι μέγιστες τιμές στους δείκτες k που αντιστοιχούν στις δύο συχνότητες που απαρτίζουν το κάθε σήμα. Η αντιστοίχιση γίνεται σύμφωνα με την σχέση

$$\omega = \frac{2\pi}{N} \cdot k = \frac{2\pi}{1000} \cdot k \quad (1.1)$$

που προκύπτει από τον ορισμό του DFT (Εξ. (1)).

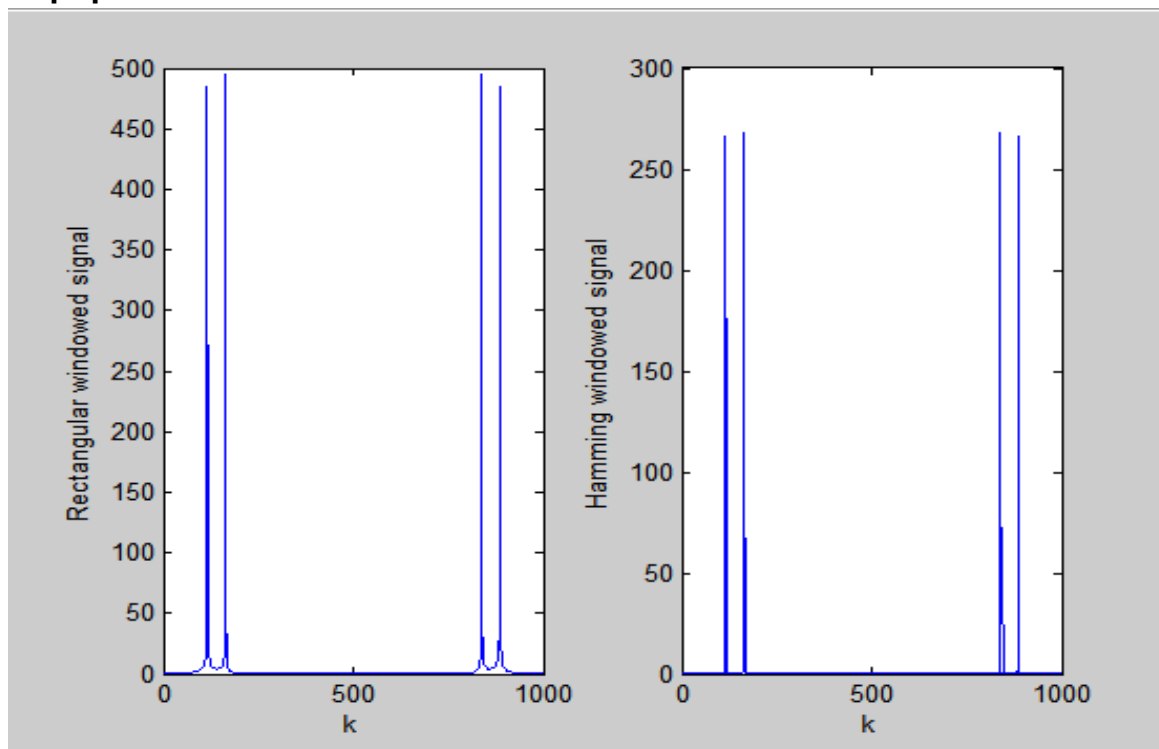
1.3.

Έπειτα ο αριθμός μητρώου μου 03112074, θα μεταφραστεί ως σήμα διαδοχικών τηλεφωνικών τόνων, με 1 – 1 αντιστοίχιση των πλήκτρων με τους παραπάνω αριθμούς. Δημιουργείται δηλαδή ένας πίνακας με καθένα τόνο δαιχωρίζοντας τους με 100 μηδενικά δείγματα. Έπειτα το σήμα αυτό ηχογραφείται στο αρχείο **03112074_tone_sequence.wav** με τη βοήθεια της **wavwrite()** (με συχνότητα δειγματοληψίας 8000Hz). Το αρχείο βρίσκεται στον παραδοτέο συμπιεσμένο φάκελο.

1.4.

Στο συγκεκριμένο ερώτημα δημιουργούμε i) ένα τετραγωνικό παράθυρο με τη βοήθεια της `ones()` και ii) ένα παράθυρο Hamming μήκους 1000 το καθένα. Στη συνέχεια, παίρνω παράθυρα του σήματος που είχαμε στο τέλος του 1.3 στα σημεία που βρίσκεται κάθε τόνος και υπολογίζω τους κατάλληλους Fourier. Προφανώς, προκύπτουν 8 παραθυροποιημένα σήματα, όσοι και οι τόνοι που περιέχονται στο αρχικό σήμα.

Παρουσιάζονται ένα σήμα επεξεργασμένο με παράθυρο hamming και ένα με τετραγωνικό.



Συγκρίνοντας τις δύο μεθόδους εξάγεται το συμπέρασμα ότι μέσω του παραθύρου Hamming προκύπτει ακρίβεστερο αποτέλεσμα έχουμε για την ακρίβεια λιγότερους κυματισμούς απ' ότι έχουμε με το τετραγωνικό παράθυρο.

1.5.

Όπως αναφέρθηκε προηγουμένως οι συχνότητες συνδέονται με τους δείκτες k σύμφωνα με τη σχέση (1.2). Λύνοντας ως προς k πρόκειται η ισοδύναμη

$$\text{μορφή } k = \frac{1000\omega}{2\pi}.$$

Έτσι υπολογίζονται για τις χαμηλές και για τις ψηλές συχνότητες του δοσμένου πίνακα συχνοτήτων τα αντίστοιχα k . Τα αποτελέσματα φαίνονται στον επόμενο πίνακα.

Συχνότητες γραμμής ($\omega_{\text{row}} - k_{\text{row}}$)		Συχνότητες στήλης ($\omega_{\text{col}} - k_{\text{col}}$)	
ω	K	Ω	K
0.5346	86	0.9273	149
0.5906	95	1.0247	164
0.6535	105	1.1328	181
0.7217	116		

Πίνακας αντιστοίχισης συχνοτήτων ω σε δείκτες k

Για το παραπάνω ερώτημα λόγω της πολυπλοκότητας του παρατίθεται και ο τρόπος που δημιουργήθηκε ο κώδικας.

Δημιούργησα ένα νέο πίνακα οποίος περιέχει τους τόνους του touch-pad με σειρά 1234567890 με 100 μηδενικά ανάμεσα στον καθένα .

Συνεχίζω και πολλαπλασιάζω κάθε τόνο του πίνακα αυτού με ένα παράθυρο hamming μήκους 1000 δειγμάτων και δημιουργώ 10 παραθυροποιημένα σήματα των οποίων βρίσκω τον dft. Παρατηρώντας τα διαγράμματα των παραπάνω μεταβλητών, συμπεραίνω ότι για να ποάρω τα ζητούμενα k αρκεί να αναζητήσω τις κορυφές με ύψος μεγαλύτερο του 205. Ωστόσο επειδή έχω τις θεωρητικές τιμές των k στον παραπάνω πίνακα καθολικά μικρότερες από το 2000 κι ο DFT είναι συμμετρικός μπορώ να θεωρήσω γενικότερα ότι μου αρκούν τα k μικρότερα από το 500. Αυτά αποθηκεύονται στον x πίνακα. Τελικά ο πίνακας x είναι μια λίστα από k όπου ανά δυο τιμές θα είναι τα ζητούμενα k για κάθε τόνο .Επίσης στις περιττές θέσεις του πίνακα x είναι τα k τα «υψηλόσυχνα», ενώ στις άρτιες τα «χαμηλόσυχνα» k.

1.6

Στο συγκεκριμένο μέρος της άσκησης δημιουργείται η συνάρτηση με όνομα **ttdecode()** που παίρνει τονικό σήμα αυτή τη φορά ως είσοδο και παράγει τονικά ψηφία. Αρχικά εντοπίζονται τα σημεία του σήματος στα οποία βρίσκονται οι τόνοι. Η θέση του πρώτου δείγματος κάθε τόνου αποθηκεύεται σε έναν πίνακα ώστε να είναι εύκολη η μετέπειτα πρόσβαση σε αυτόν τον τόνο. Αφού διαβαστεί ολόκληρο το σήμα εισόδου το μέγεθος του παραπάνω πίνακα θα είναι ίσο με το των αριθμός των τόνων που περιέχονται στο σήμα εισόδου.

Στη συνέχεια επεξεργαζόμαστε κάθε τόνο ξεχωριστά. Δημιουργούμε νέο διάνυσμα που περιέχει μόνο τα 1000 δείγματα του εκάστοτε τόνου, υπολογίζουμε τον DFT του διανύσματος. Έπειτα κρατάμε μόνο το πρώτο μισό του μετασχηματισμένου σήματος καθώς γνωρίζουμε ότι ο μετασχηματισμός DFT του σήματος μήκους 1000 είναι συμμετρικός. Θα έχω δύο σημείο εμφάνισης μέγιστης τιμής αφού πρόκειται για δύο ημίτονα διαφορετικών συχνοτήτων τα οποία εντοπίζονται μέσω της **find()** .

Τα δύο σημεία, που εντοπίζονται μέσω της συνάρτησης **find()**, αποτελούν τους δείκτες k που αντιστοιχούν στις συχνότητες των δύο ημιτόνων. Έτσι, εισάγονται ως είσοδος σε μία βοηθητική συνάρτηση **Search()** η οποία υπολογίζει ποιο πλήκτρο αντιστοιχεί στη συγκεκριμένη δυάδα. Η έξοδός της αποθηκεύεται σε ένα νέο διάνυσμα που ονομάζεται **Vector** και περιέχει την πληροφορία για το ποιο πλήκτρο έχει πατηθεί κάθε φορά.

Αφού πραγματοποιηθεί η επεξεργασία για κάθε τόνο ξεχωριστά η συνάρτηση **ttdecode()** επιστρέφει ως έξοδο το **Vector**. Για να διαπιστωθεί η σωστή λειτουργία της συνάρτησης **ttdecode()** θέτουμε ως είσοδο το σήμα που δημιουργήθηκε στο ερώτημα 1.3. Πράγματι, λαμβάνουμε το παρακάτω αποτέλεσμα:

Vector =0 6 2 2 4 1 0 0

1.7.

Στο τελευταίο μέρος της άσκησης φορτώνουμε το αρχείο “*my_touchtones.mat*” από το συμπληρωματικό υλικό της άσκησης “*dsp15 lab1 Data.zip*” με χρήση της εντολής ***load()***. Έτσι, προστίθενται δύο σήματα στο workspace: το *easySig* και το *hardSig*. Κάθε ένα από αυτά εισάγεται ως είσοδος στη συνάρτηση *ttdecode()* και λαμβάνουμε τα παρακάτω ψηφία:

easySig:

2 1 1 8 5 0 8 0 6 9

hardSig:

**6 6 2 2 6 6 2 2 6 6 0 0 0 0 9 9 3 3 9
9 3 3 9 9 0 0 2 2**

ΜΕΡΟΣ 2^ο

Φασματική Ανάλυση και Αίχνευση Ημιτονοειδών με τον Διακριτό Μετασχηματισμό Fourier (DFT)

Σκοπός της δεύτερου μέρους της εργαστηριακής άσκησης είναι η εξέταση της διακριτικής ικανότητας ενός συστήματος. Στα επιμέρους τμήματα της άσκησης το παράθυρο που θα χρησιμοποιηθεί είναι το παράθυρο Hamming, το οποίο ορίζεται σύμφωνα με τη σχέση:

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

Για την μελέτη αυτή, γίνεται χρήση δύο μιγαδικών ημιτονοειδών σημάτων με δοσμένες συχνότητες πλάτη και τυχαίες φασείς. Το σήμα που επεξεργαζόμαστε σύμφωνα και με την εκφώνηση της άσκησης είναι η συνέλιξη της συνάρτησης παραθυροποίησης με το άθροισμα των δύο σημάτων δηλαδή

$$y[n] = w[n] \cdot (x_1[n] + x_2[n]),$$

είναι το σήμα προς ανάλυση με παράμετρο προς εξέταση τη διαφορά των συχνοτήτων των δύο μιγαδικών εκθετικών.

Για τα τρία πρώτα τμήματα του δεύτερου μέρους θεωρούμε ότι τα ημιτονοειδή σήματα έχουν πλάτος 1 και 0,8, αντιστοίχως.

