

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΠΕΙΡΑΙΑ



Τ.Ε.Ι ΠΕΙΡΑΙΑ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΑΥΤΟΜΑΤΙΣΜΟΥ

ΑΝΑΓΝΩΡΙΣΗ ΦΩΝΗΤΙΚΩΝ ΕΝΤΟΛΩΝ ΜΕΣΩ ΜΑΤΛΑΒ

Όνομα Επιβλέποντα Καθηγητή
Δ.Τσελές

Έξεταστες: Χ.Δρόσος, Μ.Παπουτσιδάκης

Φοιτητές:
Αγγελάκης Επιμενίδης
Πολυχρονόπουλος Γεώργιος
Γατής Θεόδωρος

ΜΑΡΤΙΟΣ 2016

ΤΜΗΜΑ ΑΥΤΟΜΑΤΙΣΜΟΥ
Π. Ράλλη & Θηβών 250, 12244 Αιγάλεω, Αθήνα – Ελλάδα
Τηλ. 210-5381488

Ευχαριστίες

Με την ολοκλήρωση της πτυχιακής μας εργασίας θα θέλαμε να ευχαριστήσουμε όλους όσους μας βοήθησαν για την πολύτιμη στήριξή τους. Κυρίως οφείλουμε να ευχαριστήσουμε τον επιβλέποντα καθηγητή μας κ. ο οποίος μας στήριξε και για την καθοδήγησή του καθ' όλη τη διάρκεια της πτυχιακής εργασίας. Τέλος, οφείλουμε να αφιερώσουμε την πτυχιακή μας εργασία στις οικογένειες μας και σε όλους εκείνους που ήταν δίπλα μας σε όλη αυτή την προσπάθεια μας αλλά και όλα τα χρόνια της φοίτησής μας παρέχοντας μας στήριξη και απεριόριστη κατανόηση.

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Ο/Η κάτωθι υπογεγραμμένος/η Γαζής Θεόδωρος
ΕΠΙΧΕΙΡΙΣΗΣ ΑΣΦΕΛΑΚΗΣ
Πολυτεχνολογικός Γεωργίου,
του Γεωργίου, με αριθμό μητρώου 39464 φοιτητής / τρια του
Τμήματος Μηχανικών Αυτοματισμού Τ.Ε. του Α.Ε.Ι. Πειραιά Τ.Τ. πριν αναλάβω την
εκπόνηση της Πτυχιακής Εργασίας μου, δηλώνω ότι ενημερώθηκα για τα παρακάτω:

«Η Πτυχιακή Εργασία (Π.Ε.) αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο του συγγραφέα, όσο και του Ιδρύματος και θα πρέπει να έχει μοναδικό χαρακτήρα και πρωτότυπο περιεχόμενο.

Απαγορεύεται αυστηρά οποιοδήποτε κομμάτι κειμένου της να εμφανίζεται αυτούσιο ή μεταφρασμένο από κάποια άλλη δημοσιευμένη πηγή. Κάθε τέτοια πράξη αποτελεί προϊόν λογοκλοπής και εγείρει θέμα Ηθικής Τάξης για τα πνευματικά δικαιώματα του άλλου συγγραφέα. Αποκλειστικός υπεύθυνος είναι ο συγγραφέας της Π.Ε., ο οποίος φέρει και την ευθύνη των συνεπειών, ποινικών και άλλων, αυτής της πράξης.

Πέραν των όποιων ποινικών ευθυνών του συγγραφέα σε περίπτωση που το Ίδρυμα του έχει απονεμίσει Πτυχίο, αυτό ανακαλείται με απόφαση της Συνέλευσης του Τμήματος. Η Συνέλευση του Τμήματος με νέα απόφαση της, μετά από αίτηση του ενδιαφερόμενου, του αναθέτει εκ νέου την εκπόνηση της Π.Ε. με άλλο θέμα και διαφορετικό επιβλέποντα καθηγητή. Η εκπόνηση της εν λόγω Π.Ε. πρέπει να ολοκληρωθεί εντός τουλάχιστον ενός ημερολογιακού βμήνου από την ημερομηνία ανάθεσης της. Κατά τα λοιπά εφαρμόζονται τα προβλεπόμενα στο άρθρο 18, παρ. 5 του ισχύοντος Εσωτερικού Κανονισμού.»

Ο Δηλών

~~Γαζής~~
~~Ασφελής~~
~~Γεωργίου~~

Ημερομηνία

32 / 03 / 2016

Περίληψη

Ο σκοπός αυτής της διπλωματικής εργασίας είναι η διερεύνηση των αλγορίθμων αναγνώρισης φωνής. Έγινε προγραμματισμός και προσομοίωση των σχεδιασμένων συστημάτων για αλγόριθμους αναγνώρισης φωνής στο MATLAB . Σε αυτή την εργασία σχεδιάστηκαν δύο συστήματα. Το ένα βασίζεται στις πληροφορίες σχήματος του γραφήματος συσχέτισης. Το άλλο χρησιμοποιεί το φίλτρο Wiener για να πραγματοποιήσει την αναγνώριση ομιλίας. Οι προσομοιώσεις των προγραμματισμένων συστημάτων στο MATLAB επιτυγχάνονται χρησιμοποιώντας το μικρόφωνο για να καταγραφούν οι λέξεις. Μετά την εκτέλεση του προγράμματος στο MATLAB , το MATLAB θα ζητήσει από τους ανθρώπους να καταγράψει τις λέξεις τρεις φορές . Η πρώτη και η δεύτερη ηχογραφημένη λέξη είναι διαφορετικές λέξεις που θα χρησιμοποιηθούν ως σήματα αναφοράς στα σχεδιασμένα συστήματα. Η τρίτη ηχογραφημένη λέξη είναι η ίδια λέξη με τη μία από τις πρώτες δύο ηχογραφημένες λέξεις. Μετά την εγγραφή των λέξεων, οι λέξεις θα γίνουν σήματα, πληροφορίες οι οποίες θα γίνουν δειγματοληψία και θα αποθηκευθούν στο MATLAB . Τότε το MATLAB πρέπει να είναι σε θέση να δώσει την απάντηση, όποια λέξη καταγραφεί την τρίτη φορά σε σχέση με τις δύο πρώτες λέξεις σύμφωνα με τους αλγόριθμους που έχουν προγραμματιστεί στο MATLAB. Διαφορετικοί άνθρωποι καλέστηκαν από διαφορετικές χώρες για να δοκιμάσουν τα σχεδιασμένα συστήματα. Τα αποτελέσματα των προσομοιώσεων για τα δύο σχεδιασμένα συστήματα δείχνουν ότι και τα δύο σχεδιασμένα συστήματα λειτουργούν καλά όταν οι δύο πρώτες ηχογραφήσεις και η τρίτη φορά ηχογράφησης καταγράφονται από το ίδιο πρόσωπο. Αλλά τα σχεδιασμένα συστήματα έχουν κάποια ελαττώματα , όταν οι δύο πρώτες ηχογραφήσεις και η τρίτη φορά ηχογράφησης καταγράφονται από διαφορετικούς ανθρώπους. Ωστόσο , εάν το περιβάλλον δοκιμών είναι αρκετά ήσυχο και ο ομιλητής είναι το ίδιο πρόσωπο για τρεις χρονικές ηχογραφήσεις, η πιθανότητα επιτυχίας της αναγνώρισης φωνής προσεγγίζεται στο 100%. Έτσι, τα σχεδιασμένα συστήματα πράγματι λειτουργούν καλά για την βασική αναγνώριση ομιλίας

Πίνακας Περιεχομένων

Ευχαριστίες	2
Περίληψη	3
Πίνακας Περιεχομένων	4
1° Κεφάλαιο Εισαγωγή	6
1.1 Εισαγωγή	6
1.2 Κίνητρο για την διεξαγωγή της εργασίας	6
2° Κεφάλαιο Θεωρία	8
2.1 Επίπεδο DC και θεωρία δειγματοληψίας	8
2.2 Πεδίο χρόνου σε πεδίο συχνότητας: DFT και FFT	10
2.2.1 DFT	10
2.2.2 FFT	13
2.3 Ανάλυση συχνότητας στο MATLAB για την Αναγνώριση Φωνής	15
2.3.1 Κανονικοποίηση Φάσματος	15
2.3.2 Ο αλγόριθμος συσχέτισης	17
2.3.3 Ο αλγόριθμος αυτοσυσχέτισης	21
2.3.4 Το φίλτρο FIR Wiener	22
2.3.5 Χρήση Φασματογραφήματος στο MATLAB	25
3° Κεφάλαιο Βήματα προγραμματισμού και αποτελέσματα της προσομοίωσης	33
3.1 Βήματα προγραμματισμού	33
3.1.1 Βήματα προγραμματισμού για Σχεδιασμένο Σύστημα 1	33
3.1.2 Βήματα προγραμματισμού για Σχεδιασμένο Σύστημα 2	34
3.2 Αποτελέσματα προσομοίωσης	35
3.2.1 Αποτελέσματα προσομοίωσης για το σύστημα 1	36
3.2.2 Αποτελέσματα προσομοίωσης για το σύστημα 2	45
4° Κεφάλαιο Συζήτηση και Συμπεράσματα	51
4.1 Συζήτηση	51
4.1.1 Συζήτηση για τα αποτελέσματα της προσομοίωσης του συστήματος 1	51
4.1.2 Συζήτηση για τα αποτελέσματα της προσομοίωσης του συστήματος 2	52
4.2 Συμπεράσματα	53
Αναφορές	54
Παράρτημα Α	55
Παράρτημα Β	64

1° Κεφάλαιο

Εισαγωγή

1.1 Εισαγωγή

Η αναγνώριση φωνής είναι ένα δημοφιλές θέμα στη σημερινή ζωή. Οι εφαρμογές της αναγνώρισης φωνής μπορεί να βρεθεί παντού , και οι οποίες κάνουν πιο αποτελεσματική τη ζωή μας . Για παράδειγμα, οι εφαρμογές στο κινητό τηλέφωνο , αντί να πληκτρολογήσουμε το όνομα του προσώπου που θέλουμε να καλέσουμε , μπορούμε απλά να πούμε απευθείας το όνομα του προσώπου στο κινητό τηλέφωνο και το κινητό τηλέφωνο θα καλέσει αυτόματα το εν λόγω πρόσωπο . Αν θέλουμε να στείλουμε μερικά μηνύματα κειμένου σε κάποιον, οι μπορούμε επίσης να πούμε μηνύματα στο κινητό τηλέφωνο, αντί να πληκτρολογήσουμε. Η αναγνώριση φωνής είναι μια τεχνολογία που οι άνθρωποι μπορούν να ελέγξουν το σύστημα με την ομιλία τους. Αντί να πληκτρολογήσουμε στο πληκτρολόγιο ή να χρησιμοποιήσουμε τα κουμπιά για το σύστημα , χρησιμοποιώντας την ομιλία για να ελέγξουμε το σύστημα είναι πιο βολικό . Μπορεί επίσης να μειώσει το κόστος της παραγωγής της βιομηχανίας την ίδια στιγμή. Χρησιμοποιώντας το σύστημα αναγνώρισης φωνής όχι μόνο βελτιώνει την αποτελεσματικότητα της καθημερινής ζωής, αλλά και κάνει τη ζωή των ανθρώπων πιο διαφοροποιημένη .

1.2 Κίνητρο για την διεξαγωγή της εργασίας

Σε γενικές γραμμές , ο στόχος της παρούσας διατριβής είναι η διερεύνηση των αλγορίθμων αναγνώρισης φωνής με τον προγραμματισμό και την προσομοίωση του σχεδιασμένου συστήματος στο MATLAB . Την ίδια στιγμή , ο άλλος σκοπός αυτής της διατριβής είναι να χρησιμοποιήσει την πολυμαθή γνώση για την πραγματική εφαρμογή.

Σε αυτή την εργασία, θα προγραμματίσουμε δύο συστήματα. Οι κύριοι αλγόριθμοι για αυτά τα δύο σχεδιασμένα συστήματα είναι η συσχέτιση και το φίλτρο FIR Wiener. Για να δούμε αν αυτοί οι δύο αλγόριθμοι μπορούν να λειτουργήσουν για την αναγνώριση φωνής, θα καλέσουμε διαφορετικούς ανθρώπους από διαφορετικές χώρες για να δοκιμάσουν τα σχεδιασμένα συστήματα. Για να πάρουμε αξιόπιστα αποτελέσματα, οι δοκιμές θα πρέπει να ολοκληρωθούν σε διαφορετικές καταστάσεις. Πρώτον, το περιβάλλον των δοκιμών θα είναι θορυβώδης και αθόρυβο, αντίστοιχα, για τη διερεύνηση του θορύβου για τα σχεδιασμένα συστήματα. Και οι λέξεις δοκιμών θα πρέπει να επιλεγούν από διαφορετικά ζευγάρια που είναι οι εύκολες αναγνωρισμένες λέξεις και οι δυσκολότερες. Δεδομένου ότι τα δύο σχεδιασμένα συστήματα χρειάζονται τρεις λέξεις εισόδου ομιλίας που είναι δύο λέξεις ομιλία αναφοράς και μια λέξη ομιλίας στόχος, γι 'αυτό είναι σημαντικό να ελέγξουμε αν τα δύο σχεδιασμένα συστήματα λειτουργούν καλά όταν οι λέξεις ομιλίας αναφοράς και οι λέξεις ομιλία στόχος καταγράφονται από διαφορετικά πρόσωπα.

2° Κεφάλαιο

Θεωρία

Αυτό το τμήμα θεωρίας παρουσιάζει κάποιους ορισμούς και πληροφορίες για αυτή την διατριβή. Χρειαστήκαμε κάποιες υποχρεωτικές πληροφορίες για την υποστήριξη της έρευνας μας. Με την αξιοποίηση της θεωρητικής γνώσης, πετύχαμε το στόχο της παρούσας διατριβής. Συμπεριλαμβανομένων του DC επιπέδου και της δειγματοληπτικής θεωρίας, DFT, FFT, της ομαλοποίησης του φάσματος, του αλγόριθμου συσχέτισης, του αλγόριθμου αυτοσυσχέτισης, του φίλτρου FIR Wiener, την χρήση της λειτουργίας φασματογράμματος πήραμε τα επιθυμητά σήματα.

2.1 Επίπεδο DC και θεωρία δειγματοληψίας

Όταν κάνει την ανάλυση επεξεργασίας σήματος, οι πληροφορίες του επιπέδου DC για το σήμα στόχου δεν είναι τόσο χρήσιμες εκτός αν το σήμα εφαρμοστεί στο πραγματικό αναλογικό κύκλωμα, όπως του AD μετατροπέα, ο οποίος έχει την απαίτηση της παρεχόμενης τάσης. Κατά την ανάλυση των σημάτων στο πεδίο συχνοτήτων, το επίπεδο DC δεν είναι τόσο χρήσιμο. Μερικές φορές το μέγεθος του επιπέδου DC στον πεδίο συχνοτήτων θα επηρεάσει την ανάλυση όταν το σήμα στόχος είναι πιο συγκεντρωμένο στην χαμηλή ζώνη συχνοτήτων. Σε κατάσταση WSS για τη στοχαστική διαδικασία, η διακύμανση και η μέση τιμή του σήματος δεν θα αλλάξει, κατά την αλλαγή του χρόνου. Έτσι, προσπαθούμε να μειώσουμε αυτή την επίδραση από την αφαίρεση της μέσης τιμής των καταγεγραμμένων σημάτων. Αυτό θα αφαιρέσει τις μηδενικές συνιστώσες συχνότητας για το επίπεδο DC στο φάσμα συχνοτήτων.

Σε αυτή την εργασία, δεδομένου ότι χρησιμοποιώντας το μικρόφωνο γίνεται καταγραφή της αναλογικής ομιλίας του ατόμου μέσω του υπολογιστή, έτσι ώστε η ποιότητα των δεδομένων του σήματος ομιλίας θα αποφασίσει άμεσα την ποιότητα της αναγνώρισης φωνής. Η

συχνότητα δειγματοληψίας είναι ένας από τους καθοριστικούς παράγοντες για την ποιότητα των δεδομένων.

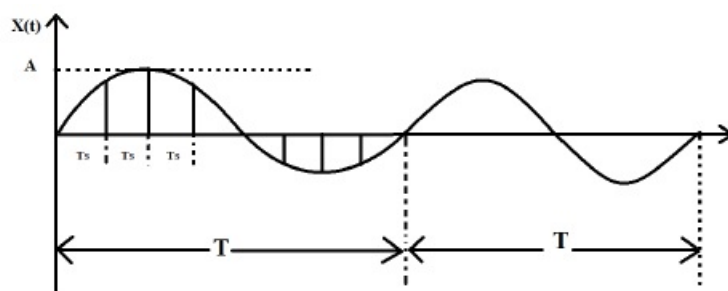
Γενικά, το αναλογικό σήμα μπορεί να παρασταθεί ως

$$x(t) = \sum_{i=1}^N A_i \cos(2\pi f_i t + \phi_i) \quad (1)$$

Αυτό το αναλογικό σήμα αποτελείται στην πραγματικότητα από πολλές συνιστώσες διαφορετικών συχνοτήτων. Υποθέτοντας ότι υπάρχει μόνο μία συνιστώσα συχνότητας σε αυτό το αναλογικό σήμα, και δεν έχει καμία μετατόπιση φάσης. Έτσι, αυτό το αναλογικό σήμα γίνεται :

$$x(t) = A \cos(2\pi f t) \quad (2)$$

Το αναλογικό σήμα δεν μπορεί να εφαρμοστεί απ'ευθείας στον υπολογιστή. Είναι αναγκαίο να δοκιμάσουμε το αναλογικό σήμα $x(t)$ στο σήμα διακριτού χρόνου $x(n)$, το οποίο ο υπολογιστής μπορεί να χρησιμοποιήσει για να επεξεργαστεί. Γενικότερα, το διακριτό σήμα $x(n)$ θεωρείται πάντοτε ως μία αλληλουχία σήματος ή ενός διάνυσματος. Έτσι το MATLAB μπορεί να κάνει τον υπολογισμό για το σήμα διακριτού χρόνου. Το ακόλουθο Σχήμα 1 είναι η δειγματοληψία του αναλογικού σήματος σε σήμα διακριτού χρόνου:



Σχήμα 1: Η δειγματοληψία του αναλογικού σήματος

Όπως φαίνεται στο Σχήμα 1 παραπάνω, το χρονικό διάστημα του αναλογικού σήματος $x(t)$ είναι T . Η περίοδος δειγματοληψίας του διακριτού σήματος χρόνου είναι T_s . Υποθέτοντας ότι το αναλογικό σήμα δειγματοληπτείται από τον αρχικό χρόνο 0, έτσι ώστε το δείγμα του σήματος να μπορεί να γραφτεί ως διάνυσμα $x(n) = [x(0), x(1), x(2), x(3), x(4) \dots x(N-1)]$.