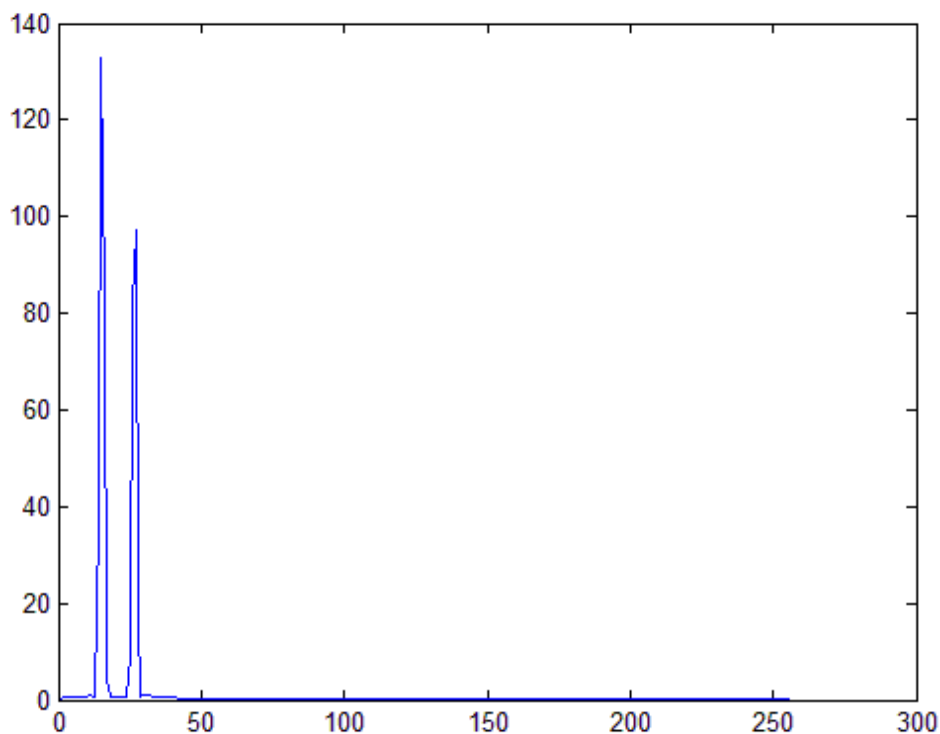
Σημείωση : Κομμάτι του κώδικα σχετικά με την τυχαιότητα της γωνίας έχει μπει σε σχόλια πιο συγκεκριμένα την επιλογή της από μια κατανομή με «τυχαίο τρόπο» αφού δεν είναι αυτό που μας ενδιαφέρει στην άσκηση και ορίστηκε μια τυχαία γωνία χειροκίνητα και όχι μέσω επιλογής από κάποια κατανομή ομοιόμορφη ή κανονική για παράδειγμα.

2.1.

Στο συγκεκριμένο ερώτημα κατασκευάζεται η συνάρτηση **dft()** και γίνεται ο υπολογισμός του dft ενός σήματος μέσω αυτής της συνάρτησης καθώς και μέσω της **fft()** που υπάρχει στο Matlab. Παρατηρείται ότι ο υπολογισμός μέσω της dft και της fft στο μόνο που διφέρουν είναι η πολυπλοκότητα.

2.2.

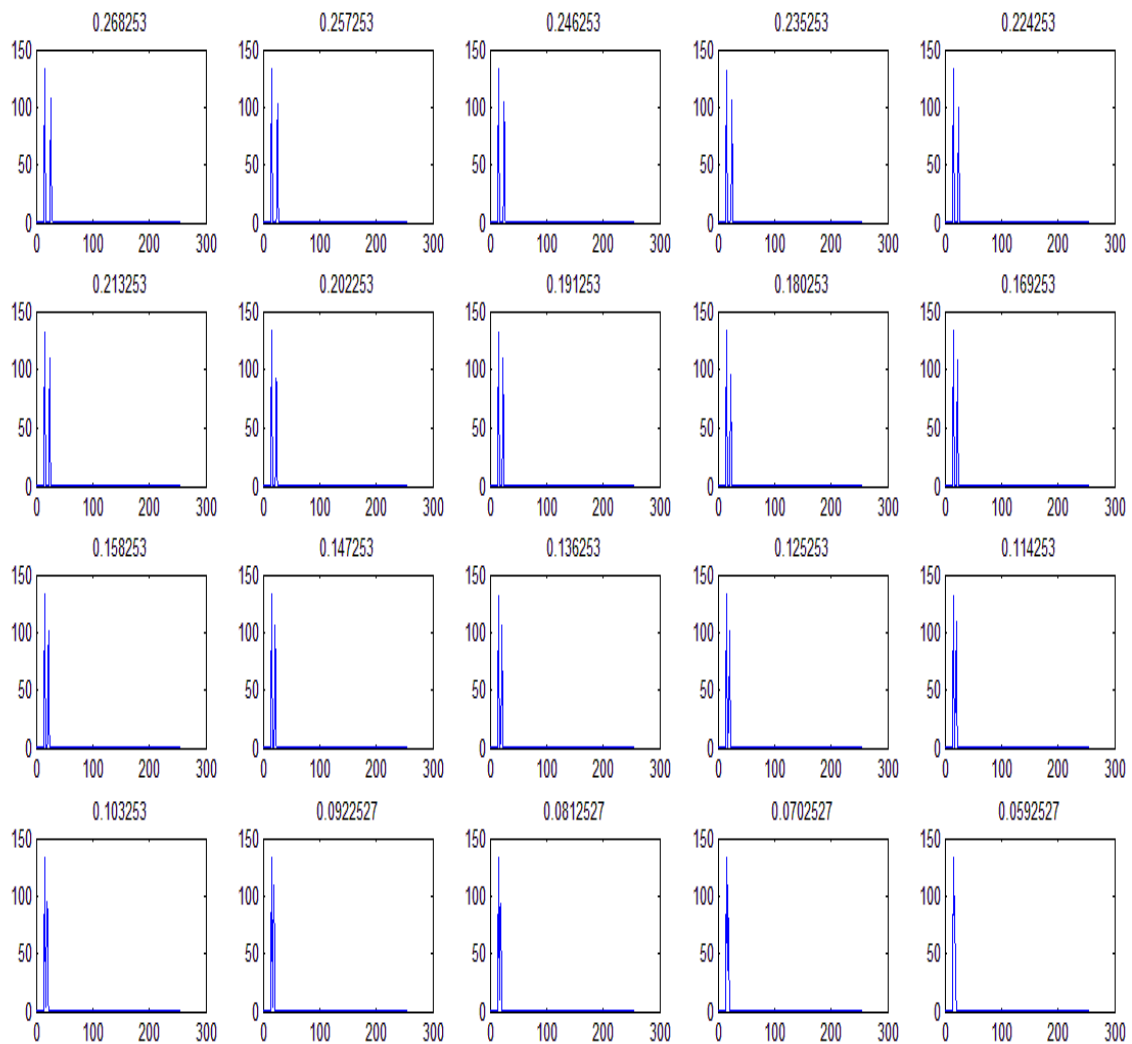
Αρχικά, θεωρώντας ότι το πρώτο σήμα έχει συχνότητα $\pi/9$ και το δεύτερο $\pi/5$ και επιλέγοντας αυθαίρετα ως φάσεις τις $\pi/42$ και $\pi/17$ αντιστοίχως, έγινε εισαγωγή του αθροίσματος των δύο σημάτων σε ένα παράθυρο Hamming μήκους 256 δειγμάτων. Χρησιμοποιώντας τον DFT μήκους επίσης 256 δειγμάτων, μετασχηματίστηκε η έξοδος y του παραθύρου και το πλάτος παρατίθεται στη συνέχεια, ως γραφική παράσταση συνάρτησης του k .



Γραφική παράσταση του πλάτους του DFT 256 σημείων του σήματος εξόδου $y[n]$ από παράθυρο 256 σημείων

Παρατηρούμε από την παραπάνω γραφική παράσταση ότι υπάρχει σαφής διάκριση μεταξύ των δύο ημιτόνων. Για να βρεθεί η οριακή τιμή του ω για την οποία συμπίπτουν οι δύο κορυφές, πραγματοποιείται σταδιακή μείωση της συχνότητας του δεύτερου σήματος και με παρατήρηση των γραφικών παραστάσεων του πλάτους του DFT γίνεται εντοπισμός αυτής της ελάχιστης τιμής.

Για την παρατήρηση των γραφικών καθώς σταδιακά αλλάζει η μια εκ των δύο συχνοτήτων δημιουργήθηκε η συνάρτηση `draw_subplots2()`, η οποία παίρνει ως παραμέτρους το πλήθος το σημείων του DFT, την αρχική συχνότητα του δεύτερου σήματος, το πρώτο σήμα και το βήμα μείωσης συχνότητας και σχεδιάζεται το ζητούμενο πλάτος και αναγράφεται ως τίτλος σε αυτό η αντίστοιχη διαφορά συχνότητας. Μειώνοντας σε κάθε επανάληψη τη συχνότητα κατά $0,011\text{rad/sec}$ και με DFT 256 σημείων προκύπτουν τα εξής:



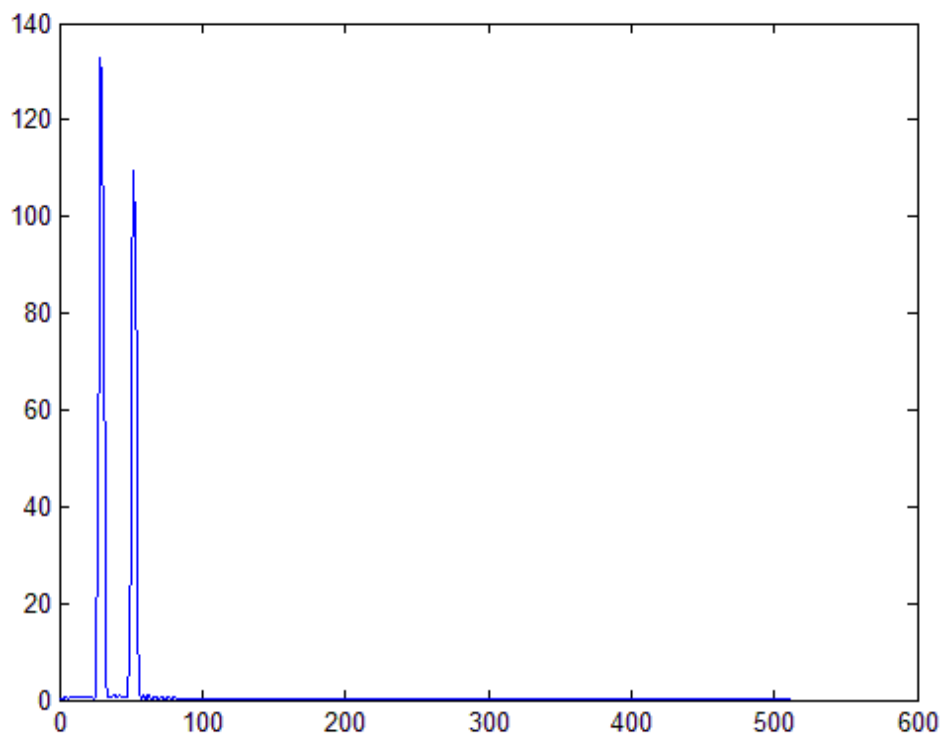
Σύνολο γραφικών παραστάσεων του πλάτους του DFT 256 σημείων του $y[n]$ για διάφορες τιμές της διαφοράς των συχνοτήτων των αρχικών σημάτων

Παρατηρώντας τις παραπάνω γραφικές παραστάσεις, γίνεται αντιληπτό ότι η ελάχιστη διαφορά συχνοτήτων για την οποία μπορούμε να διακρίνουμε τις δύο συχνότητες είναι $\Delta\omega = 0,0812527 \text{ rad/sec}$. Για μικρότερες διαφορές συχνοτήτων από αυτήν φαίνεται ότι υπάρχει επικάλυψη των δύο φασμάτων, η οποία δεν επιτρέπει τη φασματική διάκριση των προς εξέταση σημάτων.

2.3.

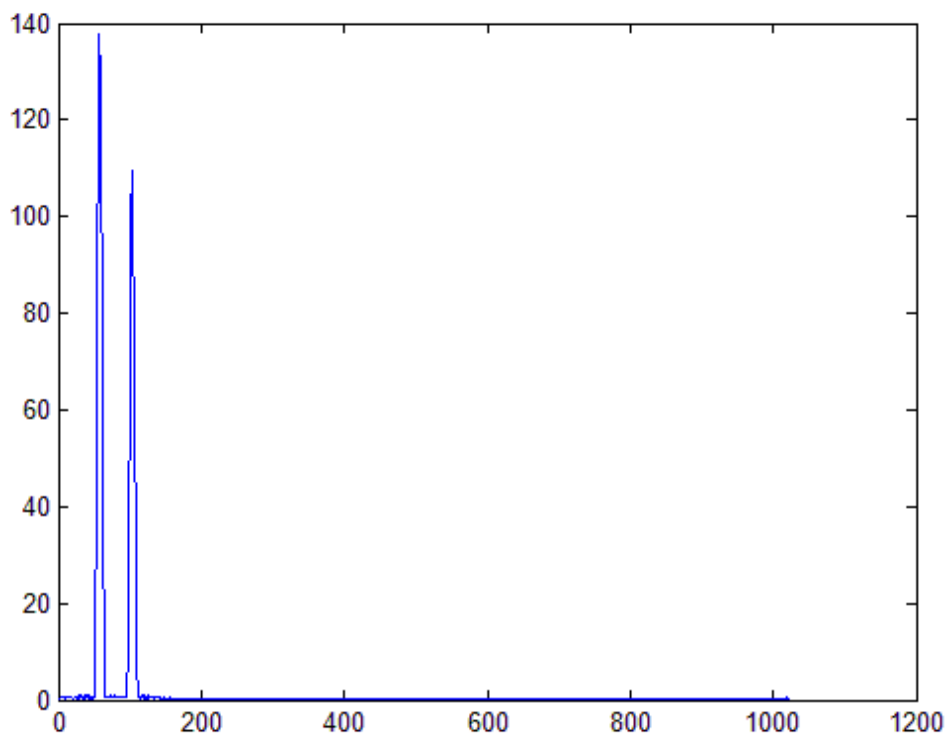
Τώρα, πράττουμε τα ίδια με το 2.2, αλλά εφαρμόζοντας zero-padding στην έξοδο του παραθύρου πριν το μετασχηματισμό της κατά Fourier, ώστε να παρατηρηθούν αλλαγές στην ευκρίνεια του σήματος και στη δυνατότητα φασματικής διάκρισης.

Αρχικά, θα εξετάσουμε την επίδραση του zero padding στην ευκρίνεια του σήματος, εφαρμόζοντας τη διαδικασία στο σήμα εξόδου ώστε να ληφθεί ο DFT 512 σημείων. Έτσι, λαμβάνουμε, μέσω του Matlab, τη γραφική παράσταση του μέτρου του DFT 512 σημείων του $y[n]$:



Γραφική παράσταση του πλάτους του DFT 512 σημείων του σήματος εξόδου $y[n]$ από παράθυρο 256 σημείων

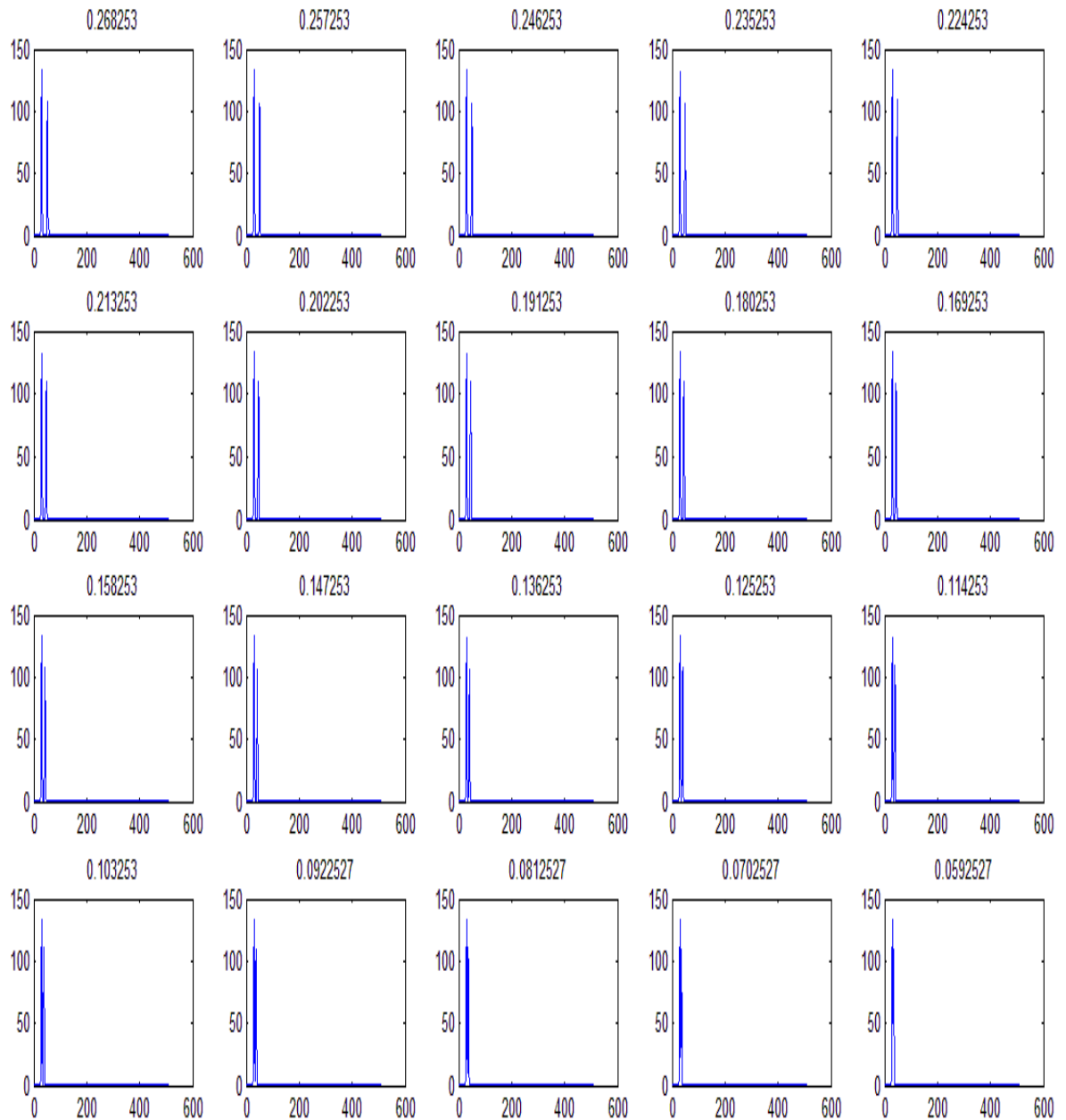
Για να σχολιαστεί με μεγαλύτερη ακρίβεια η επίδραση της διαδικασίας του zero padding στην ευκρίνεια του σήματος, θα επαναλάβουμε ώστε να προκύψει ο DFT 1024 σημείων. Στην περίπτωση αυτή θα προκύψει η ακόλουθη γραφική παράσταση:



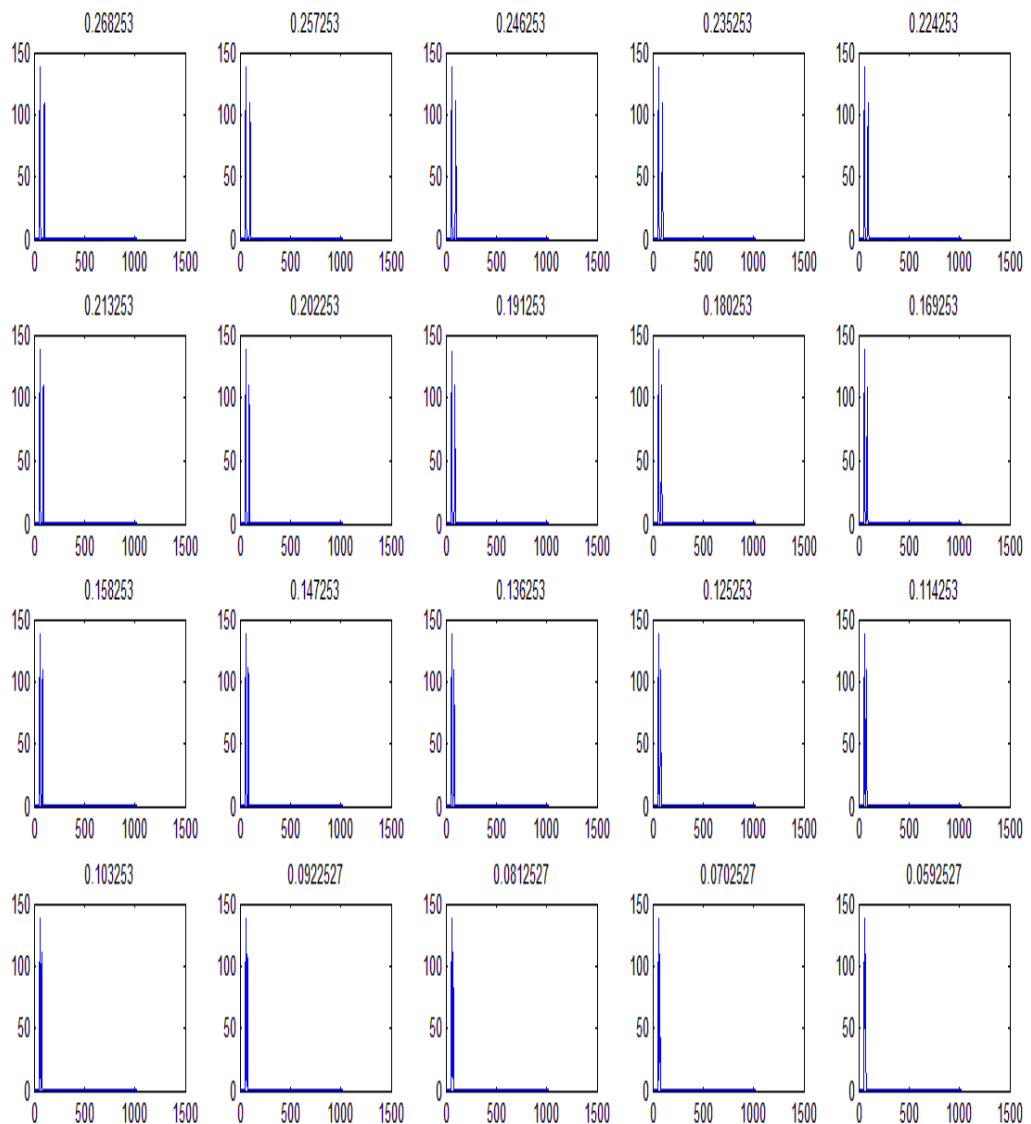
Γραφική παράσταση του πλάτους του DFT 1024 σημείων του σήματος εξόδου $y[n]$ από παράθυρο 256 σημείων

Παρατηρώντας τις δύο αυτές γραφικές παραστάσεις καθώς και την εικόνα του πρώτου ερωτήματος, διαπιστώνουμε ότι με τη διαδικασία του zero-padding μπορούμε να παραστήσουμε τον DFT με μεγαλύτερη ευκρίνεια, σε σχέση με την απεικόνιση του DFT μήκους ίσο με αυτό του σήματος. Επομένως, συμπεραίνουμε ότι υπάρχει μια σχέση αναλογίας μεταξύ του μήκους του DFT και της ευκρίνειας του φάσματος ενός σήματος.

Στη συνέχεια, θα εξετάσουμε την επίδραση του zero-padding στην δυνατότητα φασματικής διάκρισης των δύο ημιτόνων. Αυτό θα επιτευχτεί με τον ίδιο τρόπο με την ίδια επαναληπτική διαδικασία όπως και πριν, δηλαδή θα μειώνεται η συχνότητα του δεύτερου σήματος εισόδου και θα εντοπιστεί η ελάχιστη διαφορά συχνοτήτων ώστε να επιτρέπεται η διακρίση του από το άλλο ημίτονο με παρόμοιο βήμα. Τόσο για DFT 512 όσο και για 1024 σημείων τα αποτελέσματα παρατίθενται ακολούθως.