Όπως είναι γνωστό, η σχέση μεταξύ της αναλογικής συχνότητας του σήματος και ο χρόνος της περιόδου είναι αλληλοπαθή. Έτσι η συχνότητα δειγματοληψίας του δείγματος του σήματος είναι $f_s=1/T_s$.

Ας υποθέσουμε ότι το μήκος του $x(n)$ είναι N για K αρχικές χρονικές περιόδους. Τότε, η σχέση μεταξύ T και T_s είναι $N \times T_s = K \times T$. Έτσι, $N/K = T/T_s = f_s/f$, όπου N και K είναι ακέραιοι αριθμοί. Και εάν αυτό το αναλογικό σήμα γίνεται ακριβώς δειγματοληψία με το ίδιο διάστημα δειγματοληψίας και τη δειγματοληψία του σήματος που είναι περιοδική, τότε N/K είναι ακέραιος επίσης. Διαφορετικά, η δειγματοληψία του σήματος θα είναι απεριοδική.

Σύμφωνα με το θεώρημα δειγματοληψίας (θεώρημα Nyquist) [2], όταν η συχνότητα δειγματοληψίας είναι μεγαλύτερη ή ίση από 2 φορές των μέγιστων αναλογικών συχνοτήτων του σήματος, το σήμα διακριτού χρόνου μπορεί να χρησιμοποιηθεί για την επανάκτηση του αρχικού αναλογικού σήματος. Και η υψηλότερη συχνότητα δειγματοληψίας θα έχει ως αποτέλεσμα τα καλύτερα δείγματα σημάτων για ανάλυση. Θα χρειαστεί ταχύτερο επεξεργαστή για την επεξεργασία του σήματος και την εκτίμηση για περισσότερους χώρους δεδομένων. Σε μη τηλεπικοινωνιακές εφαρμογές, στις οποίες το υποσύστημα αναγνώρισης φωνής έχει πρόσβαση σε ομιλία υψηλής ποιότητας, οι συχνότητες δείγματος που έχουν χρησιμοποιηθεί είναι 10 kHz, 14 kHz και 16 kHz. Αυτά τα δείγματα συχνοτήτων δίνουν καλύτερη ανάλυση χρόνου και συχνότητας [1]. Σε αυτή την εργασία, για το MATLAB, η συχνότητα δειγματοληψίας καθορίζεται ως 16 kHz. Έτσι το μήκος του εγγεγραμμένου σήματος σε 2 δευτερόλεπτα θα είναι 32000 μονάδες χρόνου στο MATLAB. Στο επόμενο μέρος, θα παρουσιαστεί η θεωρία του DFT και FFT οι οποίες είναι σημαντικές όταν προσπαθούμε να αναλύσουμε φάσματα στον τομέα συχνοτήτων. Και είναι το κλειδί για να κάνουμε την αναγνώριση φωνής σε αυτή τη διατριβή.

2.2 Πεδίο χρόνου σε πεδίο συχνότητας: DFT και FFT

2.2.1 DFT

Ο DFT είναι η συντομογραφία του διακριτού μετασχηματισμού Fourier. Έτσι, ο DFT είναι απλώς ένα είδος Μετασχηματισμού Fourier για διακριτό χρόνο $x(n)$ αντί συνεχόμενου αναλογικού σήματος $x(t)$. Η εξίσωση Μετασχηματισμού Fourier έχει ως εξής :

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{j\omega n} \quad (3)$$

Από την εξίσωση , η κύρια λειτουργία του Μετασχηματισμού Fourier είναι να μετατρέψει τη μεταβλητή από τη μεταβλητή n στη μεταβλητή ω , πράγμα που σημαίνει τη μετατροπή των σημάτων από το πεδίο χρόνου στο πεδίο συχνοτήτων.

Υποθέτοντας ότι το σήμα εγγεγραμμένης φωνής $x(n)$ είναι μία αλληλουχία ή ένα διάνυσμα το οποίο αποτελείται από πολύπλοκες τιμές, όπως $x(n) = R + I \cdot j$, όπου το R αντιπροσωπεύει το πραγματικό μέρος της τιμής, και το I αντιπροσωπεύει το φανταστικό μέρος της τιμής. Δεδομένου ότι ο εκθέτης του συντελεστή είναι :

$$e^{j\omega n} = \cos(\omega n) + j \cdot \sin(\omega n) \quad (4)$$

Έτσι:

$$x(n) \cdot e^{j\omega n} = (R + I \cdot j) \cdot [\cos(\omega n) + j \cdot \sin(\omega n)] = R \cdot \cos(\omega n) + R \cdot j \cdot \sin(\omega n) + I \cdot \cos(\omega n) + I \cdot j \cdot \sin(\omega n) \quad (5)$$

Αναδιάταξη του πραγματικού μέρους και του φανταστικού μέρους της εξίσωσης. Παίρνουμε:

$$x(n) \cdot e^{j\omega n} = [R \cdot \cos(\omega n) + I \cdot \cos(\omega n)] + [R \cdot j \cdot \sin(\omega n) + I \cdot j \cdot \sin(\omega n)] \quad (6)$$

Έτσι, η εξίσωση (3) γίνεται:

$$X(\omega) = \sum [R \cos(\omega n) + I \cos(\omega n)] + \sum j [R \sin(\omega n) + I \sin(\omega n)] \quad (7)$$

Η εξίσωση (7), είναι επίσης κατασκευασμένη από το πραγματικό μέρος και το φανταστικό μέρος. Δεδομένου ότι σε γενικές γραμμές, η πραγματική τιμή του σήματος $x(n)$ χρησιμοποιείται. Έτσι, αν το φανταστικό μέρος είναι $I = 0$ τότε ο Μετασχηματισμός Fourier είναι:

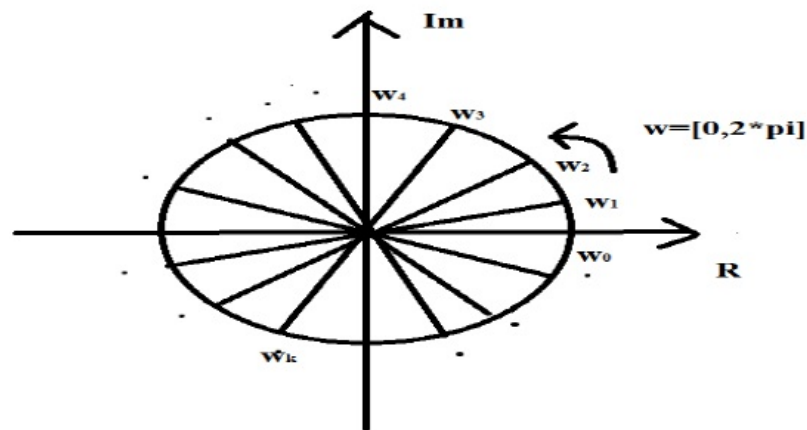
$$X(\omega) = \sum_{n=-\infty}^{\infty} [R \cos(\omega n)] + \sum_{n=-\infty}^{\infty} [j R \sin(\omega n)] \quad (8)$$

Οι αναλύσεις παραπάνω είναι τα γενικά βήματα για να προγραμματίσουμε το μετασχηματισμό Fourier προγραμματίζοντας τον συντελεστή συχνότητας υπολογισμού ο οποίος αποτελείται από το πραγματικό μέρος και το φανταστικό μέρος με το μέγεθος του σήματος. Αλλά στο MATLAB, υπάρχει μια άμεση εντολή "fft", η οποία μπορεί να χρησιμοποιηθεί άμεσα για να πάρουμε την συνάρτηση μετασχηματισμού. Και η μεταβλητή ω στην εξίσωση (3) μπορεί να αντιμετωπιστεί ως συνεχής μεταβλητή.

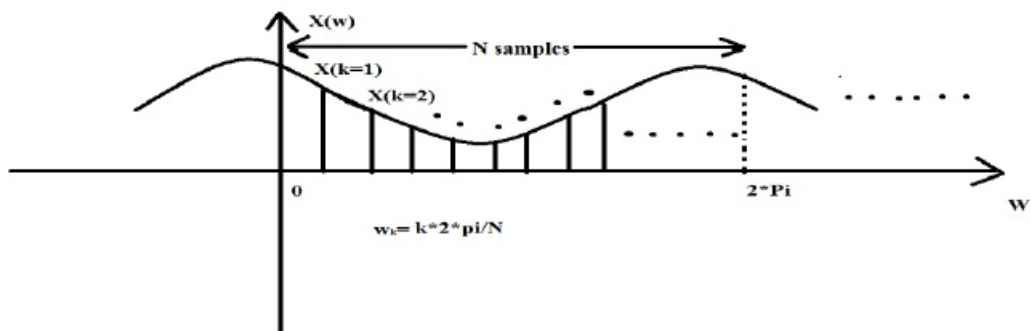
Υποθέτοντας ότι η συχνότητα ω βρίσκεται στο $[0, 2\pi]$, το $X(\omega)$ μπορεί να θεωρηθεί ως αναπόσπαστο ή ως άθροισμα σήματος όλων των συνιστωσών συχνοτήτων. Τότε η συνιστώσα συχνότητας $X(k)$ από $X(\omega)$ έγινε με δειγματοληψία ολόκληρου του διαστήματος συχνότητας $\omega = [0, 2\pi]$ από N δείγματα. Έτσι, αυτό σημαίνει ότι η συνιστώσα συχνότητας $\omega_k = k \times 2\pi/N$. Και η εξίσωση DFT για την ω_k συνιστώσα συχνότητας είναι όπως παρακάτω:

$$X(k) = X(\omega_k) = \sum_{n=-\infty}^{\infty} x(n)e^{j\omega_k n} = \sum_{n=0}^{N-1} x(n)e^{j\frac{2\pi k}{N}n}, \quad 0 \leq k \leq N-1 \quad (9)$$

Αυτή η εξίσωση χρησιμοποιείται για τον υπολογισμό του μεγέθους της συνιστώσας συχνότητας. Το κλειδί κατανόησης του DFT είναι σχετικό με τη δειγματοληψία στο πεδίο της συχνότητας. Η διαδικασία δειγματοληψίας φαίνεται πιο καθαρά στα ακόλουθα σχήματα.



Σχήμα 2: Η δειγματοληψία συχνότητας σε κύκλο



Σχήμα 3: Η δειγματοληψία συχνότητας σε άξονες

Επιπλέον, το MATLAB ασχολείται με τα δεδομένα για τα διανύσματα και τις μήτρες σημάτων. Είναι απαραίτητη η κατανόηση της γραμμικής άλγεβρας ή της διαδικασίας του πίνακα του DFT. Παρατηρώντας την εξίσωση (3), εκτός από το χειριστή άθροισης, η εξίσωση αποτελείται από 3 μέρη: έξοδος $X(\omega)$, είσοδος $x(n)$, και ο συντελεστής φάσης $e^{j\omega_k n}$. Δεδομένου ότι όλες οι πληροφορίες των συνιστωσών συχνότητας είναι από τον συντελεστή φάσης $e^{j\omega_k n}$. Έτσι ώστε ο συντελεστής φάσης μπορούν να υποδηλωθεί ως:

$$W_N^{kn} = e^{j\omega_k n}, \quad n \text{ και } k \text{ ακέραιοι από } 0 \text{ έως } N-1 \quad (10)$$

Γράφοντας τον συντελεστή φάσης σε διανυσματική μορφή είναι:

$$W_N^{kn} = e^{j\omega_k n} = [W_N^{0k}, W_N^{1k}, W_N^{2k}, W_N^{3k}, W_N^{4k}, \dots, W_N^{(N-1)k}] \quad (11)$$

Και

$$x(n) = [x(0), x(1), x(2), \dots, x(N-1)] \quad (12)$$

Έτσι, η εξίσωση (9) για τη συνιστώσα συχνότητας $X(k)$ είναι:

$$X(k) = (W_N^{kn})^H \cdot x(n) \quad (13)$$

Αυτό είναι η μορφή του διανύσματος σχετικά με τον υπολογισμό της συχνότητας με τη χρήση της μεθόδου του DFT. Όμως αν το σήμα είναι μια πραγματικά μεγάλη αλληλουχία, και ο χώρος μνήμης είναι πεπερασμένος, τότε η χρησιμοποίηση του DFT για να πάρουμε το μετασχηματισμένο σήμα θα είναι περιορισμένη. Ο ταχύτερος και πιο αποτελεσματικός υπολογισμός του DFT είναι ο FFT. Θα παρουσιάσουμε εν συντομία για τον FFT στην επόμενη ενότητα.

2.2.2 FFT

Ο FFT είναι μια συντομογραφία του γρήγορου μετασχηματισμού Fourier. Ουσιαστικά, ο FFT εξακολουθεί να είναι ο DFT για το μετασχηματισμό του σήματος διακριτού χρόνου από το πεδίο του χρόνου στο πεδίο συχνοτήτων του. Η διαφορά είναι ότι ο FFT είναι ταχύτερος και πιο αποτελεσματικός για τον υπολογισμό. Και υπάρχουν πολλοί τρόποι για να αυξηθεί η αποτελεσματικότητα υπολογισμού του DFT, αλλά ο πιο ευρέως χρησιμοποιούμενος FFT αλγόριθμος είναι ο Radix - 2 FFT αλγόριθμος [2].

Αφού FFT εξακολουθεί να είναι ο υπολογισμός του DFT, έτσι είναι βολικό να διερευνήσουμε τον FFT λαμβάνοντας υπόψη το N- DFT σημείου εξίσωσης:

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{kn}, \quad k = 0, 1, 2, \dots, N-1 \quad (14)$$

Πρώτον το $x(n)$ χωρίζεται σε δύο μέρη: $x(\text{περιττή}) = x(2m+1)$ και το $x(\text{άρτια}) = x(2m)$, όπου $m = 0, 1, 2, \dots, N/2-1$. Τότε η εξίσωση N-σημείο DFT γίνεται επίσης δύο τμήματα, για κάθε $N/2$ σημεία:

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{kn} = \sum_{m=0}^{N/2-1} x(2m)W_N^{2mk} + \sum_{m=0}^{N/2-1} x(2m+1)W_N^{(2m+1)k} = \sum_{m=0}^{N/2-1} x(2m)W_N^{2mk} + W_N^k \sum_{m=0}^{N/2-1} x(2m+1)W_N^{2mk}, \quad (15)$$

όπου $m=0, 1, 2, \dots, N/2-1$

Αφού:

$$e^{j\omega_k n} = \cos(\omega_k n) + j \sin(\omega_k n). \quad (16)$$

$$\begin{aligned} e^{j(\omega_k + \pi)n} &= \cos[(\omega_k + \pi)n] + j \cdot \sin[(\omega_k + \pi)n] \\ &= -\cos(\omega_k n) - j \cdot \sin(\omega_k n) = -[\cos(\omega_k n) + j \cdot \sin(\omega_k n)] = -e^{j\omega_k n} \end{aligned} \quad (17)$$

Τότε:

$$e^{j(\omega_k + \pi)n} = -e^{j\omega_k n} \quad (18)$$

Έτσι, όταν ο συντελεστής φάσης μετατοπίζεται σε μισή περίοδο, η τιμή του συντελεστή φάσης δεν θα αλλάξει, αλλά το πρόσημο του συντελεστή φάσης θα είναι το αντίθετο. Αυτό ονομάζεται ιδιότητα συμμετρίας [2] του συντελεστή φάσης. Έτσι:

$$W_N^{(k+\frac{N}{2})n} = -W_N^{kn} \quad (19)$$

Και

$$(W_N^{kn})^2 = -W_N^{kn} = e^{j\frac{4\pi k}{N}n} \quad (20)$$

Η εξίσωση N-σημείο DFT γίνεται τελικά:

$$X(k) = \sum_{m=0}^{N/2-1} x_1(m)W_{N/2}^{mk} + W_N^k \sum_{m=0}^{N/2-1} x_2(m)W_{N/2}^{mk} = X_1(k) + W_N^k X_2(k), \quad k=0, 1, \dots, N/2 \quad (21)$$

$$X(k + N/2) = X_1(k) - W_N^k X_2(k), \quad k=0, 1, 2, \dots, N/2 \quad (22)$$

Έτσι το N-σημείο DFT χωρίζεται σε δύο N/2-σημείο DFT. Από την εξίσωση (21), το $X_1(k)$ έχει $(N/2) \times (N/2) = (N/2)^2$ μιγαδικούς πολλαπλασιασμούς. Έτσι ώστε ο συνολικός αριθμός των μιγαδικών πολλαπλασιασμών για το $X(k)$ είναι $2 \times (N/2)^2 + N/2 = N^2/2 + N/2$. Για την αρχική εξίσωση N-σημείο DFT (14), έχει N^2 μιγαδικούς πολλαπλασιασμούς. Έτσι, στο πρώτο στάδιο, διαχωρίζοντας το $x(n)$ σε δύο μέρη καθιστά τον αριθμό των μιγαδικών πολλαπλασιασμών από N^2 σε $N^2/2 + N/2$. Ο αριθμός των υπολογισμών έχει μειωθεί κατά το ήμισυ περίπου.