Σύνολο γραφικών παραστάσεων του πλάτους του DFT 512 σημείων του $y[n]$ για διάφορες τιμές της διαφοράς των συχνοτήτων των αρχικών σημάτων



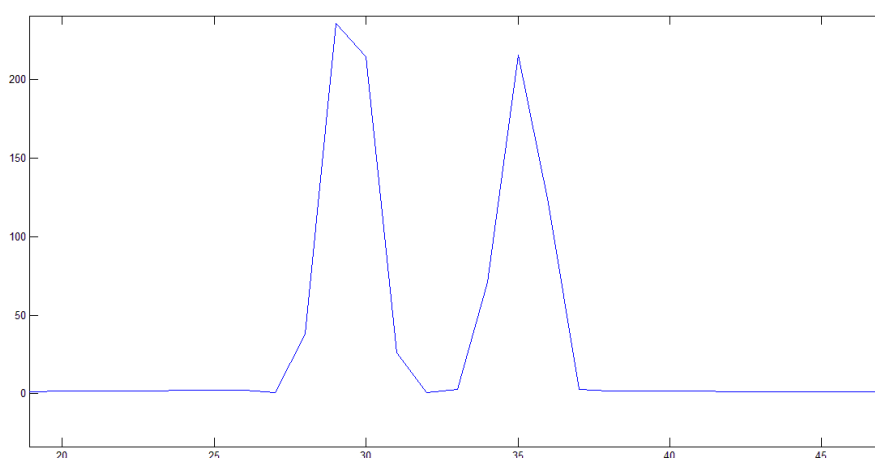
Σύνολο γραφικών παραστάσεων του πλάτους του DFT 1024 σημείων του $y[n]$ για διάφορες τιμές της διαφοράς των συχνοτήτων των αρχικών σημάτων

Παρατηρώντας τις εικόνες 2.5 και 2.6 συμπεραίνουμε ότι τόσο στην περίπτωση του DFT 512 όσο και σε αυτήν των 1024 σημείων η ελάχιστη διαφορά συχνοτήτων, ώστε να μπορούμε να παρατηρήσουμε το φάσμα και των δύο ημιτόνων, είναι $\Delta\omega = 0,702527\text{rad/sec}$ και ίσως λίγο μεγαλύτερο. Η τιμή αυτή είναι ελάχιστα μικρότερη από αυτήν που βρέθηκε χωρίς να έχουμε κάνει zero-padding. Ωστόσο, οι τιμές αυτές είναι σχεδόν ίσες. Με δεδομένο ότι στις δύο περιπτώσεις η ελάχιστη διαφορά συχνοτήτων είναι ίδια και η απόκλιση της από αυτήν του πρώτου πειράματος είναι αμελητέα, μπορούμε να συμπεράνουμε ότι η διαδικασία του zero-padding δεν επηρεάζει τη

δυνατότητα φασματικής διάκρισης. Επομένως, χρησιμοποιούμε τη διαδικασία αυτή αν επιθυμούμε να έχουμε μεγαλύτερη ευκρίνεια φάσματος, αλλά όχι όταν θέλουμε μεγαλύτερη δυνατότητα φασματικής διάκρισης.

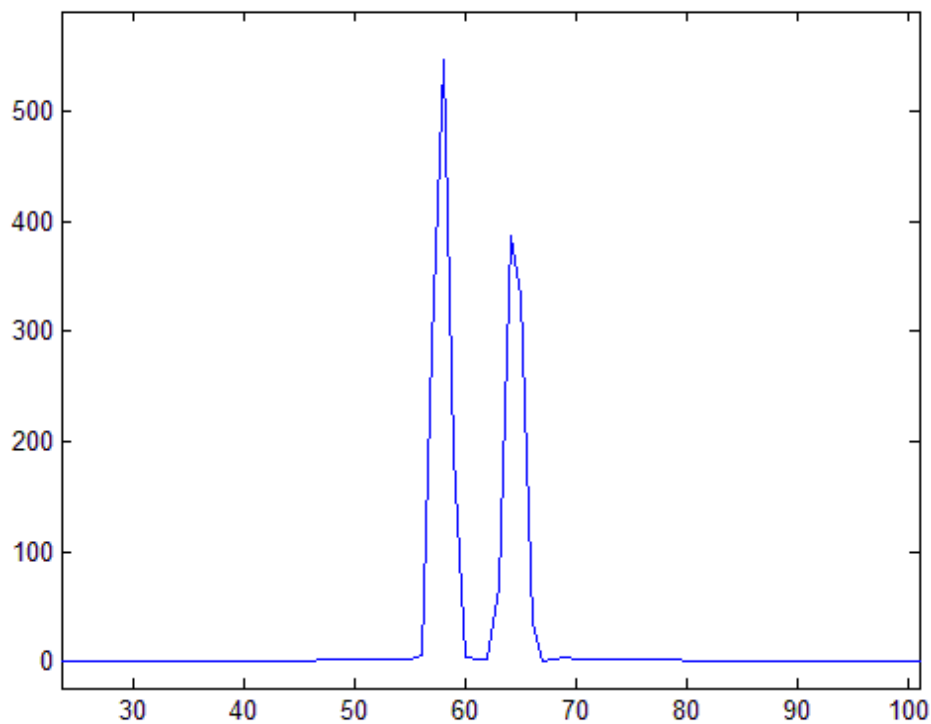
2.4.

Στο τρίτο μέρος της άσκησης θα εξεταστεί αν επηρεάζεται η δυνατότητα φασματικής διάκρισης στη περίπτωση που πάρουμε περισσότερα δείγματα από τα σήματα εισόδου για την επεξεργασία τους. Ουσιαστικά λοιπόν αυξάνουμε το μήκος του δείγματος μας. Για τη λήψη περισσότερων δειγμάτων των προς επεξεργασία σημάτων, θα χρησιμοποιήσουμε μεγαλύτερα παράθυρα. Στο ερώτημα 2.3 βρέθηκε ότι, στην περίπτωση του DFT 512 δειγμάτων, η ελάχιστη διαφορά συχνότητων που επιτρέπει τη φασματική διάκριση είναι $0,07\text{rad/sec}$. Θεωρούμε ότι το πρώτο σήμα εισόδου έχει την ίδια συχνότητα με προηγουμένως και το δεύτερο τέτοια συχνότητα ώστε η διαφορά τους να είναι αυτή που αναφέρθηκε ώστε να επιτρέπεται οριακά η φασματική διάκριση. Έτσι, εισάγοντας τα σε ένα παράθυρο 512 δειγμάτων θα εξετάσουμε κατά πόσο το πλήθος των δειγμάτων επηρεάζει τη διάκριση των επιμέρους σημάτων. Στη συνέχεια, παρατίθεται το πλάτος του DFT 512 σημείων του σήματος εξόδου του παραθύρου στην περίπτωση αυτή.



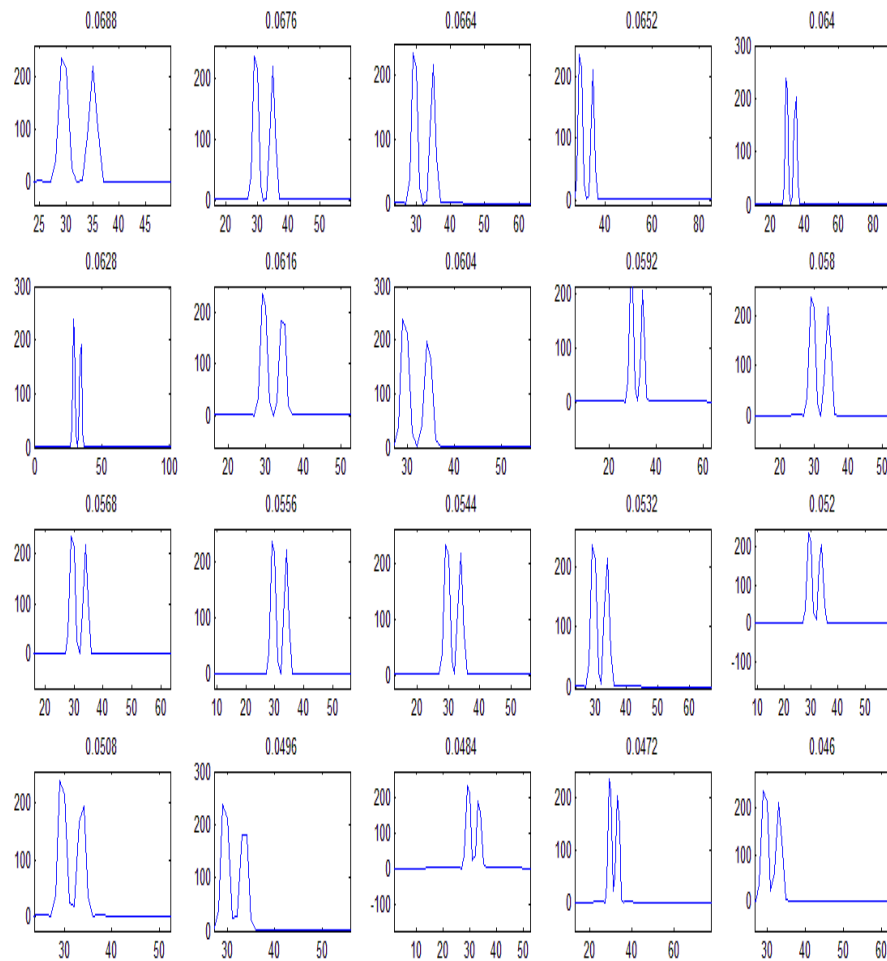
Γραφική παράσταση του πλάτους του DFT 512 σημείων του σήματος εξόδου $y[n]$ από παράθυρο 512 σημείων

Παρατηρούμε ότι πλέον η φασματική διάκριση των δύο σημάτων δεν είναι οριακή, αλλά αντίθετα φαίνονται ξεκάθαρα οι δύο κορυφές. Για παράθυρο 1024 σημείων λαμβάνουμε την ακόλουθη γραφική παράσταση:



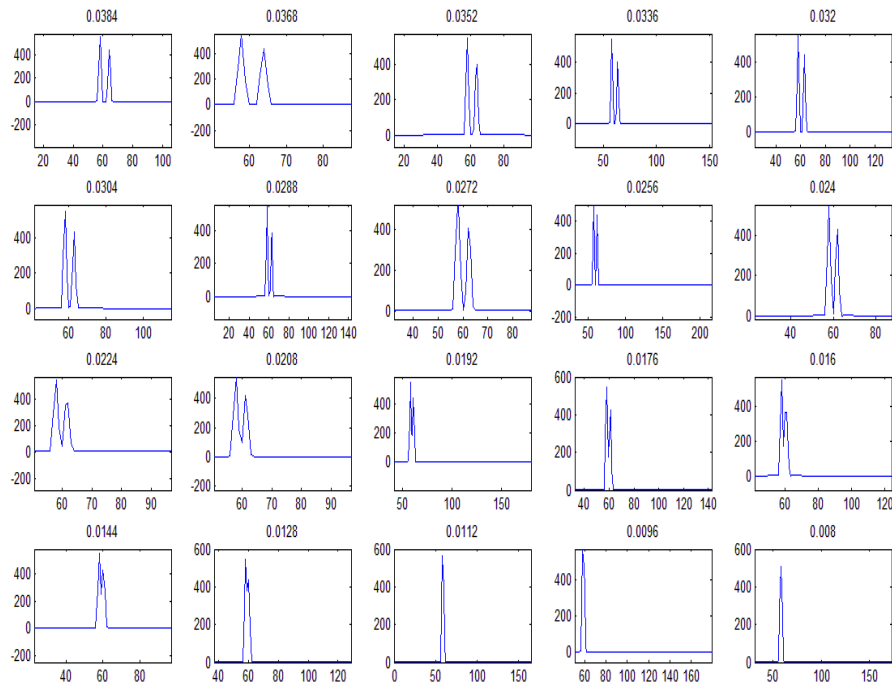
Γραφική παράσταση του πλάτους του DFT 1024 σημείων του σήματος εξόδου $y[n]$ από παράθυρο 1024 σημείων

Συγκρίνοντας τις δύο εικόνες παρατηρούμε ότι όσο μεγαλύτερος είναι ο αριθμός των δειγμάτων τόσο μεγαλύτερη είναι η ορατή διαφορά των δύο κορυφών. Για την ύπαρξη ολοκληρωμένης εικόνας επί του φαινομένου θα γίνει αναλυτική επαναληπτική διαδικασία όπως στα παραπάνω ερωτήματα με βήματα (512 δείγματα) $0,0012\text{rad/sec}$, (1024 δείγματα) $0,0016\text{rad/sec}$. Ακολουθώντας, παρατίθενται τα αντίστοιχα σύνολα γραφικών παραστάσεων που βοηθούν στον εντοπισμό της ζητούμενης διαφοράς συχνότητων.