Αυτή είναι η μέθοδος για τη μείωση των υπολογισμών από N σημεία σε $N/2$ σημεία. Έτσι, διαχωρίζοντας συνεχώς το $x_1(m)$ και $x_2(m)$ ανεξάρτητα σε περιττό και άρτιο μέρος με τον ίδιο τρόπο, οι υπολογισμοί για τα $N/2$ σημεία θα μειωθούν σε $N/4$. Στη συνέχεια, οι υπολογισμοί του DFT θα μειωθούν συνεχώς. Έτσι, το σήμα για N -σημείο DFT συνεχώς διαχωρίζεται μέχρις ότου η τελική ακολουθία σήματος να μειωθεί στην ακολουθία του ενός σημείου. Υποθέτοντας ότι υπάρχουν $N = 2^s$ σημεία DFT που πρέπει να υπολογιστούν.

Έτσι, ο αριθμός τέτοιων διαχωρισμών που μπορούν να γίνουν είναι $s = \log_2(N)$. Έτσι ώστε ο συνολικός αριθμός των μιγαδικών πολλαπλασιασμών θα μειωθεί περίπου σε $(N/2)\log_2(N)$. Για τους επιπλέον υπολογισμούς, ο αριθμός τους θα μειωθεί σε $N \log_2(N)$ [2]. Η ταχύτητα του υπολογισμού DFT έχει βελτιωθεί επειδή οι πολλαπλασιασμοί και τα αθροίσματα μειώνονται. Η κύρια ιδέα για το Radix-2 FFT είναι να διαχωρίσει την παλιά ακολουθία δεδομένων σε περιττό και άρτιο μέρος έτσι ώστε να μειώσει στο ήμισυ περίπου τους αρχικούς υπολογισμούς.

2.3 Ανάλυση συχνότητας στο MATLAB για την Αναγνώριση Φωνής

2.3.1 Κανονικοποίηση Φάσματος

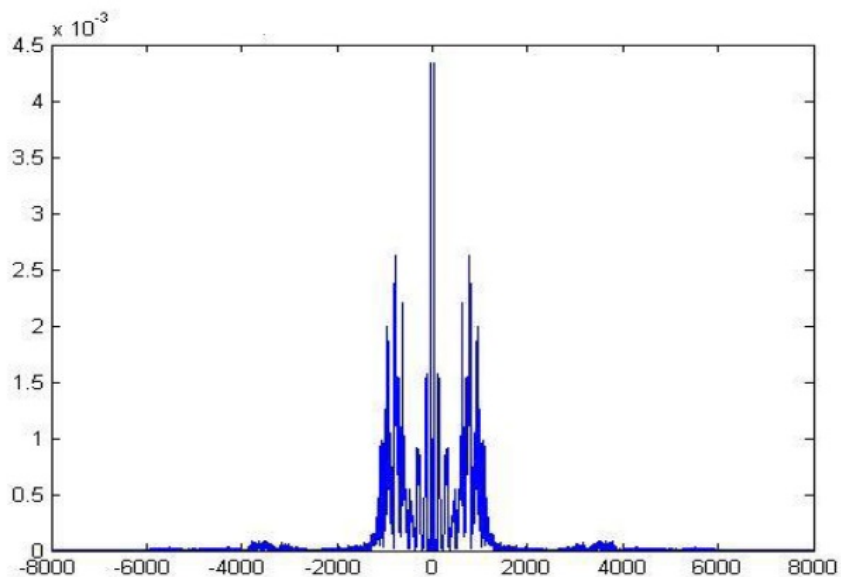
Μετά τους υπολογισμούς DFT και FFT, τα προβλήματα διερεύνησης θα αλλάξουν από τα διακριτά σήματα χρόνου $x(n)$ προς το σήμα στο πεδίο των συχνοτήτων $X(\omega)$. Το φάσμα του $X(\omega)$ είναι το ολοκλήρωμα ή το άθροισμα όλων των συνιστωσών συχνότητας. Όταν μιλάμε για τη συχνότητα του σήματος φωνής για διαφορετικές λέξεις, κάθε λέξη έχει τη δική της ζώνη συχνοτήτων, δεν είναι απλώς μια τυχαία συχνότητα. Και στη ζώνη συχνοτήτων της κάθε λέξης, το φάσμα ($|X(\omega)|$) ή η ισχύς φάσματος ($|X(\omega)|^2$) έχει μέγιστη τιμή και ελάχιστη τιμή. Συγκρίνοντας τις διαφορές μεταξύ των δύο διαφορετικών σημάτων φωνής, είναι δύσκολο να συγκρίνουμε δύο φάσματα σε διαφορετικά πρότυπα μέτρησης. Έτσι, χρησιμοποιώντας την κανονικοποίηση μπορεί να κάνει το πρότυπο μέτρησης ίδιο.

Κατά μία έννοια, η εξομάλυνση μπορεί να μειώσει το σφάλμα κατά τη σύγκριση των φασμάτων, το οποίο είναι καλό για την αναγνώριση φωνής [3]. Έτσι, πριν από την ανάλυση των διαφορών φάσματος για διαφορετικές λέξεις, το πρώτο βήμα είναι να ομαλοποιήσουμε

το φάσμα $|X(\omega)|$ από την γραμμική κανονικοποίηση. Η εξίσωση της γραμμικής κανονικοποίησης είναι όπως παρακάτω :

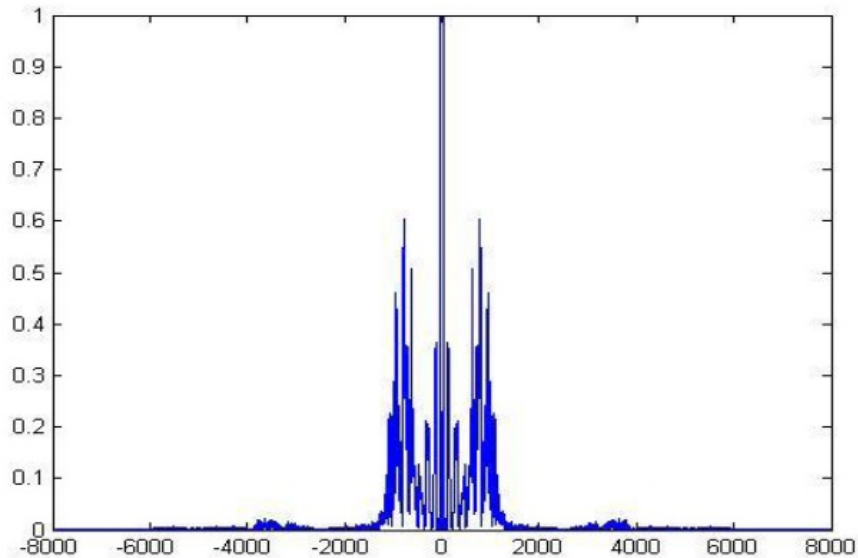
$$y=(x-\text{MinValue})/(\text{MaxValue}-\text{MinValue}) \quad (23)$$

Μετά την κανονικοποίηση, οι τιμές του φάσματος των $|X(\omega)|$, έχουν οριστεί σε διάστημα $[0,1]$. Η κανονικοποίηση αλλάζει μόνο το εύρος των τιμών του φάσματος , αλλά δεν αλλάζει το σχήμα ή τις πληροφορίες του ίδιου του φάσματος. Έτσι, η κανονικοποίηση είναι καλή για τη σύγκριση του φάσματος. Χρησιμοποιώντας το MATLAB δίνει ένα παράδειγμα για να δούμε πώς το φάσμα αλλάζει από τη γραμμική κανονικοποίηση. Πρώτον, καταγράφει ένα σήμα ομιλίας και το μετατρέπει σε FFT. Στη συνέχεια, παίρνουμε τις απόλυτες τιμές του φάσματος FFT . Το φάσμα FFT χωρίς κανονικοποίηση είναι όπως παρακάτω:



Σχήμα 4: Οι απόλυτες τιμές του φάσματος FFT χωρίς κανονικοποίηση

Το κανονικοποιημένο φάσμα είναι όπως παρακάτω:



Σχήμα 5: Οι απόλυτες τιμές του φάσματος FFT με κανονικοποίηση

Από το Σχήμα 4 και το Σχήμα 5 , η διαφορά μεταξύ των δύο φασμάτων είναι μόνο το διάστημα τιμών του φάσματος $|X(\omega)|$, το οποίο αλλάζει από $[0, 4.5 \times 10^{-3}]$ σε $[0, 1]$. Μετά την κανονικοποίηση των απόλυτων τιμών του FFT, το επόμενο βήμα του προγραμματισμού της αναγνώρισης φωνής είναι να παρατηρήσουμε τα φάσματα των τριών καταγεγραμμένων σημάτων ομιλίας και να βρούμε τους αλγόριθμους για τη σύγκριση των διαφορών μεταξύ του τρίτου καταγεγραμμένου σήματος στόχου και των δύο πρώτων καταγεγραμμένων σημάτων αναφοράς.

2.3.2 Ο αλγόριθμος συσχέτισης

Υπάρχει μία σημαντική ποσότητα δεδομένων σχετικά με τη συχνότητα της θεμελιώδους φωνής (F_0) στην ομιλία των ομιλητών οι οποίοι διαφέρουν ως προς την ηλικία και το φύλο. [4] Για τον ίδιο ομιλητή, οι διαφορετικές λέξεις έχουν επίσης διαφορετικές ζώνες συχνοτήτων οι οποίες οφείλονται στις διαφορετικές δονήσεις της φωνητικής χορδής. Και τα σχήματα των φασμάτων είναι επίσης διαφορετικά. Αυτές είναι οι βάσεις αυτής της διπλωματικής εργασίας για την αναγνώριση φωνής. Σε αυτή την εργασία, για να συνειδητοποιήσουμε την αναγνώριση λόγου, πρέπει να συγκρίνουμε τα φάσματα μεταξύ του τρίτου καταγεγραμμένου σήματος και τα δύο πρώτων καταγεγραμμένων σημάτων αναφοράς. Ελέγχοντας ποιο από τα δύο καταγεγραμμένα σήματα αναφοράς ταιριάζει καλύτερα με το τρίτο εγγεγραμμένο σήμα,

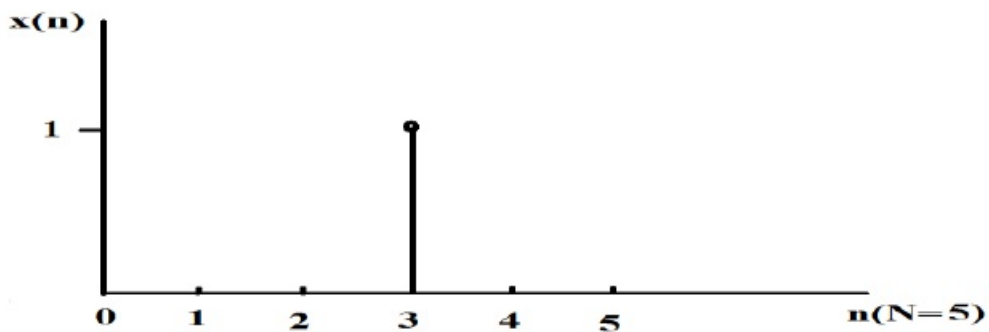
το σύστημα θα κρίνει ποια λέξη αναφοράς καταγράφεται και πάλι για τρίτη φορά. Όταν σκεφτόμαστε τη συσχέτιση των δύο σημάτων, ο πρώτος αλγόριθμος που θα πρέπει να εξεταστεί είναι ο αλγόριθμος συσχέτισης των δύο σημάτων. Η συνάρτηση συσχέτισης είναι πραγματικά χρήσιμη για να εκτιμηθεί η παράμετρος μετατόπισης [5]. Εδώ η παράμετρος μετατόπισης θα αναφέρεται ως μετατόπιση συχνότητας.

Η εξίσωση της συσχέτισης δύο σημάτων είναι η παρακάτω:

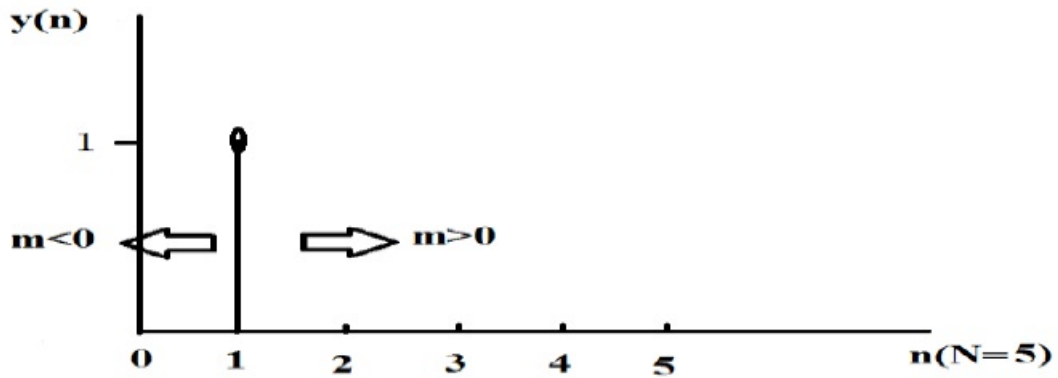
$$r_{xy} = r(m) = \sum_{n=-\infty}^{\infty} x(n)y(n+m), m = 0, \pm 1, \pm 2, \pm 3, \dots \quad (24)$$

Από την εξίσωση, η βασική ιδέα του αλγορίθμου για την συσχέτιση είναι 3 βήματα:

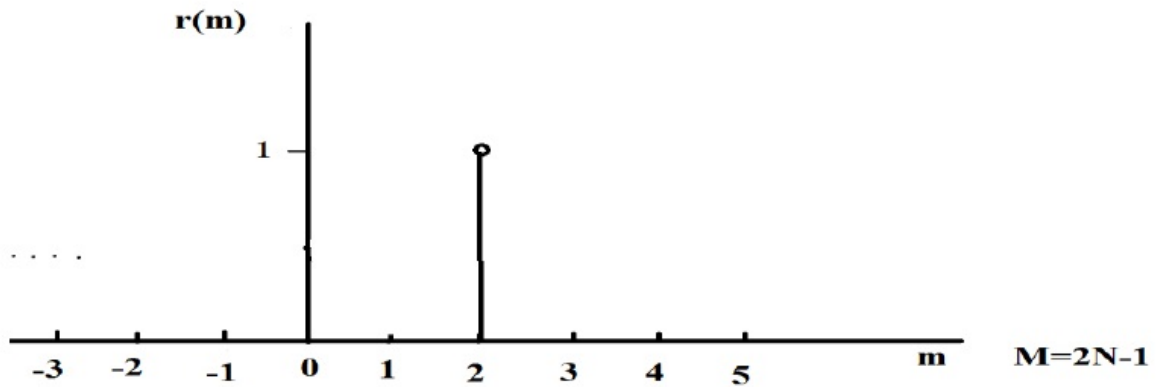
Πρώτον, να καθορίσει ένα από τα δύο σήματα $x(n)$, και να μετατοπίσει το άλλο σήμα $y(n)$ αριστερά ή δεξιά με κάποιες μονάδες χρόνου. Δεύτερον, να πολλαπλασιάσει την τιμή του $x(n)$ με την μετατόπιση του σήματος $y(n+m)$ θέσης από τη θέση του. Τέλος, να λάβει το άθροισμα όλων των αποτελεσμάτων πολλαπλασιασμού για $x(n) \times y(n+m)$. Για παράδειγμα, δύο σήματα ακολουθίας $x(n) = [0\ 0\ 0\ 1\ 0]$, $y(n) = [0\ 1\ 0\ 0\ 0]$, τα μήκη για τα δύο σήματα είναι $N = 5$. Έτσι, η συσχέτιση για $x(n)$ και $y(n)$ είναι τα σχήματα που εμφανίζονται παρακάτω:



Σχήμα 6: Η σειρά σήματος $x(n)$



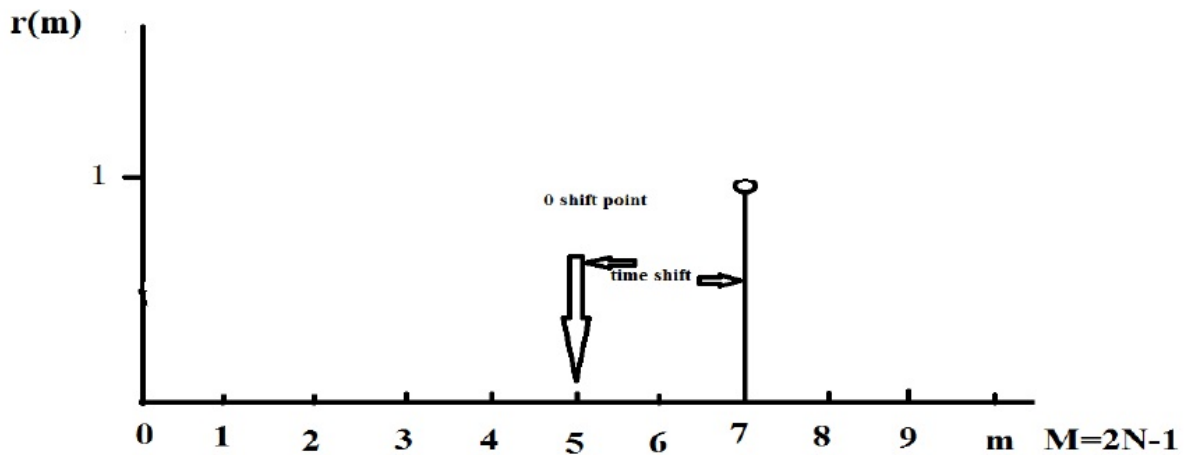
Σχήμα 7: Η ακολουθία σήματος (n) θα μετατοπιστεί προς τα αριστερά ή δεξιά κατά m μονάδες



Σχήμα 8: Τα αποτελέσματα της συσχέτισης, άθροισμα των πολλαπλασιασμών

Όπως το παράδειγμα, υπάρχει μια διακριτή χρονική μετατόπιση περίπου 2 μονάδες χρόνου μεταξύ των σημάτων $x(n)$ και $y(n)$. Από το Σχήμα 8, η συσχέτιση $r(m)$ έχει μη μηδενική τιμή, το οποίο είναι ίσο με 1 στη θέση $m = 2$. Έτσι, ο m -άξονας στο Σχήμα 8 δεν είναι πλέον ο άξονας του χρόνου για το σήμα. Είναι ο άξονας της χρονικής μετατόπισης. Δεδομένου ότι τα μήκη των δύο σημάτων $x(n)$ και $y(n)$ είναι και τα δύο $N = 5$, έτσι ώστε το μήκος του άξονα του χρόνου μετατόπισης να είναι $2N$. Όταν χρησιμοποιούμε το MATLAB για να κάνει τη συσχέτιση, το μήκος της συσχέτισης είναι ακόμα $2N$. Αλλά στο MATLAB, η αποτύπωση της συσχέτισης είναι από 0 έως $2N - 1$, όχι από $-N$ έως $+N$ πια. Τότε η θέση του σημείου μηδέν της χρονικής μετατόπισης θα μετατοπιστεί από 0 έως N . Έτσι, όταν τα δύο σήματα δεν έχουν χρονική μετατόπιση, η μέγιστη τιμή συσχέτισης τους θα είναι στη θέση $m = N$ στο MATLAB, η οποία είναι η μεσαία θέση του σημείου για το συνολικό μήκος της συσχέτισης.

Στο MATLAB, η αποτύπωση του Σχ.8 θα είναι όπως παρακάτω :



Σχήμα 9: Η συσχέτιση η οποία παρουσιάζεται στο MATLAB (μη πραγματικό Σχήμα του MATLAB)

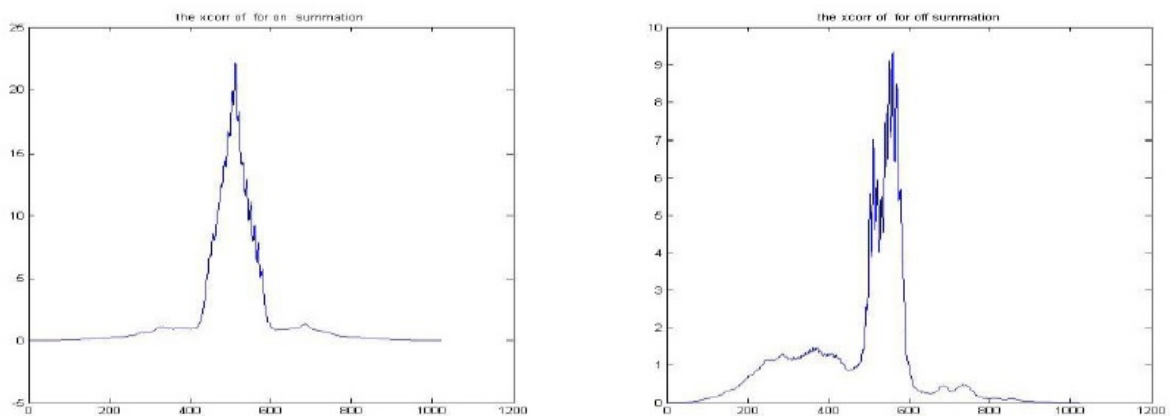
Από Σχ.9 , η μέγιστη τιμή συσχέτισης δύο σημάτων δεν είναι στο μεσαίο σημείο για το συνολικό μήκος του συσχέτισης. Όπως το παράδειγμα που δίνεται, τα μήκη των δύο σημάτων είναι $N = 5$, έτσι ώστε το συνολικό μήκος της συσχέτισης να είναι $2N = 10$. Στη συνέχεια, όταν τα δύο σήματα δεν έχουν καμία μετατόπιση χρόνο, η μέγιστη τιμή της συσχέτισης τους θα πρέπει να είναι σε $m = 5$. Αλλά στο σχήμα 9 , η μέγιστη τιμή συσχέτισης τους είναι στη θέση $m = 7$, το οποίο σημαίνει δύο αρχικά σήματα έχουν 2 μονάδες χρονική μετατόπιση σε σχέση με τη θέση μετατόπισης 0.

Από το παράδειγμα, έχουμε δυο σημαντικές πληροφορίες της συσχέτισης . Η μία είναι όταν δύο αρχικά σήματα δεν έχουν καμία μετατόπιση χρόνου, η συσχέτιση τους θα πρέπει να είναι η μέγιστη. Η άλλη πληροφορία είναι ότι η διαφορά θέσης μεταξύ της μέγιστης θέσης τιμής και το μεσαίο σημείο θέσης της συσχέτιση είναι το μήκος του χρόνου μετατόπισης για δύο αρχικά σήματα.

Τώρα υποθέτοντας ότι τα δύο καταγεγραμμένα σήματα ομιλίας για την ίδια λέξη είναι εντελώς ίδια, έτσι τα φάσματα των καταγεγραμμένων σημάτων ομιλίας είναι επίσης ίδια. Στη συνέχεια, όταν γίνεται η συσχέτιση των δύο ίδιων φασμάτων και η απεικόνιση της συσχέτισης, η γραφική παράσταση της συσχέτισης πρέπει να είναι τελείως συμμετρική σύμφωνα με τον αλγόριθμο της συσχέτισης. Ωστόσο, για την πραγματική καταγραφή ομιλίας, τα φάσματα των δύο καταγεγραμμένων σημάτων ομιλίας για την ίδια λέξη μπορεί να μην είναι τα ίδια εντελώς. Αλλά τα φάσματα τους πρέπει να είναι παρόμοια, πράγμα που σημαίνει

ότι το γράφημα συσχέτισης τους πρέπει να είναι περίπου συμμετρικό. Αυτή είναι η πιο σημαντική έννοια στην παρούσα διατριβή για την αναγνώριση φωνής κατά το σχεδιασμό του συστήματος I.

Με τη σύγκριση του επιπέδου της συμμετρικής ιδιότητας για τη συσχέτιση, το σύστημα μπορεί να πάρει την απόφαση ποιο από τα δύο καταγεγραμμένα σήματα έχουν περισσότερα παρόμοια φάσματα. Με άλλα λόγια, αυτά τα δύο καταγεγραμμένα σήματα έχουν καταγραφεί πιθανότατα από την ίδια λέξη. Το παρακάτω σχήμα είναι το αποτέλεσμα προσομοίωσης στο MATLAB για τις συσχετίσεις:



Σχήμα 10: Οι γραφικές παραστάσεις των συσχετίσεων

Οι πρώτες δύο ηχογραφημένες λέξεις αναφοράς είναι "hahaha" και "meat" , και η τρίτη ηχογραφημένη λέξη είναι "hahaha" και πάλι . Από το σχήμα 10, η πρώτη απεικόνιση είναι για τη συσχέτιση μεταξύ του τρίτου καταγεγραμμένου σήματος ομιλίας "hahaha" και του σήματος αναφοράς "hahaha". Η δεύτερη απεικόνιση είναι για τη συσχέτιση μεταξύ του τρίτου καταγεγραμμένου σήματος ομιλίας "hahaha" και του σήματος αναφοράς "κρέας". Αφού η τρίτη ηχογραφημένη λέξη είναι "hahaha" , έτσι η πρώτη απεικόνιση είναι πραγματικά πιο συμμετρική και ομαλότερη από τη δεύτερη απεικόνιση.

Στα μαθηματικά, αν θέσουμε την λειτουργία του φάσματος συχνοτήτων ως συνάρτηση $f(x)$, σύμφωνα με την αξονική συμμετρική ιδιότητα είναι: για τη συνάρτηση $f(x)$, εάν τα x_1 και x_3 συμμετρικοί άξονες για το $x = x_2$, τότε $f(x_1) = f(x_3)$. Για τη σύγκριση αναγνώρισης φωνής, μετά τον υπολογισμό της συσχέτιση των δύο καταγεγραμμένων φασμάτων συχνότητας, πρέπει να βρεθεί η θέση της μέγιστης τιμής της συσχετίσεως και να χρησιμοποιηθούν οι τιμές δεξιά στη θέση μέγιστης τιμής για να μειωθούν οι τιμές αριστερά

στη θέση μέγιστης τιμής. Παίρνουμε την απόλυτη τιμή της διαφοράς και βρίσκουμε το μέσο τετραγωνικό σφάλμα της απόλυτης τιμής. Αν δύο σήματα ταιριάζουν καλύτερα, τότε η συσχέτιση είναι πιο συμμετρική. Και αν η συσχέτιση είναι περισσότερο συμμετρική, τότε το μέσο τετραγωνικό σφάλμα θα πρέπει να είναι μικρότερο. Με τη σύγκριση αυτού του σφάλματος, το σύστημα αποφασίζει ποια λέξη αναφοράς καταγράφεται την τρίτη φορά. Οι κωδικοί για το μέρος αυτό βρίσκονται στο Παράρτημα.

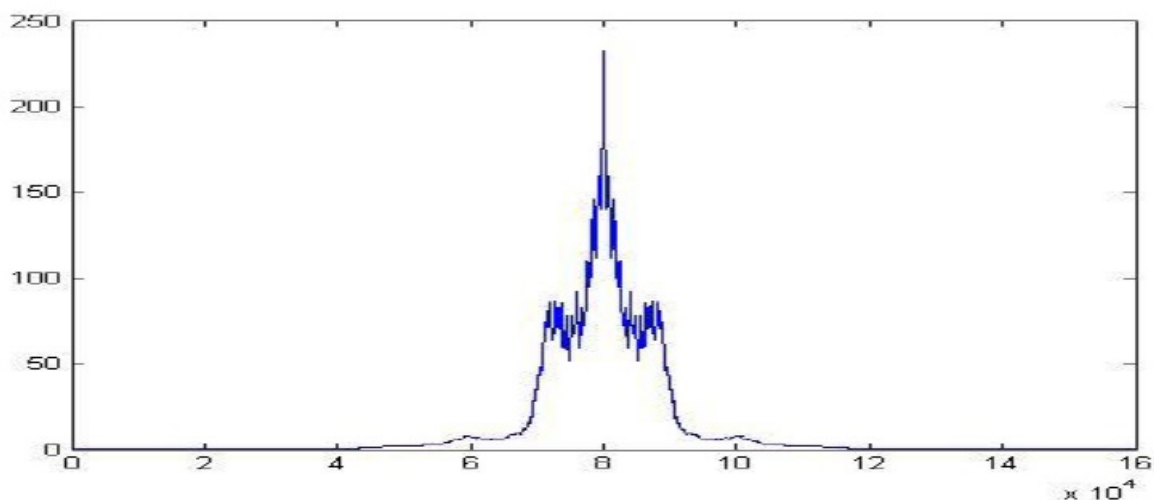
2.3.3 Ο αλγόριθμος αυτοσυσχέτισης

Στο προηγούμενο μέρος, αναφέραμε τον αλγόριθμο συσχέτισης. Βλέπε την εξίσωση (24), η αυτοσυσχέτιση μπορεί να θεωρηθεί ως ο υπολογισμός της συσχέτισης του ίδιου του σήματος, αντί δύο διαφορετικών σημάτων. Αυτός είναι ο ορισμός της αυτοσυσχέτισης στο MATLAB. Ο αλγόριθμος αυτοσυσχέτισης είναι ο αλγόριθμος για να μετρήσουμε πόσο το σήμα είναι αυτο-συσχετιζόμενο με τον εαυτό του.

Η εξίσωση για την αυτό-συσχέτιση είναι:

$$r_x(k) = r_{xx}(k) = \sum_{n=-\infty}^{\infty} x(n)x(n+k) \quad (25)$$

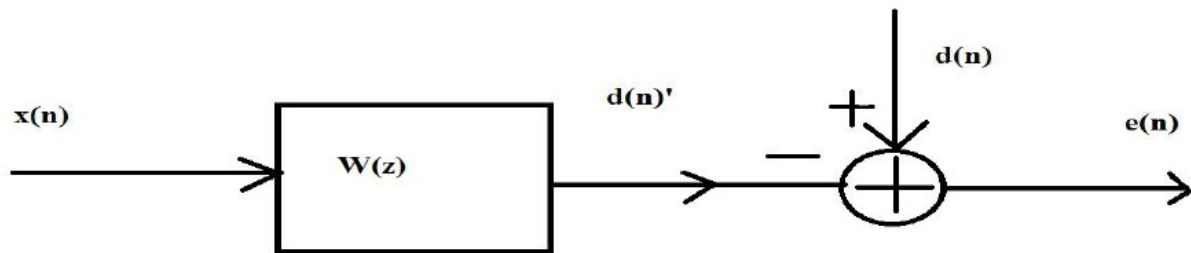
Η παρακάτω εικόνα είναι η γραφική παράσταση της αυτοσυσχέτισης του φάσματος συχνοτήτων $|X(\omega)|$



Σχήμα 11: Η αυτοσυσχέτιση για $|X(\omega)|$

2.3.4 Το φίλτρο FIR Wiener

Το φίλτρο FIR Wiener χρησιμοποιείται για την εκτίμηση του επιθυμητού σήματος $d(n)$ από τη διαδικασία παρακολούθησης $x(n)$ για να πάρει το εκτιμώμενο σήμα $d(n)'$. Υποτίθεται ότι $d(n)$ και $x(n)$ συσχετίζονται και το σφάλμα της εκτίμησης είναι το $e(n) = d(n) - d(n)'$. Το φίλτρο FIR Wiener λειτουργεί όπως το σχήμα που φαίνεται παρακάτω :



Σχήμα 12: Το φίλτρο Wiener

Από σχ.12, το σήμα εισόδου του Wiener φίλτρου είναι $x(n)$. Ας υποθέσουμε ότι το φίλτρο συντελεστών είναι $w(n)$. Έτσι, η έξοδος $d(n)'$ είναι η συνέλιξη του $x(n)$ και $w(n)$:

$$d(n)' = w(n) * x(n) = \sum_{l=0}^{p-1} w(l)x(n-l) \quad (26)$$

Στη συνέχεια, το σφάλμα εκτίμησης είναι :

$$e(n) = d(n) - d(n)' = d(n) - \sum_{l=0}^{p-1} w(l)x(n-l) \quad (27)$$

Ο σκοπός του Wiener φίλτρου είναι να επιλέξει τον κατάλληλο βαθμό του φίλτρου και να βρει τους συντελεστές του φίλτρου με τους οποίους το σύστημα μπορεί να πάρει την καλύτερη εκτίμηση. Με άλλα λόγια, με τους κατάλληλους συντελεστές, το σύστημα μπορεί να ελαχιστοποιήσει το μέσο τετραγωνικό σφάλμα :

$$\xi = E \left\{ |e(n)|^2 \right\} = E \left\{ |d(n) - d(n)'|^2 \right\} \quad (28)$$

Ελαχιστοποιείται το μέσο τετραγωνικό σφάλμα, προκειμένου να πάρει τους κατάλληλους συντελεστές φίλτρου, μια επαρκής μέθοδος που υπάρχει για να γίνει αυτό είναι να πάρει την παράγωγο του ξ να είναι μηδενική σε σχέση με το $w^*(k)$. Όπως στην ακόλουθη εξίσωση:

$$\frac{\partial \xi}{\partial w^*(k)} = \frac{\partial}{\partial w^*(k)} E \{ e(n)e^*(n) \} = E \left\{ e(n) \frac{\partial e^*(n)}{\partial w^*(k)} \right\} = 0 \quad (29)$$

Από την εξίσωση (27) και την εξίσωση (29) , γνωρίζουμε ότι:

$$\frac{\partial e^*(n)}{\partial w^*(k)} = -x^*(n-k) \quad (30)$$

Έτσι, η εξίσωση (29) γίνεται :

$$\frac{\partial \xi}{\partial w^*(k)} = E \left\{ e(n) \frac{\partial e^*(n)}{\partial w^*(k)} \right\} = -E \{ e(n)x^*(n-k) \} = 0 \quad (31)$$

Τότε έχουμε:

$$E \{ e(n)x^*(n-k) \} = 0 \quad , k=0, 1, \dots, p-1 \quad (32)$$

Η εξίσωση (32) είναι γνωστή ως αρχή ορθογωνιότητας ή το θεώρημα σχεδίασης [6] .