Σύνολο γραφικών παραστάσεων του πλάτους του DFT 512 σημείων του $y[n]$ (από παράθυρο 512 δειγμάτων) για διάφορες τιμές της διαφοράς των συχνοτήτων των αρχικών σημάτων

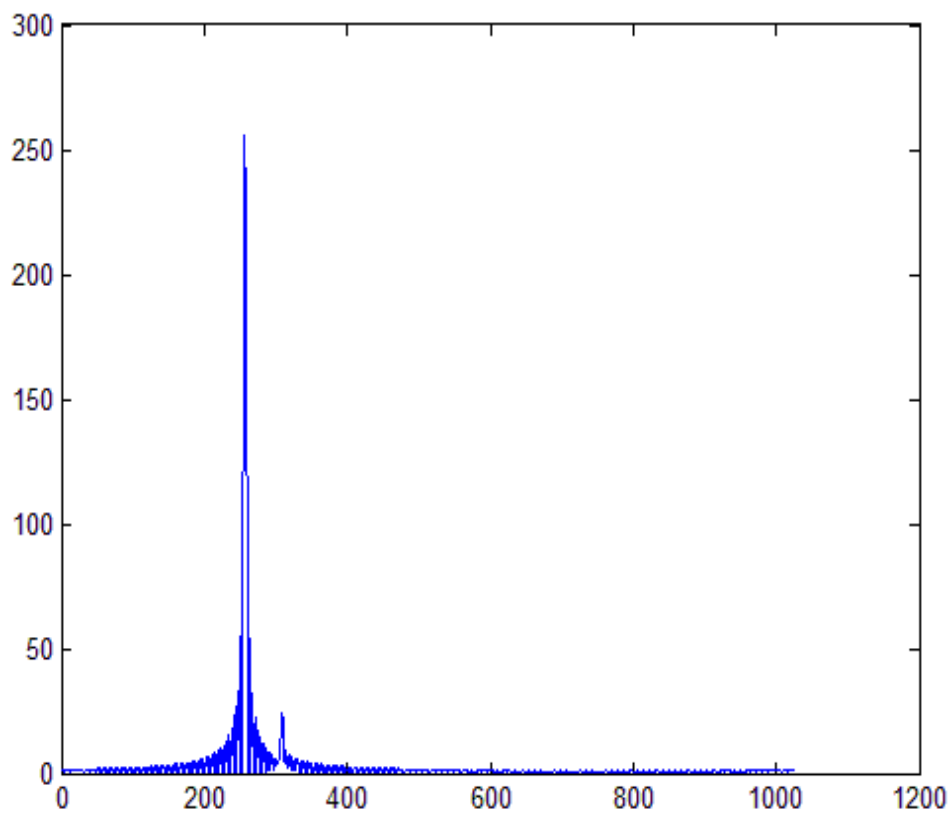
Από την παραπάνω εικόνα παρατηρούμε ότι η μικρότερη διαφορά συχνοτήτων που επιτρέπει την φασματική διάκριση είναι $\Delta\omega = 0,0652\text{rad/sec}$. Αντίστοιχα, για το σύνολο των 1024 σημείων $\Delta\omega = 0,0336\text{rad/sec}$. Σύμφωνα με τις προηγούμενες παρατηρήσεις, διαπιστώνουμε ότι αν θέλουμε να επιτύχουμε φασματική διάκριση δύο ημιτόνων για παραπλήσιες συχνότητες, τότε μπορούμε να χρησιμοποιήσουμε DFT που κατασκευάζεται με περισσότερα σημεία των σημάτων εισόδου.



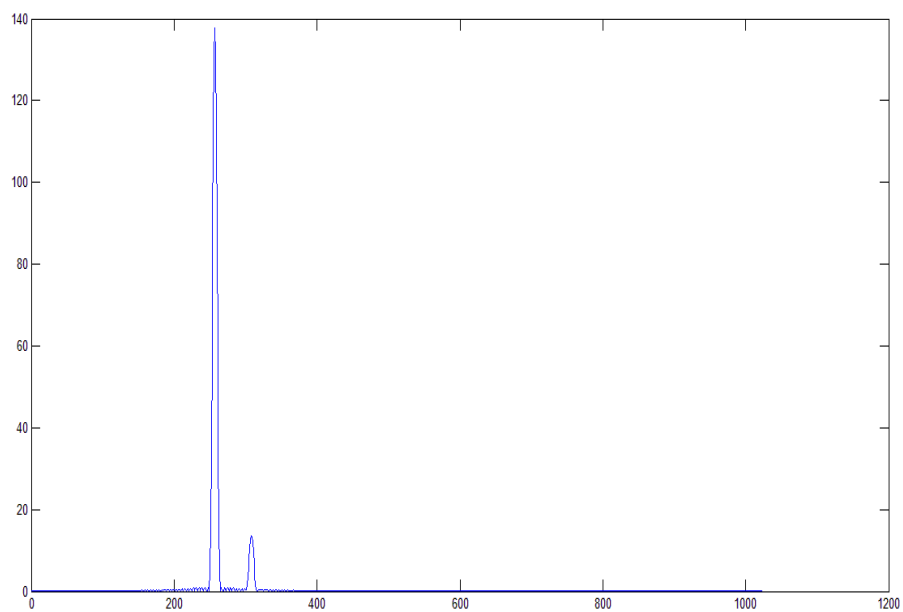
Σύνολο γραφικών παραστάσεων του πλάτους του DFT 1024 σημείων του $y[n]$ (από παράθυρο 1024 δειγμάτων) για διάφορες τιμές της διαφοράς των συχνοτήτων των αρχικών σημάτων

2.5.

Στο τελευταίο μέρος της άσκησης εξετάζεται η επίδραση του είδους του παραθύρου στο φάσμα της εξόδου. Θεωρώντας ως σήματα εισόδου τα $x_1[n]$ και $x_2[n]$ με συχνότητες $\omega_1 = 0,5\pi$ και $\omega_2 = 0,6\pi$ και πλάτη $A_1 = 1$ και $A_2 = 0,1$, αντιστοίχως, θα παρασταθεί γραφικά το πλάτος του DFT 1024 σημείων της εξόδου $y[n]$ από παράθυρα τετραγωνικό και Hamming 256 σημείων. Οι γραφικές παραστάσεις αυτές παρατίθενται στη συνέχεια.



Γραφική παράσταση του πλάτους του DFT 1024 σημείων του σήματος εξόδου $y[n]$ από τετραγωνικό παράθυρο 256 σημείων



Γραφική παράσταση του πλάτους του DFT 1024 σημείων του σήματος εξόδου $y[n]$ από hamming παράθυρο 256 σημείων

Παρατηρούμε ότι, στην περίπτωση του παραθύρου Hamming, το φάσμα της εξόδου αποτελείται ακριβώς από δύο κρουστικές στις συχνότητες των δύο ημιτόνων. Αντίθετα, στην περίπτωση του τετραγωνικού παραθύρου, παρατηρούμε κάποιους κυματισμούς του φάσματος μέχρι αυτό να μηδενιστεί. Ουσιαστικά διαπιστώνουμε την μικρή επίδραση που έχει το παράθυρο hamming στην αναπαράσταση του σήματος. Γεγονός που προκύπτει από το ότι το πλάτος του παραθύρου hamming σβήνει ομαλά και η κορύφωση βρίσκεται μόνο στον κεντρικό λοβό ενώ στο τετραγωνικό κεντρικός και πλάγιοι λοβοί έχουν ίδιο πλάτος. Συνεπώς αιτιολογούνται οι κυματισμοί που παρατηρούμε στο φάσμα του τετραγωνικού παραθύρου.

Συνοψίζοντας:

- ✚ Η διαδικασία του zero-padding βελτιώνει την ευκρίνεια του φάσματος, αλλά όχι τη διακριτική ικανότητα ενός συστήματος.
- ✚ Μεγαλύτερο μήκος παραθύρωσης βελτιώνει σημαντικά τη διακριτική ικανότητα ενός συστήματος.
- ✚ Το φάσμα της εξόδου ενός τετραγωνικού παραθύρου παρουσιάζει διακυμάνσεις σε σχέση με αυτό που προκύπτει από ένα παράθυρο Hamming.

ΜΕΡΟΣ 3^ο

Χαρακτηριστικά Βραχέος Χρόνου Σημάτων Φωνής (Ενέργεια και Ρυθμός Εναλλαγής Προσήμου)

Οι μετρήσεις βραχέος χρόνου, οι οποίες πραγματοποιούνται σε ένα μετακινούμενο παράθυρο του σήματος. Ορίζονται η ενέργεια βραχέος χρόνου:

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

Πρέπει να σημειωθεί λόγω του καθοριστικού ρόλου που έπαιξε στην επίλυση της άσκησης ότι ουσιαστικά είναι η συνέλιξη των επιμέρους τετραγώνων των δύο σημάτων πιο συγκεκριμένα το παρπάνω είναι ουσιαστικά το συνελικτικό άθροισμα.

Αντίστοιχα ο ρυθμός εναλλαγής προσήμου ορίζεται από τον τύπο:

$$Z_n = \sum_{m=-\infty}^{+\infty} |\operatorname{sgn}[x(m)] - \operatorname{sgn}[x(m-1)]| \cdot w(n-m)$$

Δηλαδή η συνέλιξη που αποδίδεται παρακάτω

$$Z_n = |\operatorname{sgn}[x(n)] - \operatorname{sgn}[x(n-1)]| * w(n)$$

Γνωρίζουμε πως η συνέλιξη στο πεδίο του χρόνου μας δίνει γινόμενο στο πεδίο του dft. Στην ουσία θα πάρουμε τον διακριτό μετασχηματισμό Fourier κάνοντας χρήση της συνάρτησης fft του Matlab και αφού βρούμε το αποτέλεσμα που θέλουμε σε μορφή DFT θα πάρουμε τον αντίστροφο μετασχηματισμό DFT (ifft) εφαρμόζοντας τας κατάλληλο zero padding ώστε το σήμα να φτάσει στο αρχικό του μήκος μειωμένο κατά μία θέση.

3.1.

Οι παραπάνω τύποι θα εφαρμοστούν μέσω του Matlab στο σήμα φωνής **speech_utterance.wav** που μας δίνεται. Θα υπολογιστούν τα παραπάνω μεγέθη για δειγματοληψία μήκη παραθύρων και με τη βοήθεια των διαγραμμάτων και των μεγεθών αυτών θα εντοπίσουμε έμφωνους και άφωνους ήχους.

Διαβάζουμε το δοσμένο αρχείο (**wave read()**) με δοσμένη συχνότητα δειγματοληψίας ίση με 16kHz. Επιθυμούμε το παράθυρο Hamming που θα επιλέξουμε να έχει δείγματα τα οποία θα έχουν προέλθει από τη συνάρτηση συνεχούς χρόνου του Hamming με ίδια συχνότητα δειγματοληψίας με αυτή του αρχείου ήχου. Επομένως έτσι βρίσκουμε την περίοδο δειγματοληψίας

$$T_s = \frac{1}{f_s} \text{ άρα } T_s = 0.0625 \text{ msec} \text{ Έπειτα διαιρούμε αυτό το μέγεθος με το μήκος}$$

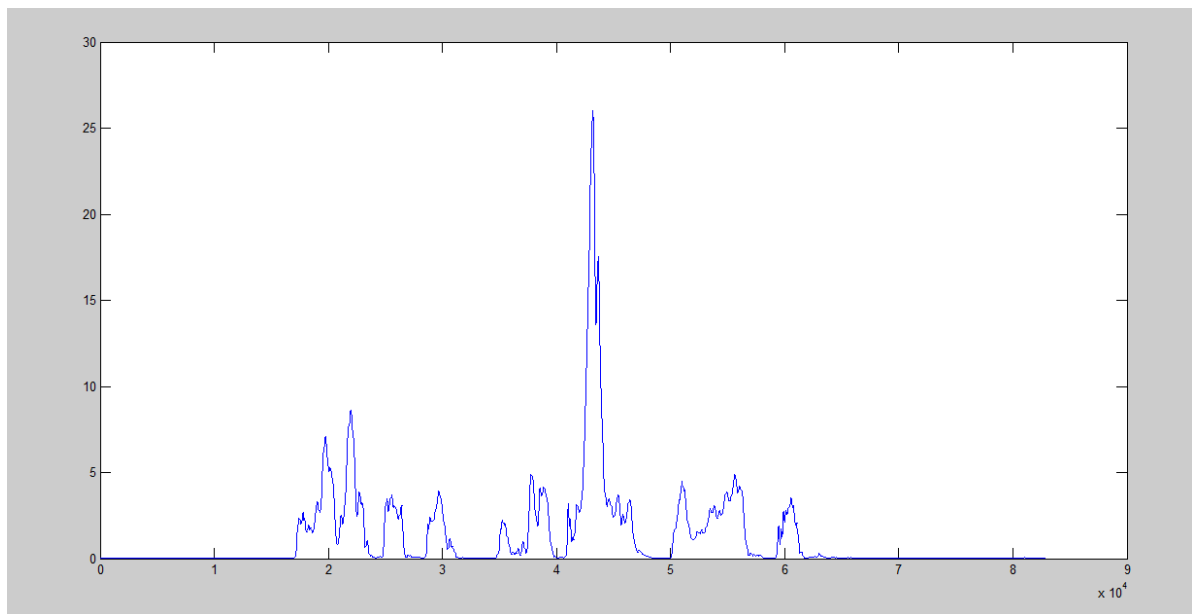
παραθύρου hamming βρίσκουμε το επιθυμητό πλήθος δειγμάτων.

Για τον λόγο αυτό γράφουμε μία συνάρτηση για την εύρεση της ενέργειας βραχέος χρόνου και άλλη μία για τον υπολογισμό του ρυθμού εναλλαγής προσήμου (zero crossing rate).