Από την εξίσωση (27) , έχουμε:

$$E \{ e(n)x^*(n-k) \} = E \left\{ \left[d(n) - \sum_{l=0}^{p-1} w(l)x(n-l) \right] x^*(n-k) \right\} = 0 \quad (33)$$

Η αναδιάταξη της εξίσωσης (33) :

$$E \{ d(n)x^*(n-k) \} - E \left\{ \sum_{l=0}^{p-1} w(l)x(n-l)x^*(n-k) \right\} = r_{dx} - \sum_{l=0}^{p-1} w(l)r_x(k-l) = 0 \quad (34)$$

Τέλος , η εξίσωση έχει ως εξής:

$$\sum_{l=0}^{p-1} w(l)r_x(k-l) = r_{dx} \quad k=0, 1, \dots, p-1 \quad (35)$$

Με $r_x(k) = r_x^*(-k)$, η εξίσωση μπορεί να γραφεί σε μορφή πίνακα :

$$\begin{bmatrix} r_x(0) & r_x^*(1) & \cdots & r_x^*(p-1) \\ r_x(1) & r_x(0) & \cdots & r_x^*(p-2) \\ r_x(2) & r_x(1) & \cdots & r_x^*(p-3) \\ \vdots & \vdots & & \vdots \\ r_x(p-1) & r_x(p-2) & \cdots & r_x^*(0) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \\ w(2) \\ \vdots \\ w(p-1) \end{bmatrix} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \\ r_{dx}(2) \\ \vdots \\ r_{dx}(p-1) \end{bmatrix} \quad (36)$$

Η εξίσωση του πίνακα (36) είναι στην πραγματικότητα η εξίσωση Wiener - Hopf [6] :

$$R_x w = r_{dx} \quad (37)$$

Στην παρούσα διατριβή, η εξίσωση Wiener - Hopf μπορεί να λειτουργήσει για την αναγνώριση φωνής. Από την εξίσωση (37), το σήμα εισόδου $x(n)$ και το επιθυμητό σήμα $d(n)$ είναι τα μόνα πράγματα που πρέπει να ξέρουμε. Στη συνέχεια, χρησιμοποιώντας το $x(n)$ και $d(n)$ βρίσκουμε το r_{dx} αυτοσυσχέτισης. Ταυτόχρονα, χρησιμοποιώντας το $x(n)$ βρίσκουμε το $r_x(n)$ αυτόματης συσχέτισης και χρησιμοποιώντας το $r_x(n)$ σχηματίζουμε τον πίνακα R_x στο MATLAB . Όταν έχουμε την R_x και r_{dx} , μπορεί να βρεθεί άμεσα από τους συντελεστές του φίλτρου. Με τους συντελεστές του φίλτρου, μπορεί συνεχώς να πάρει το ελάχιστο μέσο τετραγωνικό-σφάλμα^ξ. Από τις εξισώσεις (27) , (28) και (32), το ελάχιστο μέσο τετραγωνικό-σφάλμα^ξ είναι :

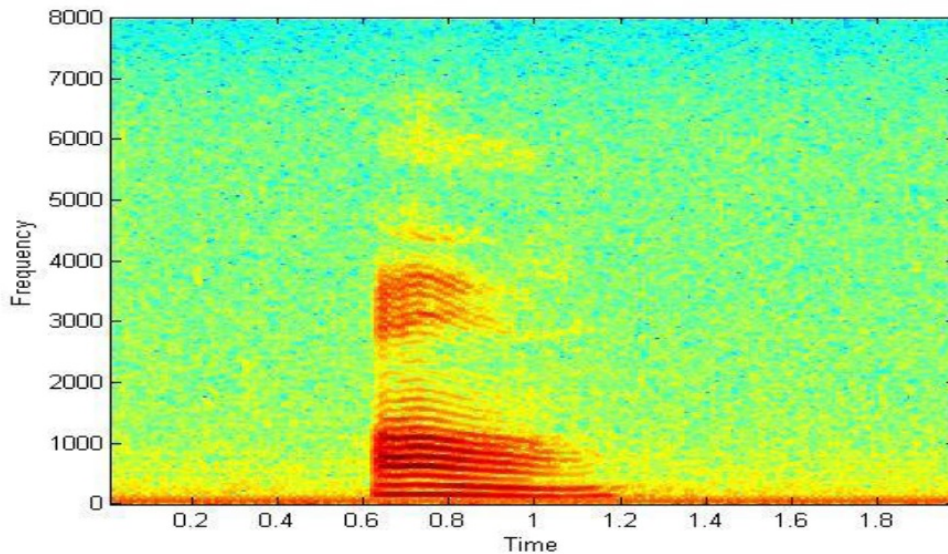
$$\xi_{\min} = E \{ e(n)d^*(n) \} = E \left\{ \left[d(n) - \sum_{l=0}^{p-1} w(l)x(n-l) \right] d^*(n) \right\} = r_d(0) - \sum_{l=0}^{p-1} w(l)r_{dx}^*(l) \quad (38)$$

Εφαρμόζουμε τη θεωρία του Wiener φίλτρου για την αναγνώριση φωνής. Αν θέλουμε να χρησιμοποιήσουμε την Wiener-Hopf εξίσωση, είναι απαραίτητο να γνωρίζουμε δύο δεδομένες συνθήκες : το ένα είναι το επιθυμητό σήμα $d(n)$ και το άλλο είναι το σήμα εισόδου $x(n)$. Σε αυτή την εργασία, υποτίθεται ότι τα καταγεγραμμένα σήματα είναι ευρείας-αίσθησης στάσιμες διαδικασίες. Τότε τα δύο πρώτα καταγεγραμμένα σήματα αναφοράς μπορεί να χρησιμοποιηθούν ως σήματα εισόδου $x_1(n)$ και $x_2(n)$. Το τρίτο καταγεγραμμένο σήμα ομιλίας μπορεί να χρησιμοποιηθεί ως το επιθυμητό σήμα $d(n)$. Πρέπει να βρούμε την καλύτερη εκτίμηση του επιθυμητού σήματος στο φίλτρο Wiener. Έτσι, η διαδικασία της εφαρμογής του Wiener φίλτρου για την αναγνώριση φωνής μπορεί να θεωρηθεί η χρήση των δύο πρώτων καταγεγραμμένων σημάτων αναφοράς για την εκτίμηση του τρίτου καταγεγραμμένου επιθυμητού σήματος . Αφού το ένα από τα δύο σήματα αναφοράς $x_1(n)$, $x_2(n)$ είναι καταγεγραμμένα για την ίδια λέξη όπως η λέξη που καταγράφηκε την τρίτη φορά. Έτσι, χρησιμοποιώντας το ένα από τα δύο σήματα αναφοράς που είναι καταγεγραμμένο για την ίδια λέξη στην τρίτη καταγραφή του χρόνου που είναι το σήμα εισόδου του Wiener φίλτρου θα έχει το μικρότερο εκτιμώμενο ελάχιστο μέσο τετραγωνικό σφάλμα ξ_{\min} σύμφωνα με την εξίσωση (38) . Μετά τον ορισμό των ρόλων των τριών σημάτων που καταγράφονται στο σχεδιασμένο σύστημα 2, το επόμενο βήμα είναι απλά να βρούμε τις αυτοσυσχετίσεις των σημάτων αναφοράς , τα οποία είναι $r_{x_1}(n)$, $r_{x_2}(n)$ και να βρούμε τις συσχετίσεις για το τρίτο καταγραφόμενο σήμα φωνής με τα δύο πρώτα καταγραφόμενα σήματα αναφοράς, τα οποία είναι $r_{dx_1}(n)$, $r_{dx_2}(n)$. Και τη χρήση $r_{x_1}(n)$, $r_{x_2}(n)$ για την

κατασκευή του πίνακα R_{x_1} , R_{x_2} . Τέλος, σύμφωνα με την εξίσωση Wiener - Hopf (37) , υπολογίζουμε τους συντελεστές του φίλτρου, τόσο για τα δύο σήματα αναφοράς αλλά και να βρούμε τις μέσες τιμές του ελάχιστου μέσου τετραγωνικού σφάλματος σε σχέση με τους δύο συντελεστές του φίλτρου. Συγκρίνοντας τα ελάχιστα μέσα τετραγωνικά σφάλματα, το σύστημα θα δώσει την απόφαση ποιο από το ένα από τα δύο καταγραφόμενα σήματα αναφοράς θα είναι η λέξη που καταγράφεται την τρίτη φορά. Η καλύτερη εκτίμηση θα είναι η μικρότερη μέση τιμή ξ_{min} .

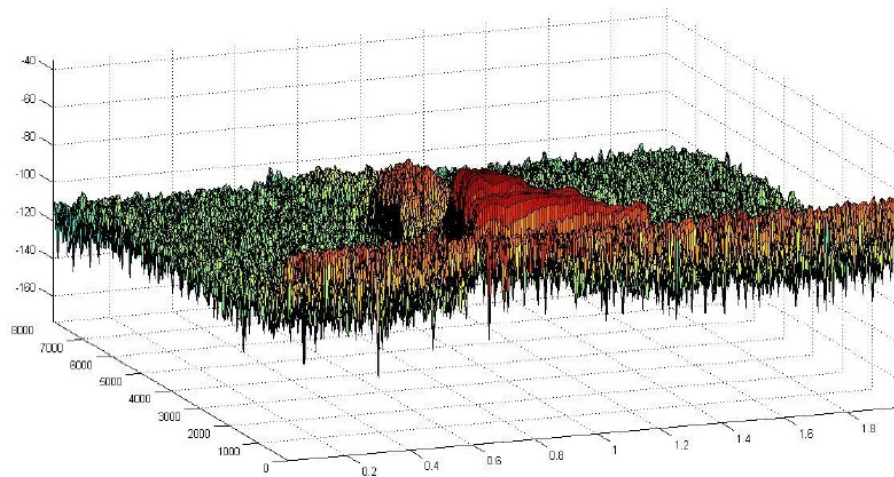
2.3.5 Χρήση Φασματογραφήματος στο MATLAB

Το φασματογράφημα είναι μια αποτύπωση χρόνου-συχνότητας το οποίο περιέχει ενέργεια κατανομής πυκνότητας ταυτόχρονα σε σχέση με τους δύο άξονες συχνότητας και του άξονα του χρόνου. Στο MATLAB, είναι εύκολο να πάρουμε το φασματόγραμμα του σήματος φωνής ορίζοντας κάποιες μεταβλητές: τη συχνότητα δειγματοληψίας, το μήκος του χρόνου του μετασχηματισμού Fourier και το μήκος του παραθύρου. Από τα προηγούμενα του παρόντος κειμένου, έχουν αναφερθεί ο DFT και FFT. Ο STFT είναι, πρώτον, για τη χρήση της συνάρτησης παραθύρου για να περικόψει το σήμα στο πεδίο του χρόνου, γεγονός που καθιστά τον χρονικό άξονα, σε διάφορα τμήματα. Αν το παράθυρο είναι ένα διάνυσμα, τότε ο αριθμός των τμημάτων είναι ίσο με το μήκος του παραθύρου. Τότε υπολογίζεται ο μετασχηματισμός Fourier της περικομμένης αλληλουχίας με καθορισμένο FFT μήκος (nfft). Το σχήμα 13 παρακάτω είναι το φασματόγραμμα για το καταγραφόμενο σήμα ομιλίας στο MATLAB, με οριζόμενα $fs = 16000$, $nfft = 1024$, το μήκος του παραθύρου είναι 512, και το μήκος της επικάλυψης είναι 380. Είναι αναγκαίο να σημειωθεί ότι το μήκος του παραθύρου πρέπει να είναι μικρότερο ή ίσο από το $\frac{1}{2}$ του μήκους του STFT (nfft) κατά τον προγραμματισμό στο MATLAB.



Σχήμα 13: Το φασματογράφημα της ηχογραφημένης λέξης "on"

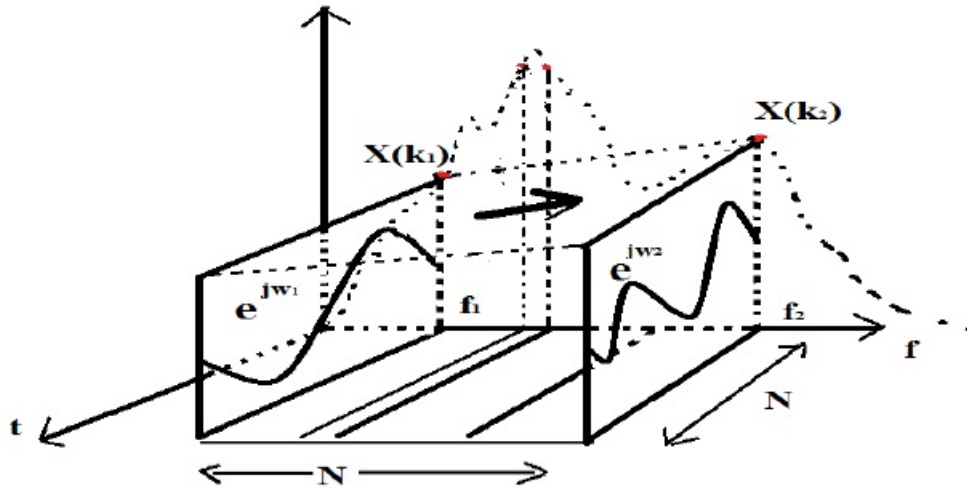
Από το Σχήμα 13, ο Χ-άξονας είναι ο χρόνος και ο Υ- άξονας είναι η συχνότητα. Η ανάλυση του χρώματος αντιπροσωπεύει την κλίση της κατανομής ενέργειας. Το βαθύτερο χρώμα είναι η υψηλότερη κατανομή ενέργειας στην εν λόγω ζώνη. Από το Σχήμα 13, η μεγαλύτερη ενέργεια βρίσκεται στην ζώνη χαμηλής συχνότητας. Το παρακάτω σχήμα είναι σχεδιασμένο στο MATLAB για ένα τρισδιάστατο φασματογράφημα της ίδιας καταγραφόμενης λέξης "on".



Σχήμα 14: Το τρισδιάστατο φασματογράφημα της ηχογραφημένης λέξης "on"

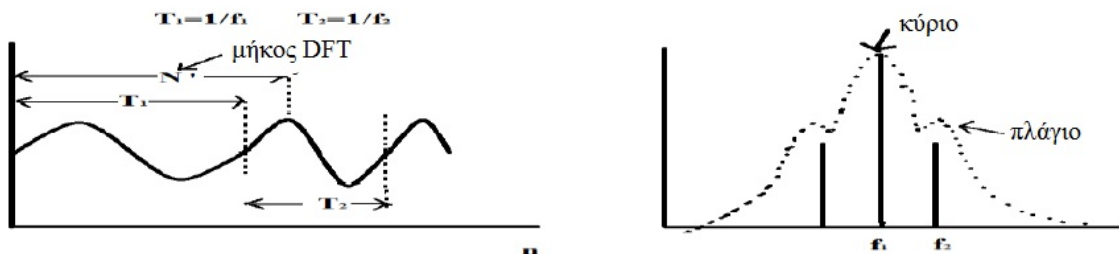
Βασικά το Σχήμα 14 είναι ακριβώς το ίδιο με το Σχήμα 13, εκτός του ότι μπορούμε να δούμε την κατανομή ισχύος από τα ύψη των ηλεκτροκίνητων «βουνών». Τώρα εξετάζοντας την αναγνώριση φωνής, το θέμα εδώ δεν είναι πώς μοιάζει η γραφική παράσταση, αλλά η λειτουργία της λήψης του φασματογραφήματος. Για τη διαδικασία της λήψης φασματογράμματος στο MATLAB χρησιμοποιούμε το παράθυρο για να περικόψουμε το

χρόνο σε μικρά μέρη του χρόνου και υπολογίζουμε το STFT. Γι 'αυτό είναι σκόπιμο να χρησιμοποιηθεί η λειτουργία του φασματογράμματος στο MATLAB για να πάρουμε το φάσμα συχνοτήτων καθαρότερο και πιο αξιόπιστο. Πρώτον δείτε την εικόνα όπως παρακάτω:



Σχήμα 15: Τρισδιάστατη σχέση γραφήματος του DFT

Από το Σχήμα 15, το φάσμα στο πεδίο συχνότητας μπορεί να θεωρηθεί σαν αναπόσπαστο ή σαν άθροισμα όλων των επιπέδων των στοιχείων της συχνότητας. Για κάθε επίπεδο συνιστώσας συχνότητας, το ύψος του επιπέδου του στοιχείου της συχνότητας είναι ακριβώς το σήμα σε όλο το πεδίο του χρόνου πολλαπλασιάζοντας το συσχετιζόμενο παράγοντα φάσης συχνότητας e^{jw} . Από το Σχήμα 15, εάν το σήμα στο πεδίο του χρόνου είναι ένα καθαρό περιοδικό σήμα, τότε η συνιστώσα συχνότητας θα είναι ένα τέλειο ενιαίο επίπεδο συνιστώσας χωρίς να αγγίζει άλλο επίπεδο συχνότητας, όπως e^{jw_1} και e^{jw_2} επίπεδα που φαίνονται στο Σχήμα 15. Είναι σταθερές και δεν θα επηρεάσει το ένα το άλλο. Αλλά αν το σήμα είναι απεριοδικό σήμα, δείτε το σχήμα παρακάτω :



Σχήμα 16: Μη περιοδικό σήμα παράγει τη διαρροή από DFT για μεγάλη ακολουθία μήκους

Από το σχήμα 16 , το σήμα είναι ένα απериοδικό σήμα, η συχνότητα αλλάζει μετά από μια περίοδο T_1 . Αν εμείς εξακολουθούμε να αντιμετωπίζουμε αυτό το απериοδικό σήμα ως ένα ενιαίο επίπεδο, και άμεσα υπολογίσουμε το DFT του, τότε το αποτέλεσμα του DFT για ακολουθία δεδομένων με το μήκος του $N-T_1$ είναι να μετακινήσει την ισχύ της συνιστώσας συχνότητας του στη συνιστώσα συχνότητα η οποία έχει την ίδια συχνότητα με αυτή της ακολουθίας των δεδομένων. Το αποτέλεσμα της DFT είναι ένα φάσμα ισχύος. Η συμπεριφορά της ισχύος που ρέει ονομάζεται διαρροή . Δεδομένου ότι το σήμα είναι διακριτό στην πραγματική επεξεργασία σήματος, μία θέση χρόνου έχει μια κατάσταση τιμής. Και κατά την εγγραφή του σήματος φωνής , το σήμα φωνής είναι ένα σύνθετο σήμα, το οποίο περιέχει πολλές συχνότητες . Έτσι, το καταγεγραμμένο σήμα ομιλίας θα είναι απериοδικό σήμα λόγω της αλλαγής της προφοράς, θα έχει τη διαρροή στο φάσμα συχνοτήτων, συμπεριλαμβανομένης της ισχύος της παρεμβολής θορύβου. Από το σχήμα 16, μετά από το χρόνο T_1 η συχνότητα του σήματος αλλάζει κατά το χρονικό διάστημα T_2 . Όσο η συχνότητα αλλάζει από απериοδικό σήμα, το φάσμα δεν θα είναι ομαλό, το οποίο δεν είναι καλό για την ανάλυση.