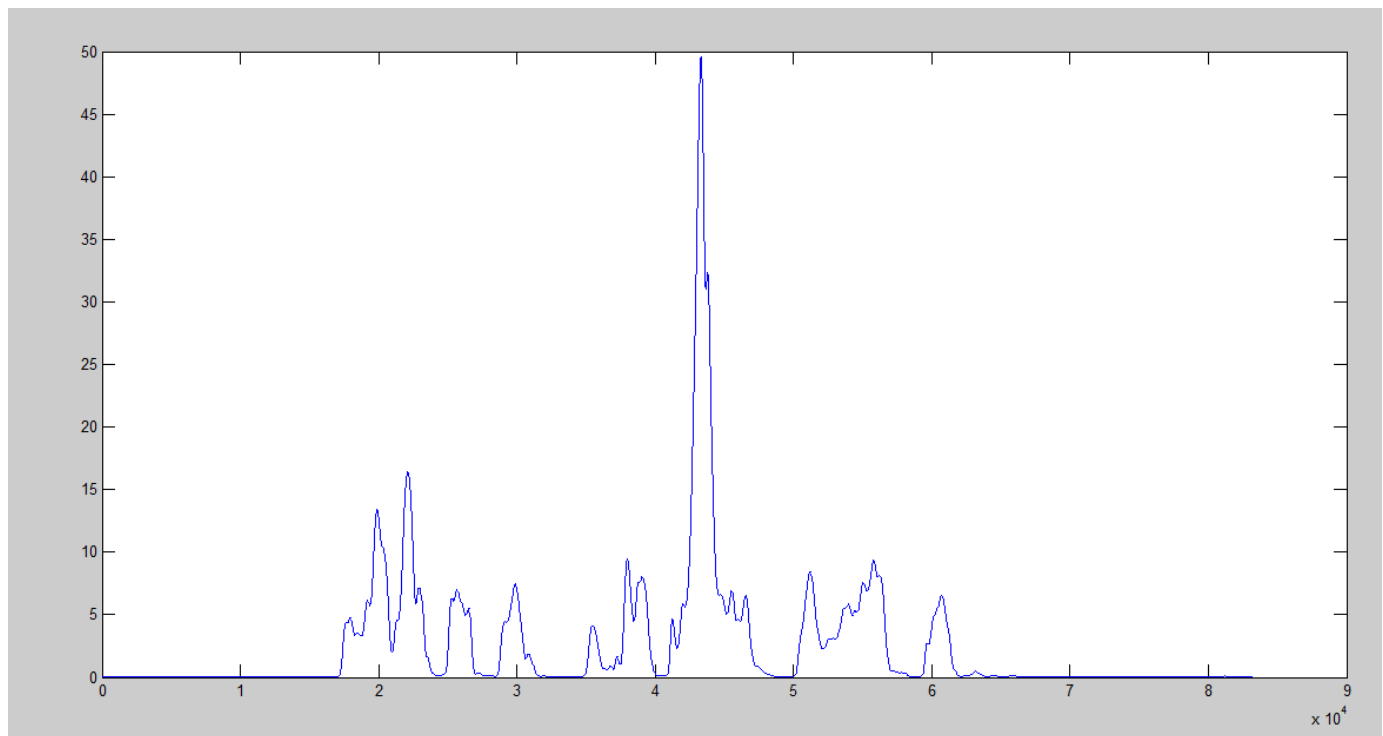
Οι συναρτήσεις καλούνται για τα διάφορα σε μήκος παράθυρα που ζητούνται στην άσκηση. Παρακάτω παρουσιάζονται οι ανάλογες γραφικές παραστάσεις:

Γραφική παράσταση ενέργειας βραχέος χρόνου για:

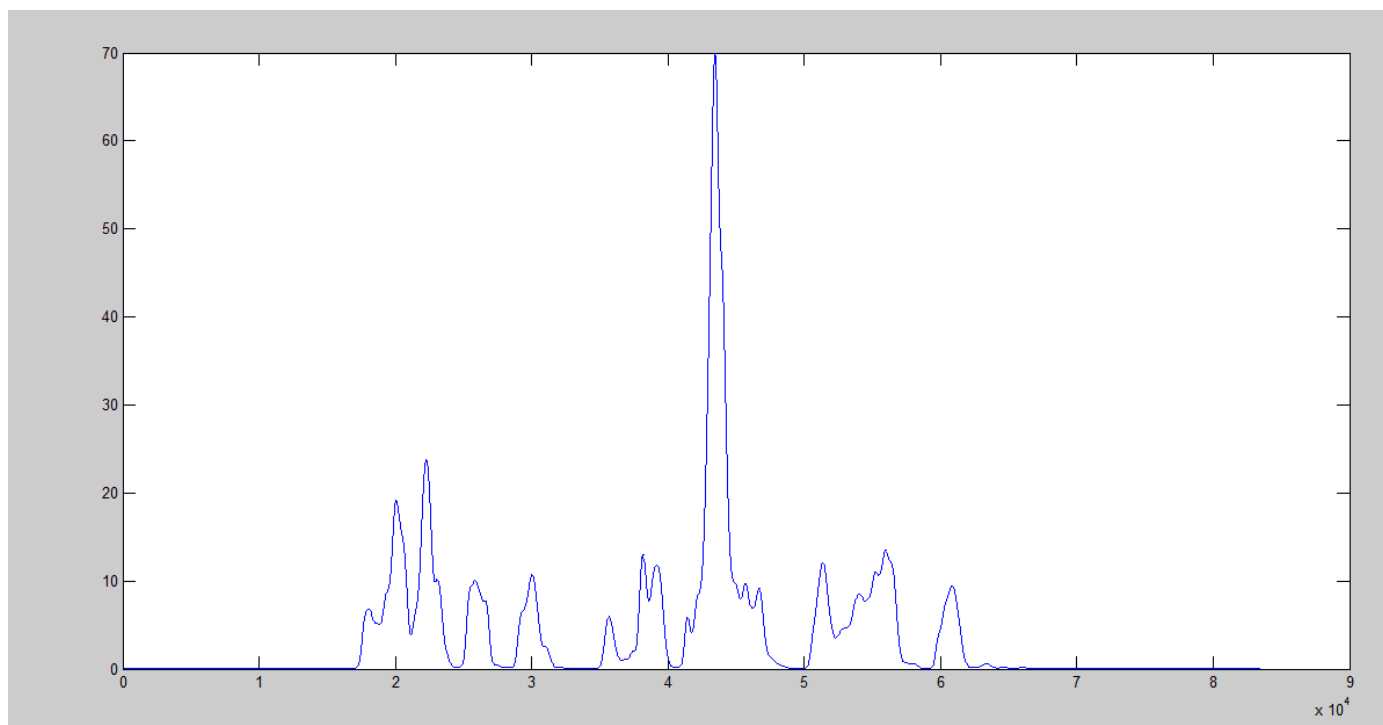
Παράθυρο Hamming διάρκειας 20msec:



Παράθυρο Hamming διάρκειας 40msec:

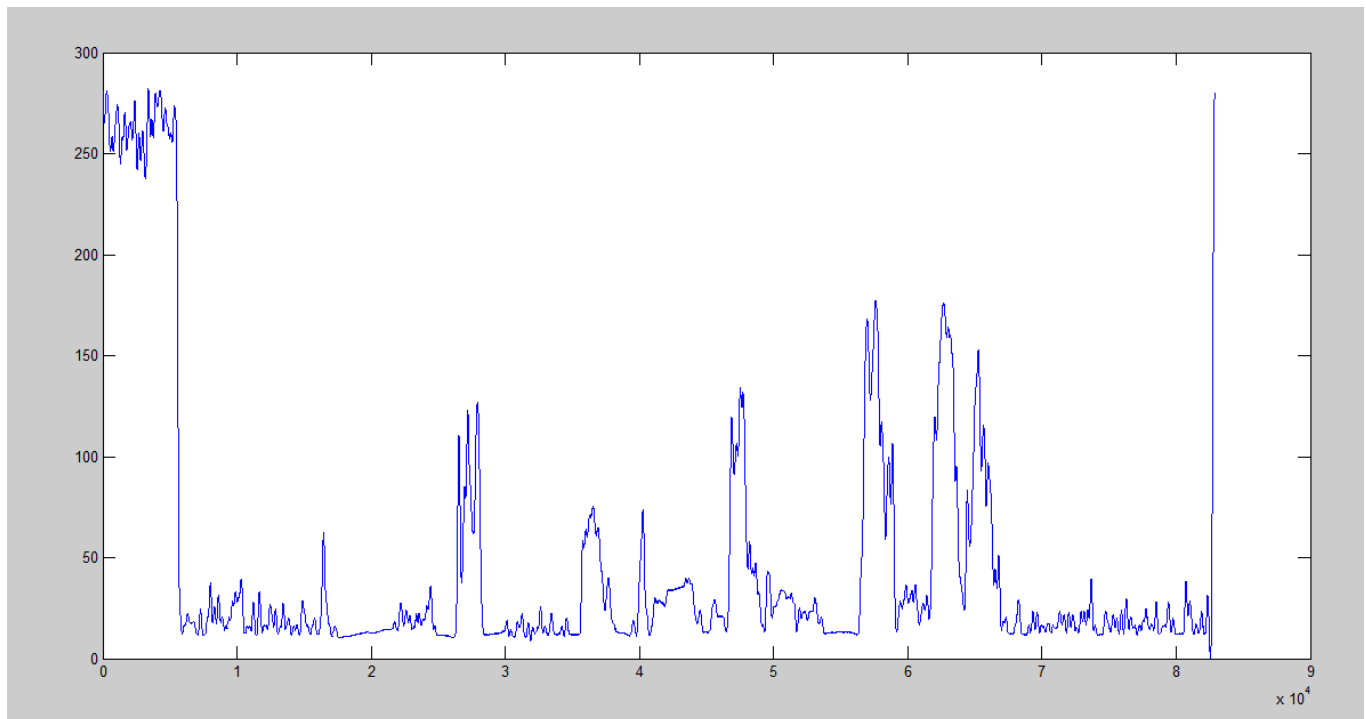


Παράθυρο Hamming διάρκειας 60msec:

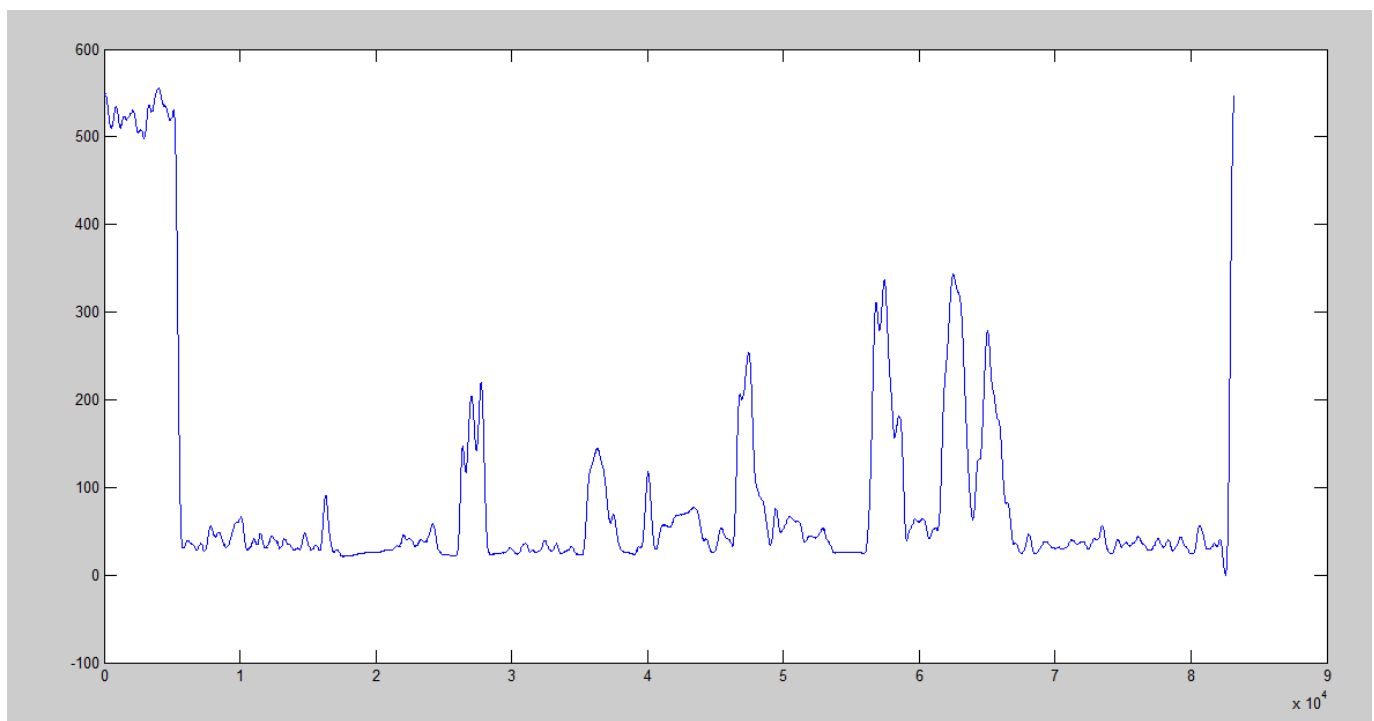


Γραφική παράσταση zero crossing rate για:

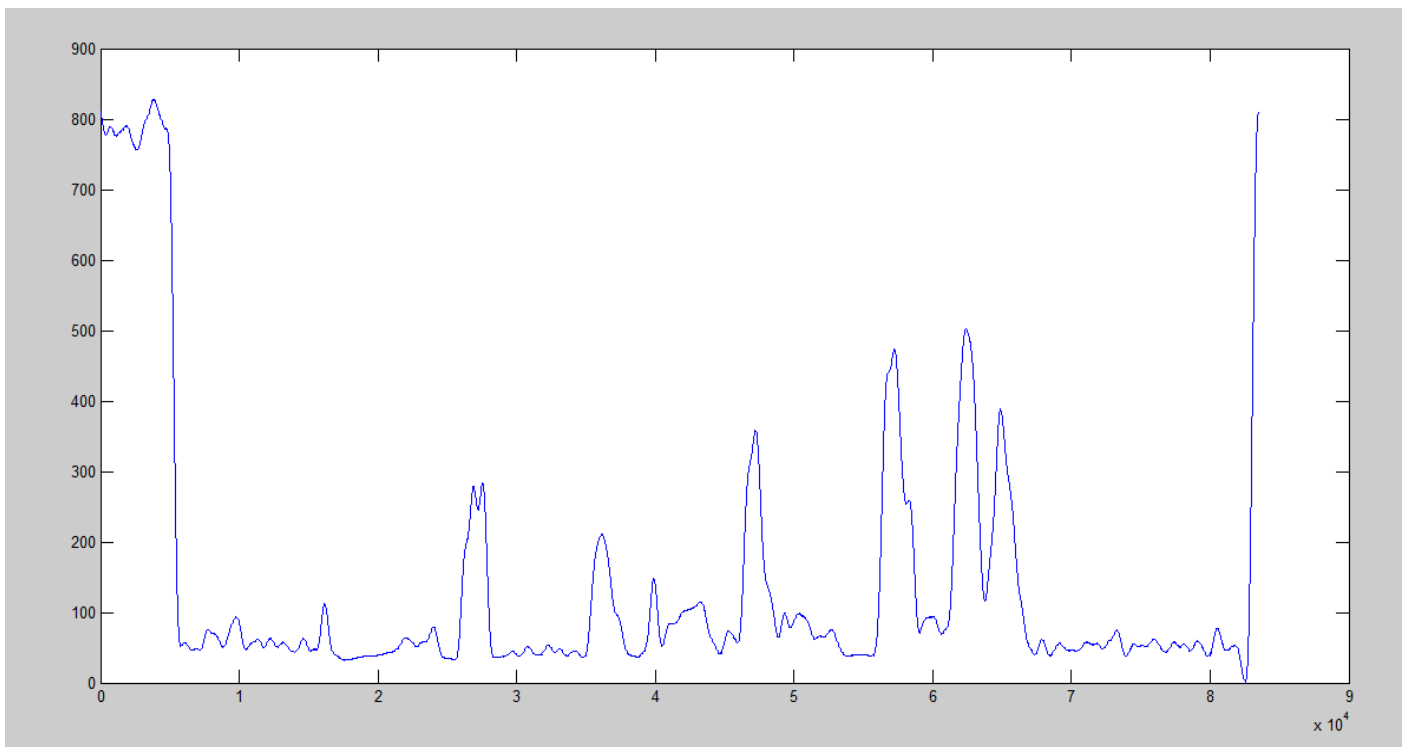
Παράθυρο Hamming διάρκειας 20msec:



Παράθυρο Hamming διάρκειας 40msec:



Παράθυρο Hamming διάρκειας 60msec:



Παρατηρώντας τα παραπάνω διαγράμματα συμπεραίνουμε τα εξής:

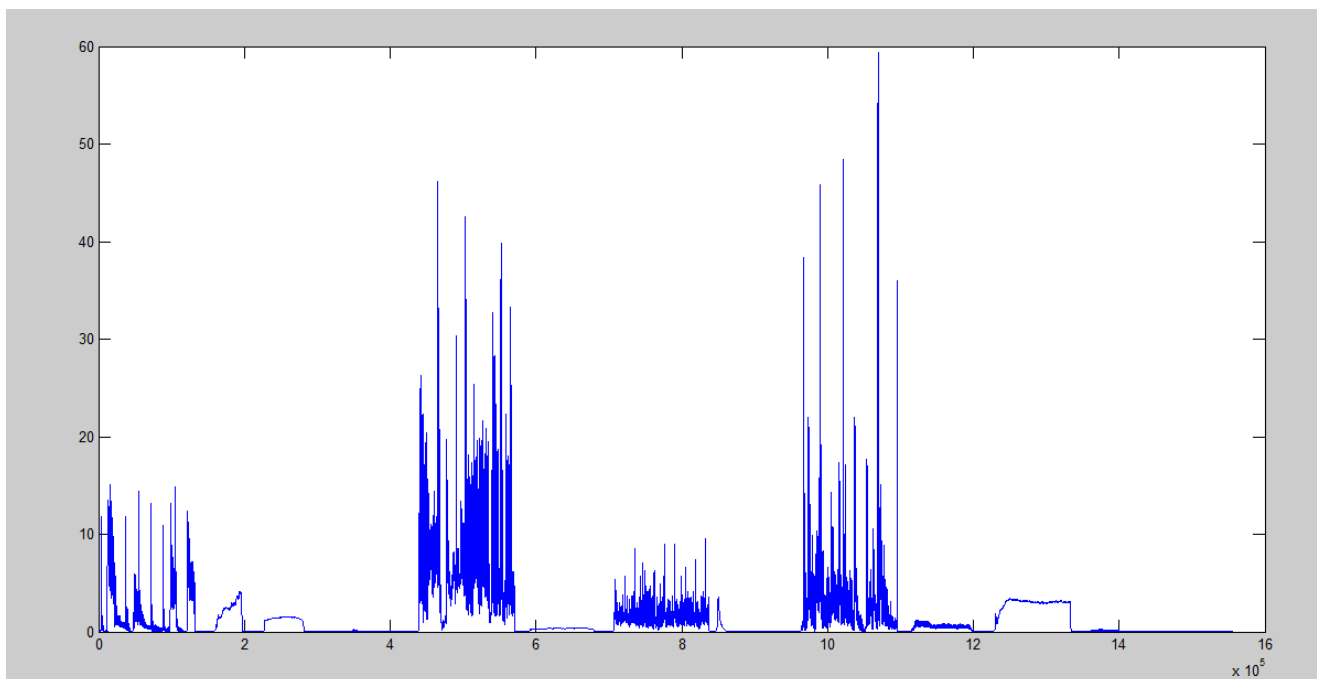
- ✓ Αρχικά παρατηρούμε ότι τα μεγαλύτερα παράθυρα έχουν πιο ομαλοποιημένες γραφικές και των δύο μεγεθών. Στα διαγράμματα των μικρότερων παραθύρων παρατηρούνται εντονότερες μεταβολές και αυξομειώσεις.
- ✓ Επίσης δαιφωτιζόμαστε σχετικά με τους έμφωνους και άφωνους ήχους. Κατά την παρατήρηση έμφωνων ήχων π.χ. αα έχουν άυξηση της ενέργειας βραχέος χρόνου αντιθέτως έχουμε μείωση για τους άφωνους ήχους άρα χαμηλή ενέργεια βραχέος χρόνου. Στα σημεία στα οποία η ενέργεια βραχέος χρόνου είναι μηδενική ή ο zero crossing rate είναι αρκετά αυξημένος έχουμε σιγή.

Ερώτημα 3.2.:

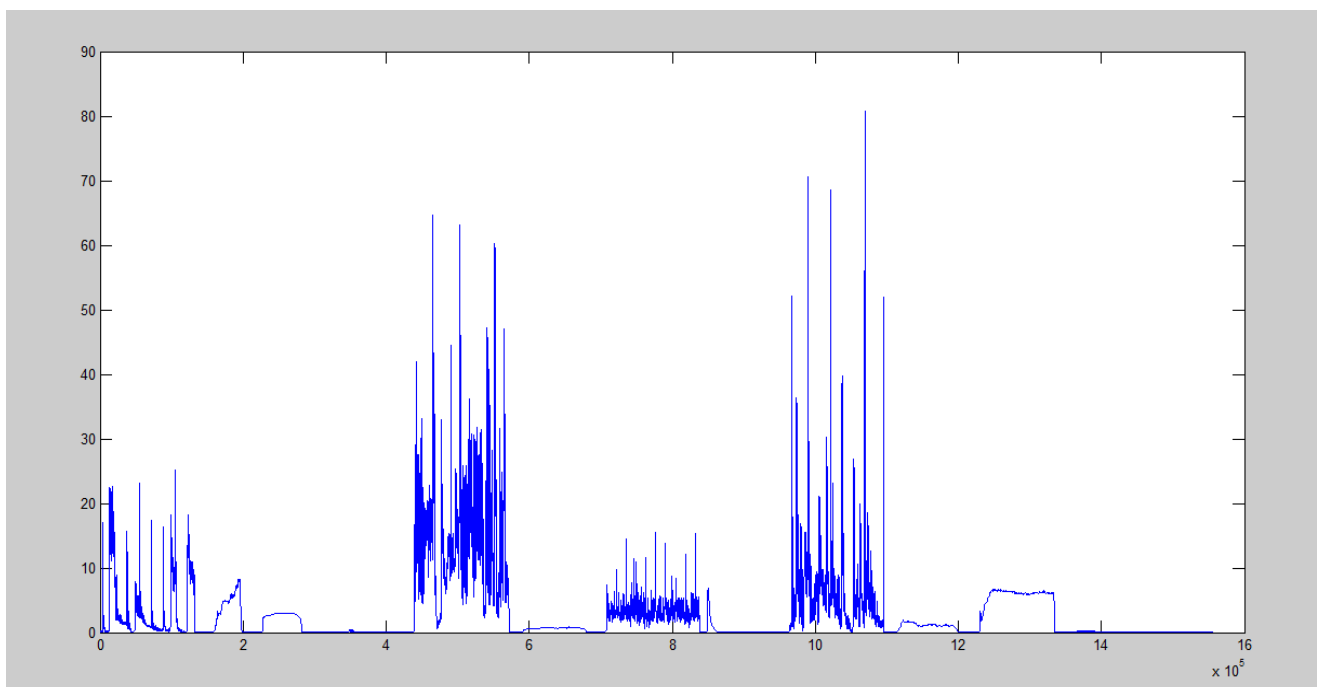
Την ίδια διαδικασία με παραπάνω ακολουθούμε αλλά αυτή τη φορά επεξεργαζόμαστε το σήμα **music.wav** που μας δίνεται. Παρακάτω δίδονται οι αντίστοιχες γραφικές παραστάσεις όπως προέκυψαν από το Matlab για τα διάφορα μήκη παραθύρων.