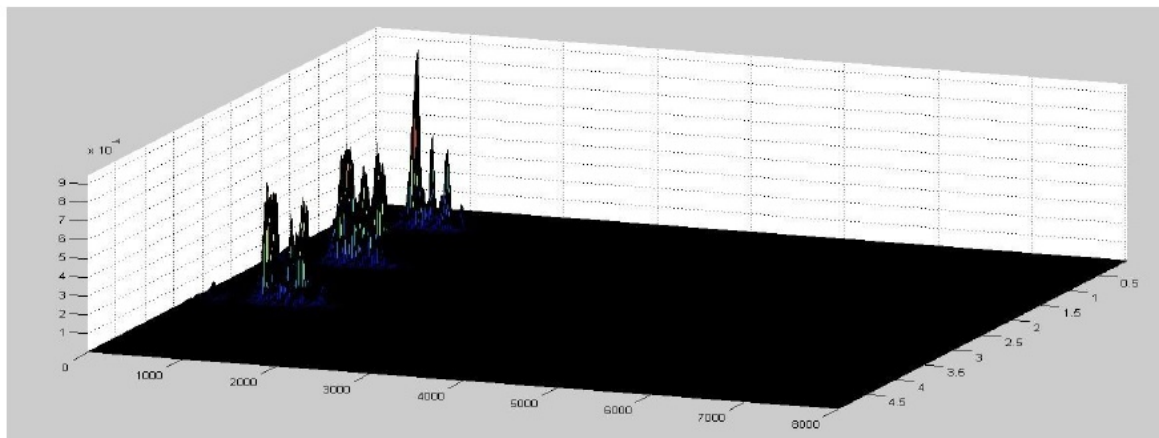
Χρησιμοποιώντας τα παράθυρα μπορούν να βελτιώσουν την κατάσταση αυτή. Τα παράθυρα είναι λειτουργίες σταθμίσεως που εφαρμόζονται σε δεδομένα για να μειωθεί η διαρροή του φάσματος που συνδέεται με πεπερασμένα διαστήματα παρατήρησης [8]. Είναι καλύτερο να χρησιμοποιήσουμε το παράθυρο για να περικόψουμε τη μακρά ακολουθία σήματος σε σύντομη χρονική ακολουθία. Για σύντομη χρονική ακολουθία, το σήμα μπορεί να θεωρηθεί ως απериοδικό σήμα και το σήμα έξω από το παράθυρο θεωρείται ως μηδενικές τιμές. Στη συνέχεια, γίνεται ο υπολογισμός του DFT ή FFT για τα περικομμένα δεδομένα σήματος. Αυτό ονομάζεται σύντομο χρονικό διάστημα μετασχηματισμού Fourier (STFT). Συνεχίζοντας κατά μήκος του άξονα του χρόνου, έως ότου το παράθυρο να περικοπεί σε όλο το φάσμα.

Με αυτόν τον τρόπο, το παράθυρο δεν θα μειώσει μόνο την διαρροή της συνιστώσας συχνότητας, αλλά επίσης κάνει το φάσμα ομαλότερο.

Αφού η κινούμενη βαθμίδα του παραθύρου είναι πάντοτε μικρότερη από το μήκος του παραθύρου έτσι το αποτέλεσμα φάσματος θα έχει τις επικαλύψεις. Οι επικαλύψεις δεν είναι κακές για την ανάλυση. Όσο περισσότερες επικαλύψεις, τόσο καλύτερη ανάλυση του STFT, που σημαίνει ότι το φάσμα που προκύπτει είναι πιο αξιόπιστο.

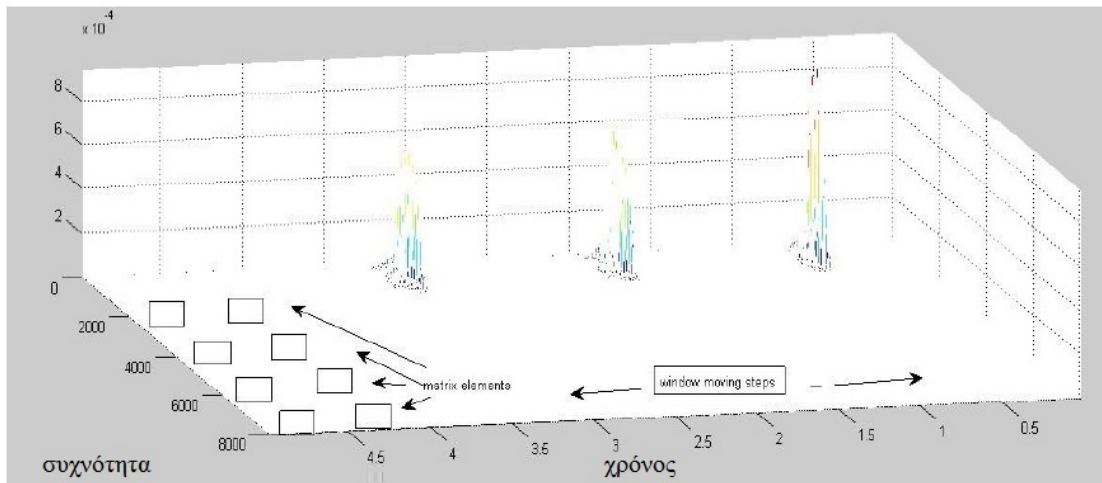
Χρησιμοποιώντας τη λειτουργία φασματογράμματος στο MATLAB μπορεί να ολοκληρώσει αυτή τη διαδικασία, η οποία επιστρέφει έναν πίνακα με τη χρήση "specgram" λειτουργίας σε περιβάλλον MATLAB. Έτσι, το "specgram" μπορεί να χρησιμοποιηθεί απευθείας ως λειτουργία "παράθυρο" για να πάρουμε το φιλτραρισμένο σήμα ομιλίας. Μετά τη χρήση του "specgram", μπορούμε να πάρουμε χρήσιμες και αξιόπιστες πληροφορίες των καταγεγραμμένων σημάτων τόσο στο πεδίο του χρόνου όσο και στο πεδίο της συχνότητας. Το επόμενο βήμα που πρέπει να εξεταστεί είναι απλά να συγκρίνουμε τα φάσματα του τρίτου καταγεγραμμένου σήματος με τα δύο πρώτα καταγεγραμμένα σήματα αναφοράς υπολογίζοντας την συσχέτιση ή χρησιμοποιώντας το σύστημα φίλτρου Wiener όπως παρουσιάσαμε προηγουμένως. Αυτό είναι το πώς λειτουργεί το φασματογράμμα για την αναγνώριση φωνής σε αυτή τη διατριβή.

Κατά τη χρήση του "s = specgram (r, NFFT , fs , Hanning)" ως εντολή στο MATLAB, θα επιστρέψει ένα πίνακα, στον οποίο τα στοιχεία είναι όλα μιγαδικοί αριθμοί. Χρησιμοποιούμε το MATLAB για να σχεδιάσουμε το φασματογράμμα για καλύτερη κατανόηση.



Σχήμα 17 : Το φασματογράφημα της λέξης " ha ... ha ... ha "

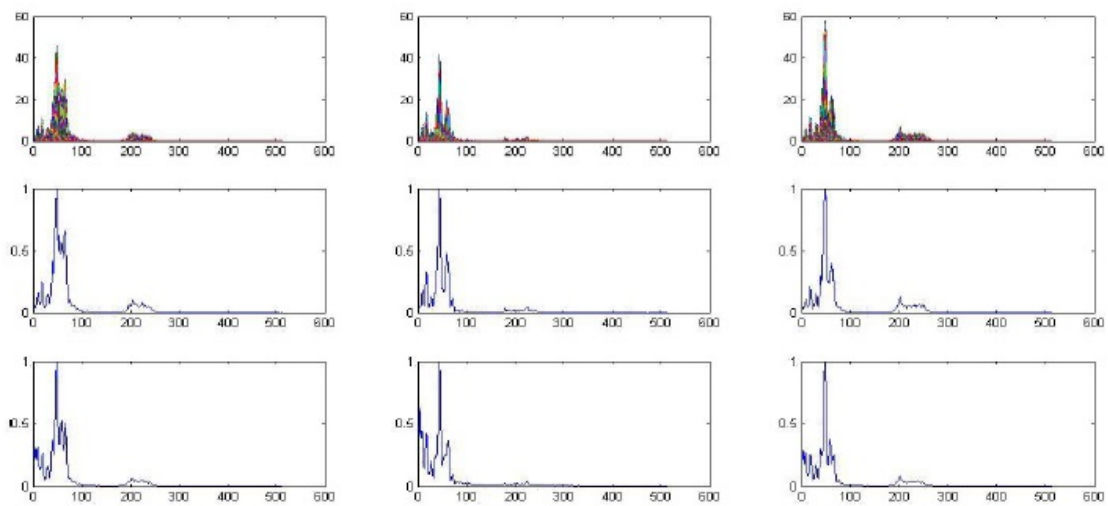
Για να είναι καλύτερη η κατανόηση του σχήματος, τροποποιούμε το σχήμα, όπως παρακάτω:



Σχήμα 18: Το τροποποιημένο σχήμα του σχήματος 17

Από Σχ.18, το μήκος του διανύσματος της κάθε σειράς είναι σχετικό προς τα κινητά στάδια του παραθύρου. Με τον έλεγχο των μεταβλητών πληροφοριών στο MATLAB αυτού παραδείγματος, ο πίνακας είναι ένας πίνακας 513×603 . Δεδομένου ότι η συχνότητα δειγματοληψίας για το σύστημα καταγραφής έχει οριστεί ως $f_s = 16000$, έτσι ώστε το μήκος του σήματος φωνής είναι $16000 \times 5 = 80000$ (που καταγράφονται σε 5 δευτερόλεπτα). Και το μήκος του παραθύρου Hanning ορίζεται ως 512. Η ρύθμιση του μήκους επικάλυψης είναι 380. Και για τη DFT / FFT περιοδική επέκταση, η συνάρτηση παραθύρου υπολογίζει στην πραγματικότητα το μήκος του $512 + 1 = 513$. Έτσι το κινούμενο μήκος είναι $513 - 380 = 133$. Έτσι, ο αριθμός του χρόνου υπολογίζεται ως $80000 / 133 \approx 602$, το οποίο είναι σχεδόν το ίδιο με τον αριθμό των στηλών για τον πίνακα στο MATLAB. Έχει αποδειχθεί ότι το κινούμενο παράθυρο χωρίζεται το χρονικό διάστημα του αρχικού σήματος από 80000 σε σύντομο χρονικό διάστημα 603. Έτσι, για να μετρήσουμε τον αριθμό των στηλών για το πλέγμα είναι στην πραγματικότητα για να δούμε τη θέση του χρόνου. Και η μέτρηση του αριθμού των γραμμών του πίνακα είναι στην πραγματικότητα για να δούμε τη θέση συχνότητας. Έτσι, για το στοιχείο $S_{ij} = A$ στον πίνακα, το "i" είναι η θέση συχνότητας (ο αριθμός των σειρών), και το "j" είναι η θέση του χρόνου (ο αριθμός των στηλών). "A" είναι το αποτέλεσμα FFT για αυτό το χρονικό βήμα. Από την προηγούμενη συζήτηση, ο FFT/DFT θα έχει ως αποτέλεσμα μιγαδικούς αριθμούς. Έτσι το "A" είναι ένας μιγαδικός αριθμός. Για να βρούμε το μέγεθος του φάσματος (ύψος του φάσματος) του FFT / DFT, πρέπει να πάρουμε την απόλυτη αξία, $|A|$. Υποθέτοντας ότι ο επιστρεφόμενος πίνακας είναι ένας πίνακας $M \times N$, κατά τη σύγκριση των φασμάτων μεταξύ του τρίτου καταγεγραμμένου

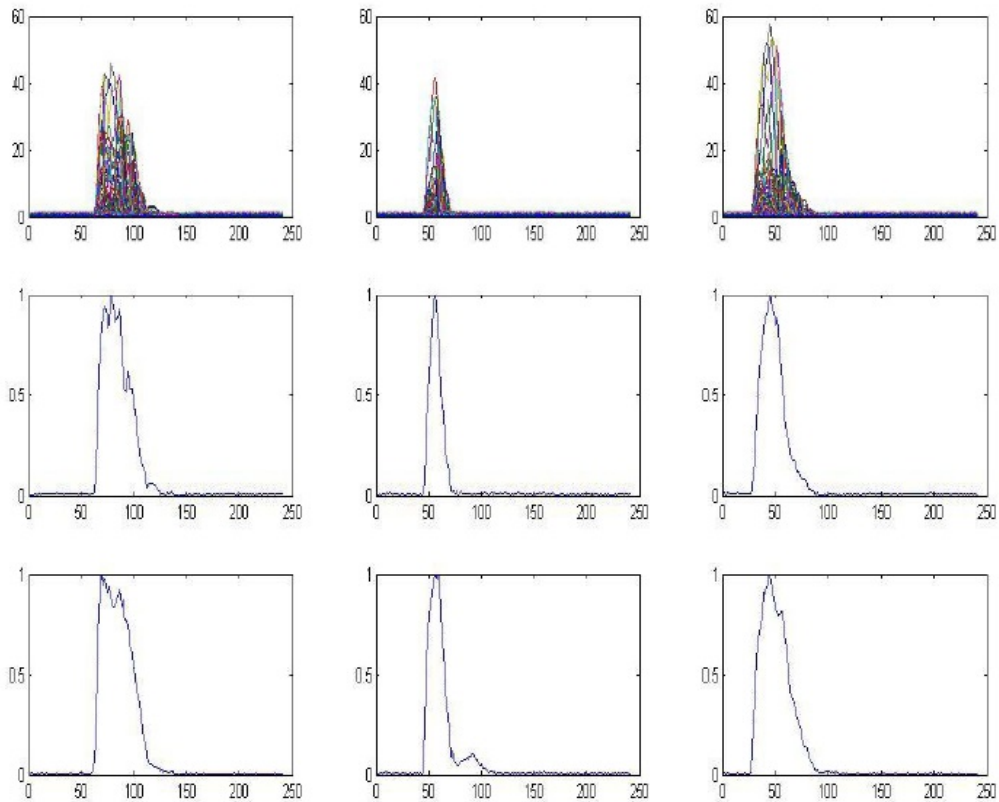
σήματος ομιλίας και των δύο πρώτων σημάτων εγγραφής αναφοράς, αυτό φαίνεται από τον άξονα συχνότητας (ο αριθμός των σειρών M). Για μία μόνο συχνότητα (ενιαία γραμμή) , το διάγραμμα γραμμής δεν περιέχει μόνο ένα στοιχείο. Αυτό σημαίνει ότι τα στοιχεία της σειράς θα έχουν όλα τη δική τους συμβολή φάσματος για διαφορετικά τμήματα χρόνου σε αυτή την ενιαία συχνότητα (σε αυτό τον αριθμό σειράς). Έτσι, βλέποντας τον άξονα συχνοτήτων, βλέπουμε τις τιμές N' σε αυτή τη συχνότητα ή N κορυφές σε αυτή τη συχνότητα. Έτσι, κατά τη σχεδίαση για ολόκληρη τη ζώνη συχνοτήτων της ομιλίας , το φάσμα είναι πραγματικά N επικαλυπτόμενα φάσματα. Εκτελώντας τον κώδικα του προγράμματος αυτής της διατριβής στο MATLAB , φαίνονται τα φάσματα ομιλίας για τις τρεις ηχογραφήσεις , όπως παρακάτω:



Σχήμα 19: Το φάσμα όπως φαίνεται από τον «άξονα συχνοτήτων»

Από το Σχήμα 19, οι γραφικές παραστάσεις της πρώτης σειράς σχεδιάστηκαν άμεσα από τις απόλυτες τιμές του πίνακα. Οι γραφικές παραστάσεις της δεύτερης σειράς λαμβάνοντας γραφικώς την μέγιστη τιμή για κάθε διάγραμμα γραμμής του πίνακα. Οι γραφικές παραστάσεις της τρίτης σειράς σχεδιάστηκαν λαμβάνοντας το άθροισμα των στοιχείων κάθε γραμμής του. Οι γραφικές παραστάσεις της πρώτης γραμμής και της δεύτερης, όπως φαίνεται , δεν είναι ακριβώς οι παραστάσεις για το πραγματικό φάσμα συχνοτήτων. Αφού είναι ακριβώς η μέγιστη τιμή για κάθε συχνότητα, έτσι ώστε η πληροφορία των φασμάτων είναι μόνο για τη στιγμή που το μέγεθος του φάσματος είναι το μέγιστο. Με τη λήψη του υπολογισμού άθροισης κάθε σειράς, οι πληροφορίες των φασμάτων είναι για το σύνολο των τμημάτων του χρόνου και η επίδραση του θορύβου θα μειωθεί. Έτσι, η τρίτη σειρά γραφημάτων είναι αναπαραστάσεις των πραγματικών φασμάτων. Από το Σχήμα 19, οι διαφορές μεταξύ των

γραφημάτων της τρίτης σειράς και τις γραφικές παραστάσεις των άλλων δύο σειρών δεν είναι προφανείς όταν παρίστανται γραφικώς σε φάσματα. Αλλά οι προφανείς διαφορές μπορούν να προβληθούν κατά τη σχεδίαση των σημάτων στο πεδίο του χρόνου. Αυτό φαίνεται στο σχήμα παρακάτω:



Σχήμα 20 : Τα σήματα φωνής όπως φαίνονται από τον «άξονα του χρόνου»

Από το Σχήμα 20, συγκρίνουμε τις γραφικές παραστάσεις της δεύτερης στήλης. Υπάρχει ένας κυματισμός στην τρίτη σειρά περίπου σε χρόνο 100. Αλλά δεν μπορούμε να δούμε αυτό το κυματισμό των καμπυλών στην πρώτη και δεύτερη σειρά. Αυτό οφείλεται στο υψηλότερο επίπεδο θορύβου από την κυμάτωση του σήματος φωνής. Με τη λήψη της λειτουργίας άθροισης, οι κυματισμοί του σήματος θα βγει από το επίπεδο του θορύβου. Μετά τη γραμμική κανονικοποίηση, το αποτέλεσμα θα είναι σαφέστερο. Έτσι, κατά τη σύγκριση των φασμάτων των σημάτων, που χρειάζεται για να συγκρίνουμε τα φάσματα άθροισης, τα οποία είναι μεγαλύτερης ακρίβειας και αξιόπιστα.

3° Κεφάλαιο

Βήματα προγραμματισμού και αποτελέσματα της προσομοίωσης

Σε αυτή την διατριβή, υπάρχουν δύο σχεδιασμένα συστήματα για την αναγνώριση φωνής. Και τα δύο αυτά συστήματα χρησιμοποιούν τη γνώση σύμφωνα με το τμήμα Θεωρία της εργασίας αυτής, που έχει αναφερθεί προηγουμένως. Καλέσαμε φίλους για να βοηθήσουν στη δοκιμή των δύο σχεδιασμένων συστημάτων. Τρέχοντας τους κωδικούς του συστήματος κάθε φορά στο MATLAB, το MATLAB θα ζητήσει από τον χρήστη να καταγράψει τα σήματα φωνής για τρεις φορές. Οι δύο πρώτες ηχογραφήσεις χρησιμοποιούνται ως σήματα αναφοράς. Η τρίτη ηχογράφηση χρησιμοποιείται ως σήμα στόχου. Οι αντίστοιχοι κωδικοί για τα δύο συστήματα μπορούν να βρεθούν στο Παράρτημα.

3.1 Βήματα προγραμματισμού

3.1.1 Βήματα προγραμματισμού για Σχεδιασμένο Σύστημα 1

- (1) Διαμορφώνουμε τις μεταβλητές και ρυθμίζουμε τη συχνότητα δειγματοληψίας $f_s = 16000$. Χρησιμοποιούμε την "wavrecord" εντολή για την εγγραφή τριών σημάτων φωνής. Κάνουμε τις δύο πρώτες ηχογραφήσεις ως τα σήματα αναφοράς. Κάνουμε την τρίτη φωνητικής εγγραφής ως το σήμα στόχου.
- (2) Χρησιμοποιούμε τη λειτουργία "spectogram" για την επεξεργασία των καταγεγραμμένων σημάτων και να πάρουμε τα σήματα.
- (3) Για τη μεταφορά των σημάτων των πινάκων για τις γραμμές και στήλες, παίρνουμε τη λειτουργία "sum" του πίνακα και παίρνουμε ένα διάνυσμα γραμμής για κάθε αποτέλεσμα άθροισμα της στήλης. Αυτό το διάνυσμα γραμμής είναι το σήμα του φάσματος συχνοτήτων.
- (4) Κανονικοποίηση των φασμάτων συχνοτήτων από την γραμμική κανονικοποίηση.
- (5) Κάνουμε τις συσχετίσεις για το τρίτο καταγεγραμμένο σήμα με τα δύο πρώτα καταγεγραμμένα σήματα αναφοράς χωριστά.

(6) Αυτό το στάδιο είναι σημαντικό αφού ο αλγόριθμος σύγκρισης προγραμματίζεται εδώ. Πρώτον, ελέγχουμε τη μετατόπιση της συχνότητας των συσχετίσεων.

Εδώ θα πρέπει να αναφερθεί ότι η μετατόπιση συχνότητας δεν είναι η πραγματική αλλαγή συχνότητας. Είναι η επεξεργασία συχνότητας στο MATLAB. Με τον ορισμό του φάσματος για την "NFFT", που είναι το μήκος του STFT που έχει προγραμματιστεί στο MATLAB, η συνάρτηση θα επιστρέψει ένα φάσμα συχνοτήτων που είναι η εκτίμηση "nfft". Αν "NFFT" είναι μονός αριθμός, έτσι ώστε ο πίνακας να επιστρέψει $(nfft+1)/2$ σειρές. Αν "nfft" είναι άρτιος αριθμός, τότε ο πίνακας θα επιστρέψει $(nfft/2)+1$ σειρές. Αυτά ορίζονται στο MATLAB. Σειρές του πίνακα "φασματογραφήματος" είναι οι περιοχές συχνοτήτων. Εάν η διαφορά μεταξύ των απόλυτων τιμών των μετατοπίσεων συχνότητας για τους δύο συσχετίσεις είναι μεγαλύτερη ή ίση από 2, τότε το σύστημα θα δώσει την κρίση μόνο με τη μετατόπιση συχνότητας. Όσο μικρότερη μετατόπιση συχνότητας τόσο καλύτερη αντιστοιχία. Εάν η διαφορά μεταξύ των απόλυτων τιμών των μετατοπίσεων συχνότητας είναι μικρότερη από 2, τότε η διαφορά μετατόπισης συχνότητας είναι άχρηστη, σύμφωνα με την εμπειρία από τις πολλές δοκιμές. Το σύστημα χρειάζεται συνεχώς να κάνει τη σύγκριση με τη συμμετρική ιδιότητα για τις συσχετίσεις των αντίστοιχων σημάτων. Ο αλγόριθμος της συμμετρικής ιδιότητας έχει αναφερθεί στο τμήμα 2.3.2. Σύμφωνα με τη συμμετρική ιδιότητα, το MATLAB θα δώσει το αποτέλεσμα.

3.1.2 Βήματα προγραμματισμού για Σχεδιασμένο Σύστημα 2

(1) Διαμορφώνουμε τις μεταβλητές και ρυθμίζουμε τη συχνότητα δειγματοληψίας $f_s = 16000$.

(2) Χρησιμοποιούμε τη λειτουργία "wavrecord" και ηχογραφούμε τα 3 σήματα φωνής. Κάνουμε τις δύο πρώτες ηχογραφήσεις ως σήματα αναφοράς. Κάνουμε την τρίτη φωνητική εγγραφή ως το σήμα φωνής στόχου.

(3) Χρησιμοποιούμε τη λειτουργία "spectogram" για την επεξεργασία των καταγεγραμμένων σημάτων και για να πάρουμε τον πίνακα με τα σήματα.

(4) Μεταφέρουμε τα σήματα του πίνακα για τις γραμμές και στήλες, και παίρνουμε τη λειτουργία "sum" του πίνακα και παίρνουμε σαν αποτέλεσμα το διάνυσμα γραμμής για κάθε

αποτέλεσμα αθροίσματος της στήλης. Αυτό το διάνυσμα γραμμής αντιπροσωπεύει το φάσμα συχνοτήτων.

(5) Κανονικοποίηση των φασμάτων συχνοτήτων από την γραμμική κανονικοποίηση.

(6) Σε αυτό το στάδιο, το σύστημα θα είναι διαφορετικό από το σύστημα 1. Το σύστημα προγραμματίζεται για τον τρόπο λειτουργίας του φίλτρου. Πρώτον, υπολογίζουμε την αυτοσυσχετίσεις των 3 σημάτων: τα πρώτα δύο καταγεγραμμένα σήματα αναφοράς και το τρίτο καταγεγραμμένο σήμα στόχο. Δεύτερον, ορίζουμε το συνολικό αριθμό εντολής σε 20. Και χρησιμοποιούμε ένα "for" βρόχο για την ανίχνευση κάθε αποτελέσματος εντολής. Καθορίζουμε το μήκος αυτοσυσχέτισης από N σε $N + p$. Από τον ορισμό της εξίσωσης Wiener φίλτρου (36), τα μεγέθη ή τα μήκη του πίνακα αυτοσυσχέτισης και του διανύσματος αυτοσυσχέτισης θα πρέπει και τα δύο να είναι p . Δεδομένου ότι η θέση N είναι η μέγιστη τιμή θέσης της αυτοσυσχέτισης, τότε « $r(N) = r(0)$ ». Αυτό εξηγείται στο τμήμα 2.3.2, στο οποίο παρουσιάσαμε την σχέση μεταξύ της μέγιστης τιμής και της θέσης για τη συσχέτιση. Μετά τον ορισμό R_x και r_{dx} , το επόμενο βήμα είναι άμεσα για τον υπολογισμό των συντελεστών του φίλτρου για κάθε σήμα αναφοράς.

(7) Μετά την εύρεση των συντελεστών του φίλτρου για κάθε σήμα αναφοράς, υπολογίζουμε το ελάχιστο μέσο τετραγωνικού σφάλματος για κάθε σήμα αναφοράς. Συγκρίνουμε τη μέση τιμή του ελάχιστου μέσου τετραγωνικού σφάλματος για την περιοχή από το 0 έως το 20. Η καλύτερη εκτίμηση θα πρέπει να έχει το μικρότερο ελάχιστο μέσο τετραγωνικό σφάλμα. Η θεωρία του φίλτρου Wiener έχει παρουσιαστεί στο τμήμα 2.3.4.

3.2 Αποτελέσματα προσομοίωσης

Σε αυτή την ενότητα, προσομοιώσαμε δύο σχεδιασμένα συστήματα με τη βοήθεια των φίλων που προέρχονται από διαφορετικές χώρες. Τρέχουμε το πρόγραμμα και καταγράφουμε τα τρία σήματα ομιλίας. Οι δύο πρώτες ηχογραφήσεις χρησιμοποιούνται ως σήματα αναφοράς. Η τρίτη ηχογράφηση χρησιμοποιείται ως σήμα αντικειμενικού στόχου για το οποίο το MATLAB θα πρέπει να δώσει την απόφαση. Στα ακόλουθα αποτελέσματα, χρησιμοποιούμε "σήματα αναφοράς" για τις δύο πρώτες εγγραφές και χρησιμοποιούμε "σήμα στόχου" για την τρίτη εγγραφή. Οι λέξεις στα εισαγωγικά ξεχωρίζουν για το περιεχόμενο των εγγραφών. Προσπαθήσαμε να δοκιμάσουμε τα σχεδιασμένα συστήματα και για τις δύο εύκολα αναγνωρίσιμες λέξεις και δύσκολα αναγνωρίσιμες λέξεις. «Από τον χρόνο 1 στο χρόνο 10,

‘on’» σημαίνει ότι ο χειριστής προσομοίωσε 10 φορές και η τρίτη λέξη εγγραφής είναι "on" στις πρώτες 10 προσομοιώσεις. Τόσο τα περιεχόμενα των λέξεων αναφοράς όσο και η λέξη στόχος είναι γνωστά, θέλουμε να δοκιμάσουμε αν το αποτέλεσμα που δίνεται από το MATLAB είναι σωστό. Τα αποτελέσματα των στατιστικών προσομοιώσεων θα υποβληθούν σε πίνακες και θα είναι επίσης σχεδιασμένα. Στο μέρος Αποτελέσματα Προσομοίωσης, μόνο τα αποτελέσματα θα παρουσιάζονται στο ακόλουθο περιεχόμενο. Οι σχετικοί πίνακες αποτελεσμάτων θα δίνονται στο Παράρτημα Β. Αφού προγραμματίσαμε στο MATLAB να σχεδιάσουμε σχήματα για κάθε σύστημα για να βοηθήσουμε την ανάλυση κατά τη διαδικασία προσομοίωσης σε κάθε στιγμή, και τα προκύπτοντα στοιχεία για κάθε σύστημα θα παίρνονται από τις ίδιες αρχές, έτσι ώστε να βάλουμε μόνο το σχήμα προσομοίωσης μία φορά κατά την έναρξη των αποτελεσμάτων της προσομοίωσης για κάθε σύστημα.

3.2.1 Αποτελέσματα προσομοίωσης για το σύστημα 1

(1) Τα στοιχεία των πρώτων αποτελεσμάτων στατιστικής προσομοίωσης για το σύστημα 1 έχει ως εξής :

Σήματα αναφοράς : "on" και " off" :

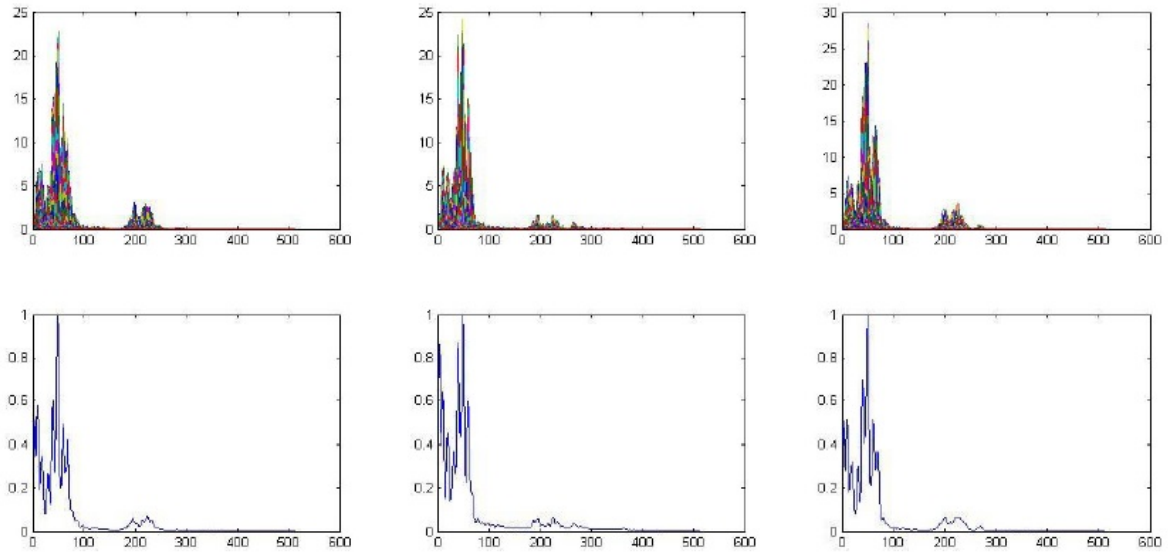
Στόχος του σήματος : Από τον χρόνο 1 έως 10 , "on" .

Από τον χρόνο 11 έως 20 , "off" .

Ομιλητής : “Δημήτρης” για τα δύο σήματα αναφοράς και το σήμα στόχο.

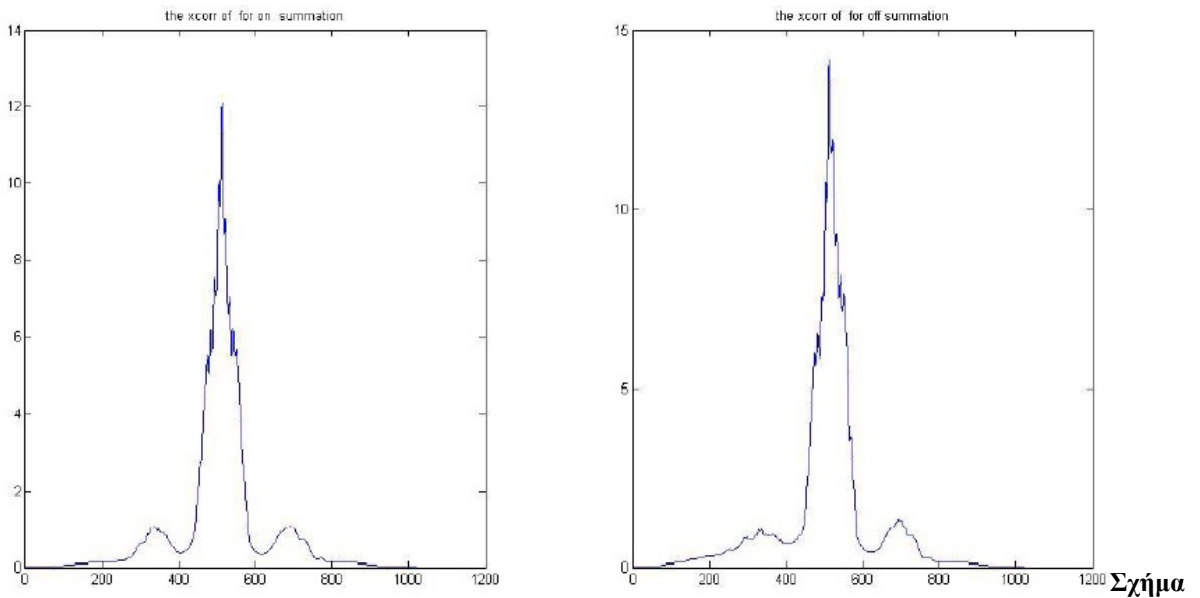
Περιβάλλον: σχεδόν κανένας θόρυβος

Το σχήμα 21 είναι για τα φάσματα συχνοτήτων για τα τρία καταγεγραμμένα σήματα, αλλά ο άξονας δεν είναι ο πραγματικός άξονας συχνότητας , αφού το σχήμα είναι από STFT. Πληροφορίες του STFT βρίσκονται στο τμήμα 2.3.5 . Η μέθοδος για να πάρουμε το ποσοστό αυτό βρίσκεται επίσης στο τμήμα 2.3.5



Σχήμα 21: Τα φάσματα συχνοτήτων για τα τρία σήματα ομιλίας: "on" , " off" , "on "

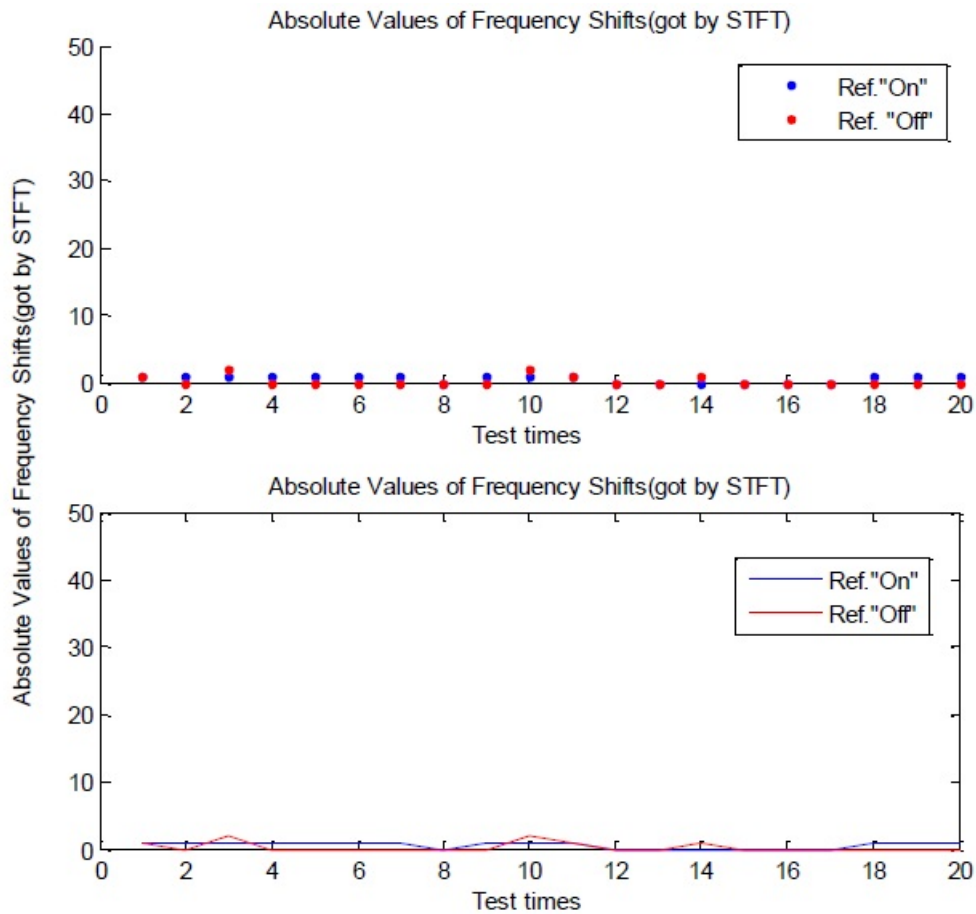
Ο αριθμός των συσχετίσεων μεταξύ του σήματος στόχου "on" και των σημάτων αναφοράς είναι όπως παρακάτω (το σήμα αναφοράς στο αριστερό σχήμα είναι "on". Το σήμα αναφοράς στο δεξί σχήμα είναι "off"):