Γραφική παράσταση ενέργειας βραχέος χρόνου για:

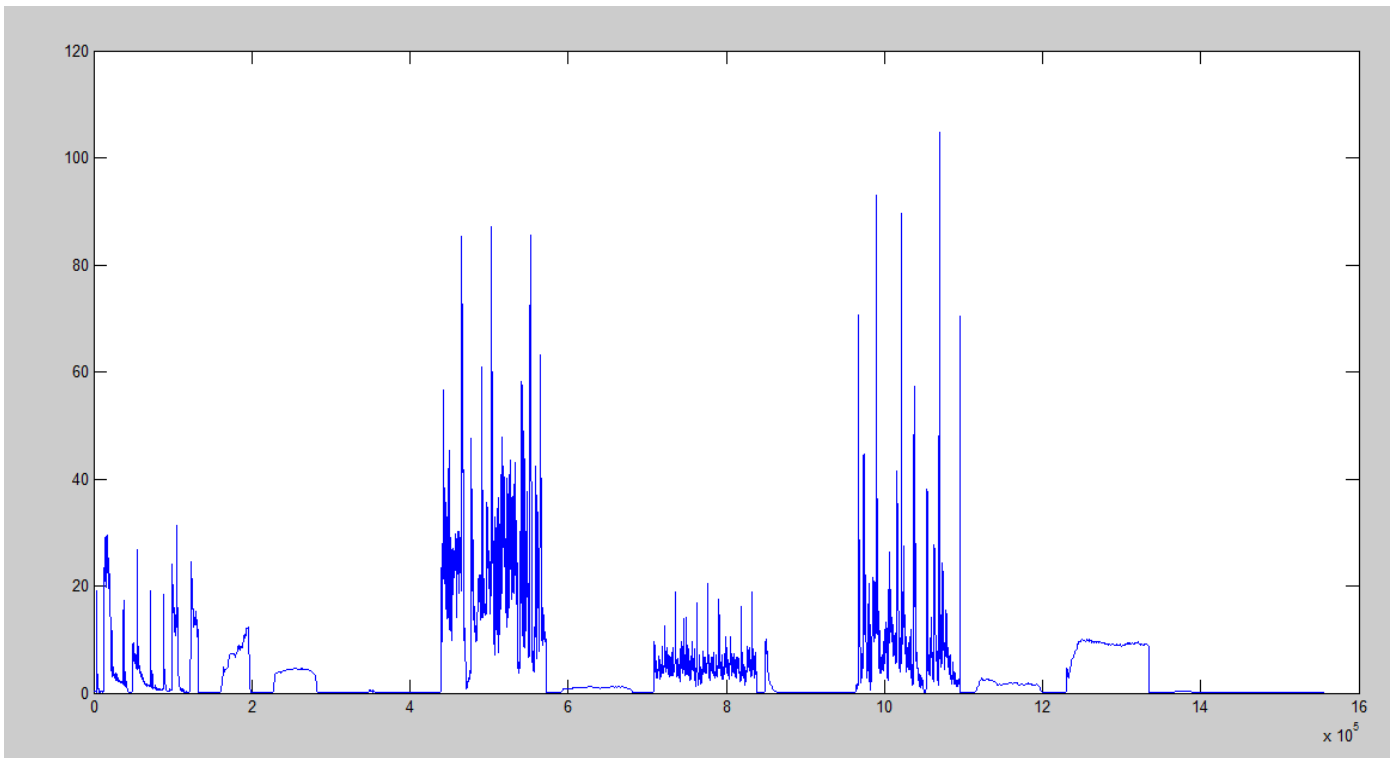
Παράθυρο Hamming διάρκειας 20msec



- **Παράθυρο Hamming διάρκειας 40msec**

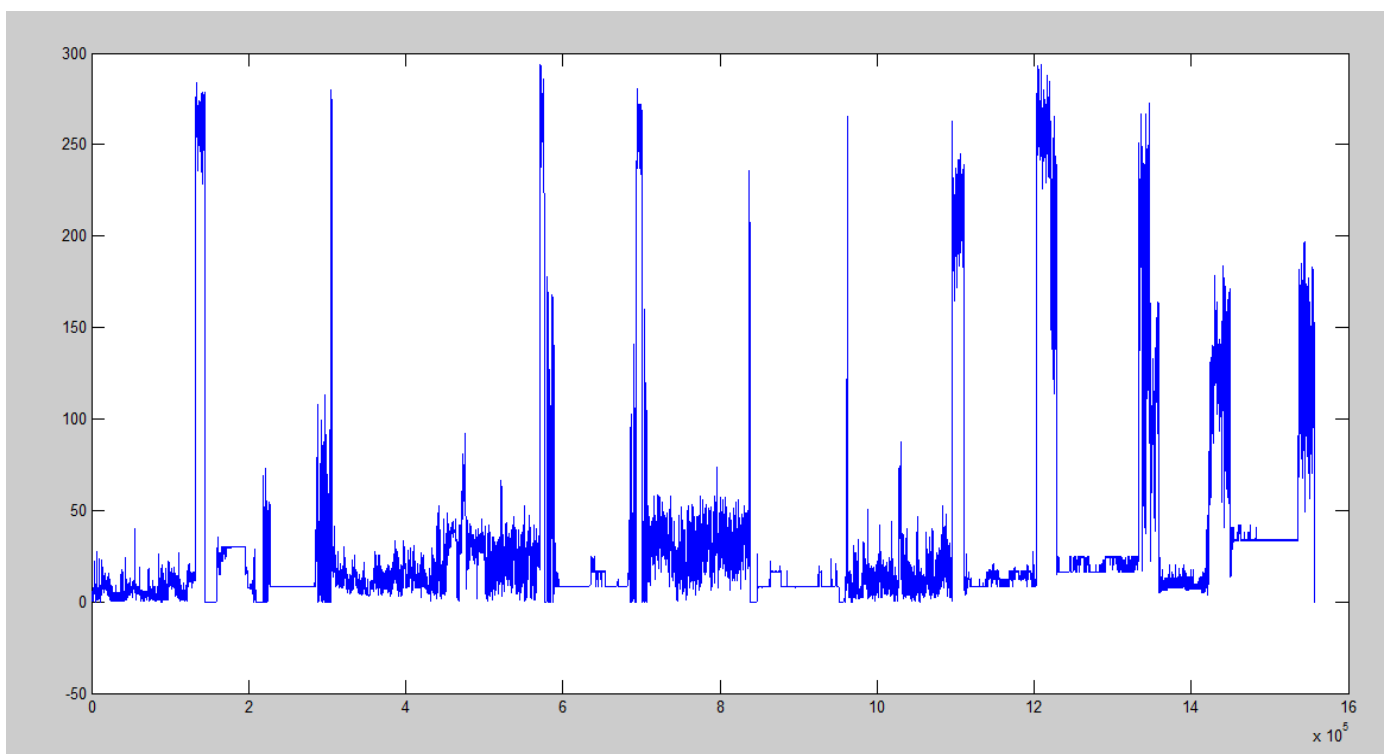


Παράθυρο Hamming διάρκειας 60msec:

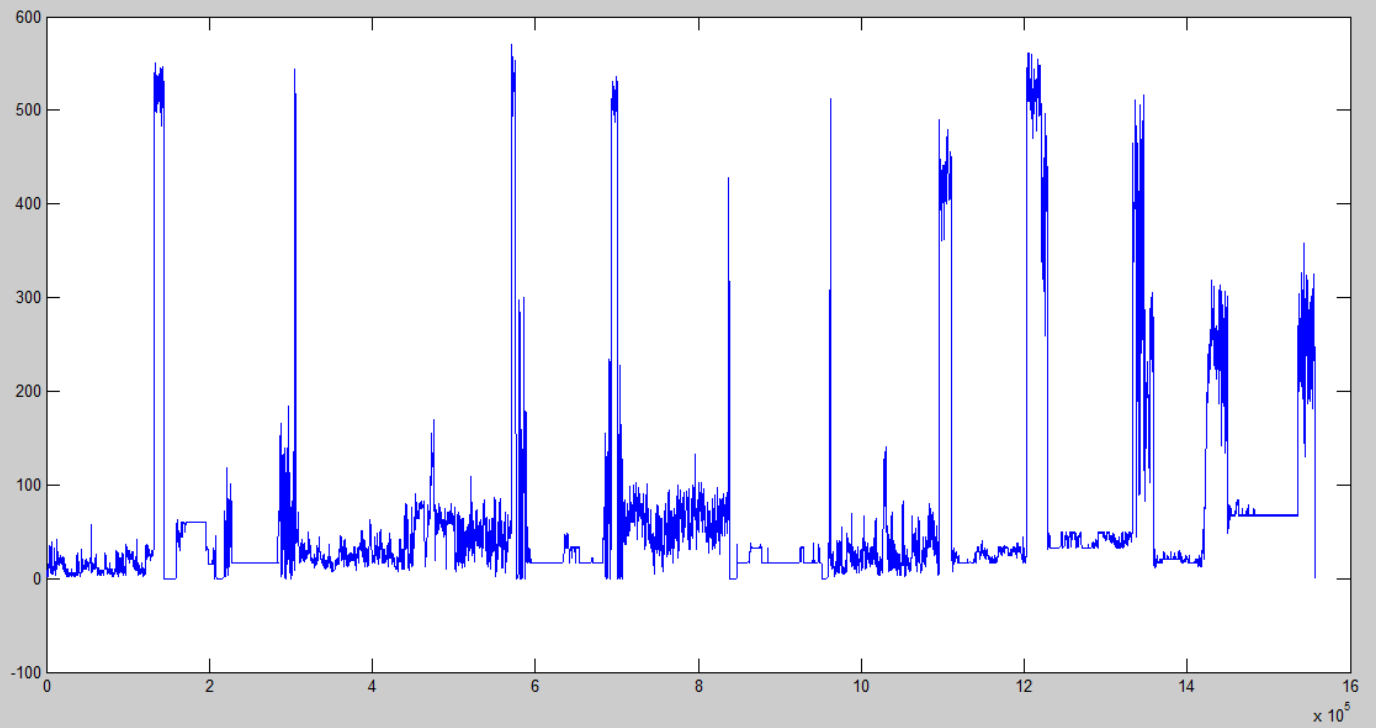


Γραφική παράσταση ρυθμού εναλλαγής προσήμου για:

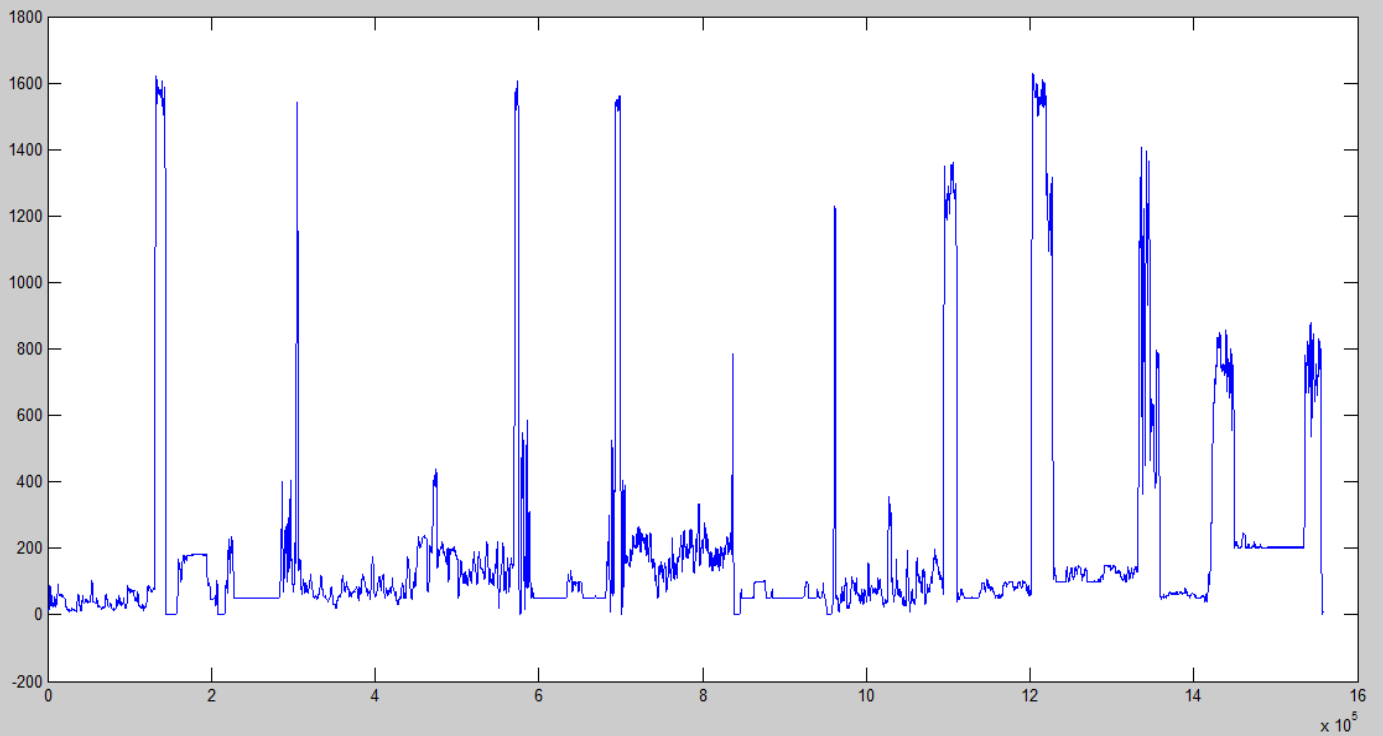
Παράθυρο Hamming διάρκειας 20msec:



Παράθυρο Hamming διάρκειας 40msec:



Παράθυρο Hamming διάρκειας 60msec:



Παρατηρούνται τα ίδια αποτελέσματα με το προηγούμενο αρχείο απλώς αυτό που αλλάζει στο παρόν είναι ότι το αρχείο είναι μεγαλύτερου μεγέθους και τα προηγούμενα συμπεράσματα είναι ορατά σε μεγαλύτερο φάσμα και εύρος παρατήρησης στο συχνοτικό πεδίο.