22: Οι συσχετίσεις μεταξύ του σήματος στόχου "on" και του σήματος αναφοράς "off"

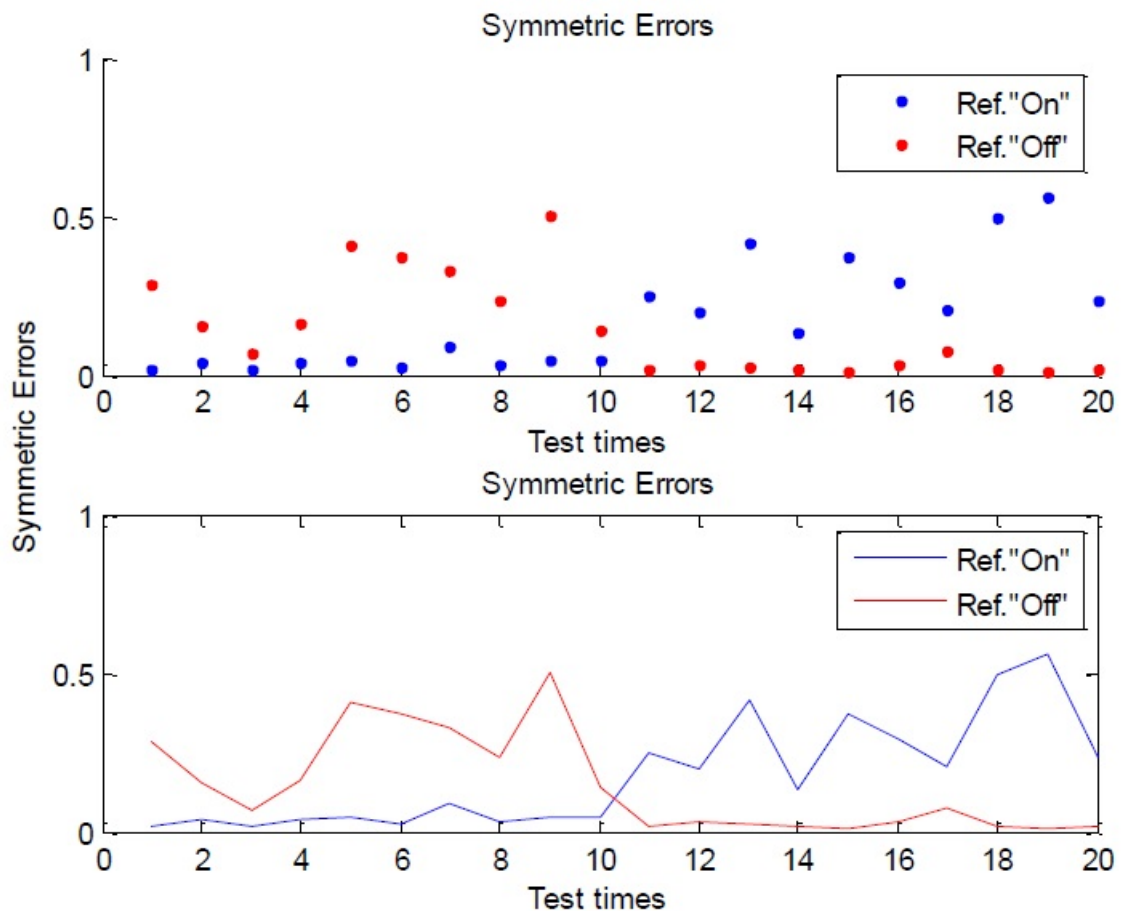
Όπως φαίνεται από το παραπάνω σχήμα 22, δεν υπάρχει μεγάλη διαφορά μεταξύ των δύο γραφημάτων , αφού η προφορά των "on" και " off" είναι παρόμοια.

Σύμφωνα με τα αποτελέσματα της προσομοίωσης (Πίνακας 1 , Παράρτημα Β), το αποτέλεσμα απεικόνισης της προσομοίωσης για την αλλαγή της συχνότητας είναι όπως παρακάτω:



Σχήμα 23 : Αλλαγή συχνότητας σε 20 προσομοιώσεις για τις λέξεις αναφοράς "on" και "off "

Από το Σχήμα 23, φαίνεται ότι είναι δύσκολο να δώσει τα αποτελέσματα με τις μετατοπίσεις συχνότητας. Οι μετατοπίσεις συχνότητας είναι πολύ κοντά μεταξύ των λέξεων "on" και "off". Έτσι, το σχεδιασμένο σύστημα θα δώσει το αποτέλεσμα σύμφωνα με τα συμμετρικά λάθη. Το αποτέλεσμα της απεικονιζόμενης προσομοίωσης για τα συμμετρικά σφάλματα είναι όπως παρακάτω:



Σχήμα 24: Συμμετρικά λάθη σε 20 προσομοιώσεις για τις λέξεις αναφοράς "on" και "off"

Όπως φαίνεται στο Σχήμα 24, η μπλε καμπύλη δίνει το αποτέλεσμα όταν η λέξη αναφοράς είναι "on". Η κόκκινη καμπύλη είναι το αποτέλεσμα όταν η λέξη αναφοράς είναι "off". Αφού οι πληροφορίες που δόθηκαν κατά την έναρξη, η λέξη στόχος είναι "on" στις πρώτες 10 προσομοιώσεις και η λέξη στόχος είναι "off" μετά τις 10 προσομοιώσεις. Από το σχήμα 24, φαίνεται ότι κατά τις πρώτες 10 προσομοιώσεις η καμπύλη της λέξης αναφοράς "on" έχει χαμηλότερη τιμή ενώ η καμπύλη της λέξης αναφοράς "off" μετά τις 10 έχει μικρότερη αξία. Τα αποτελέσματα έδειξαν ότι, όταν το σήμα ομιλίας αναφοράς και το σήμα ομιλίας στόχος ταιριάζουν, τα συμμετρικά σφάλματα είναι μικρότερα. Τα αποτελέσματα είναι απολύτως σωστά.

(2) Οι πληροφορίες των αποτελεσμάτων της δεύτερης προσομοίωσης για το σύστημα 1 έχει ως εξής:

Σήματα αναφοράς : "Door" και "Key" :

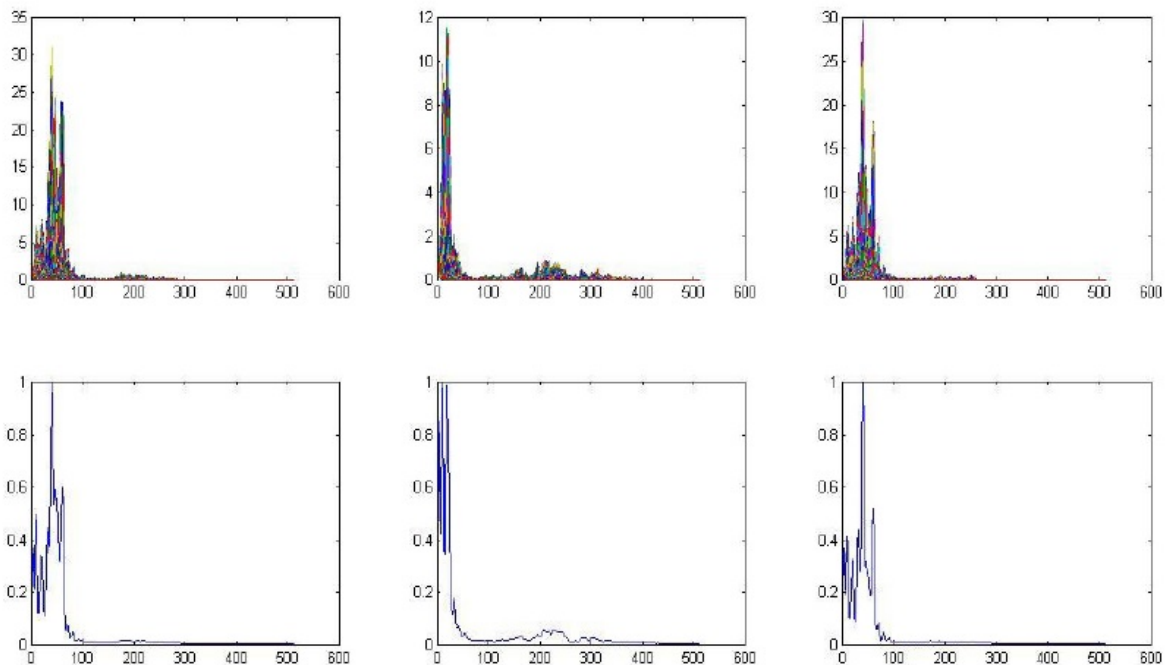
Σήμα -στόχος : Από τον χρόνο 1 έως 10 , "Door".

Από τον χρόνο 11 έως 20 , "Key".

Ομιλητής: "Δημήτρης" για τα δύο σήματα αναφοράς και το σήμα στόχο .

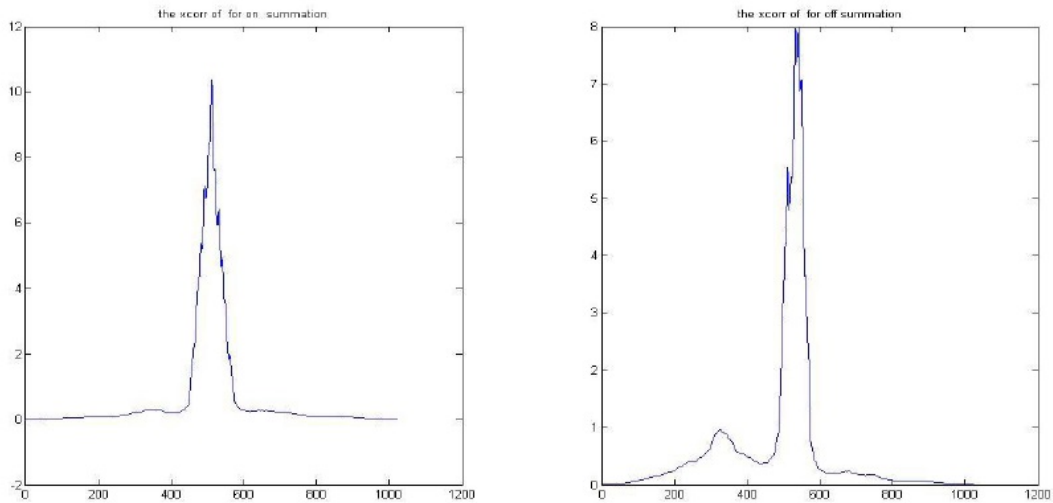
Περιβάλλον: σχεδόν κανένας θόρυβος

Το σχήμα 25 είναι για τα φάσματα συχνοτήτων για τα τρία καταγεγραμμένα σήματα που πήραμε με τον ίδιο τρόπο όπως στο σχήμα 21.



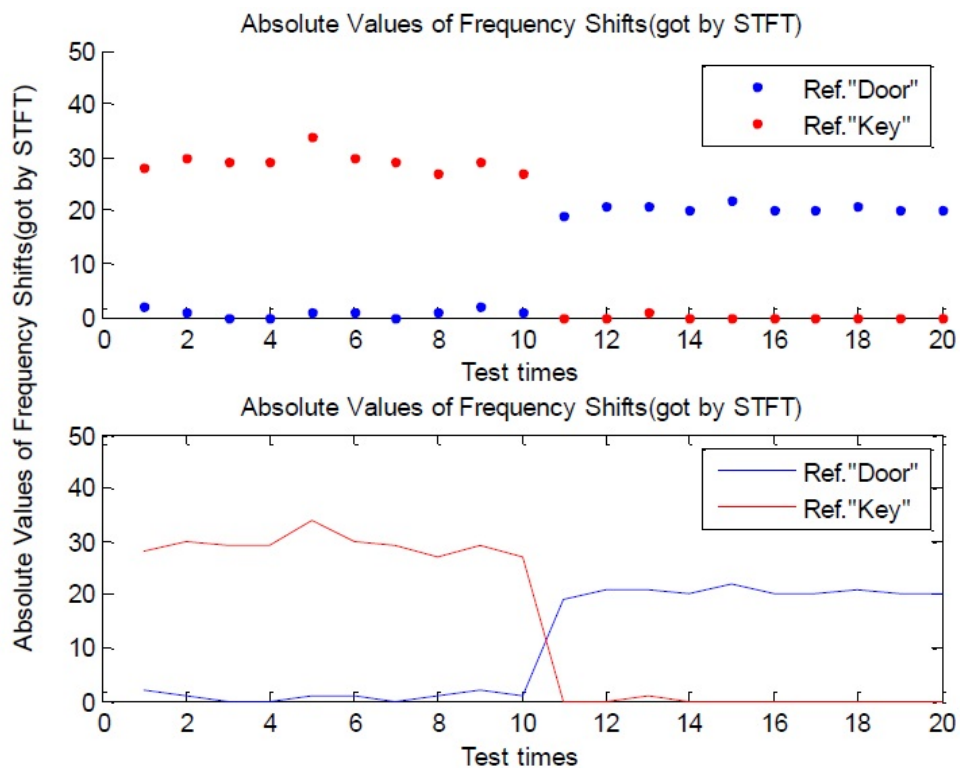
Σχήμα 25 : Τα φάσματα συχνοτήτων για τα τρία σήματα : "Door" , "Key" , και "Door"

Ο αριθμός των συσχετίσεων για το σήμα στόχο "Door" με τα σήματα αναφοράς είναι όπως παρακάτω (το σήμα αναφοράς της απεικόνισης αριστερά είναι "Door" και το σήμα αναφοράς της απεικόνισης δεξιά είναι "Key"):



Σχήμα 26: Οι συσχετίσεις για το σήμα στόχο "Door" με το σήμα αναφοράς "Key"

Στο Σχήμα 26 όπως φαίνεται παραπάνω, υπάρχει μεγάλη διαφορά μεταξύ των δύο γραφικών παραστάσεων. Αφού η προφορά των λέξεων "Door" και "Key" είναι εντελώς διαφορετική. Όπως αναφέραμε στη τμήμα 2.3.2, τα σήματα που μοιάζουν έχουν καλύτερη συμμετρική ιδιότητα της συσχέτισης. Αυτό φαίνεται από το Σχήμα 26. Σύμφωνα με τα αποτελέσματα της προσομοίωσης (Πίνακας 2, Παράρτημα Β), το απεικονιζόμενο αποτέλεσμα προσομοίωσης για την αλλαγή της συχνότητας είναι όπως παρακάτω:



Σχήμα 27: Η αλλαγή συχνότητας σε 20 προσομοιώσεις για τις λέξεις αναφοράς "Door" και "Key"

Από Σχήμα 27, φαίνεται ότι οι μετατοπίσεις συχνότητας έχουν μεγάλες διαφορές. Έτσι, το σχεδιασμένο σύστημα θα δώσει άμεσα τα αποτελέσματα σύμφωνα με τις μετατοπίσεις συχνότητας.

(3) Οι πληροφορίες των αποτελεσμάτων της τρίτης προσομοίωσης για το σύστημα 1 έχει ως εξής:

Σήματα αναφοράς : "on" και " off" :

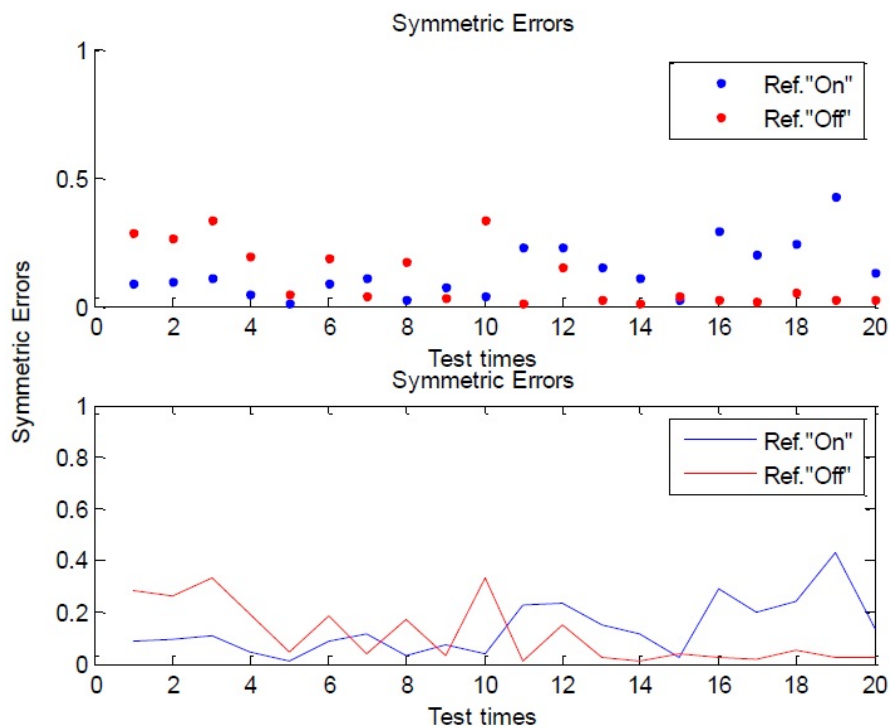
Στόχος του σήματος : Από τον χρόνο 1 έως 10 , "on" .

Από τον χρόνο 11 έως 20, "off" .

Ομιλητής : “Γιάννης” για τα δύο σήματα αναφοράς και το σήμα στόχο.

Περιβάλλον: υπάρχει κάποιος θόρυβος

Οι λέξεις "on" και "off" έχουν μικρές διαφορές μετατόπισης συχνότητας (Σχήμα 23 και Πίνακας 1), έτσι το σχεδιασμένο σύστημα θα δώσει μόνο τα αποτελέσματα με τα συμμετρικά σφάλματα. Το απεικονιζόμενο αποτέλεσμα προσομοίωσης (τα δεδομένα δίνονται στον Πίνακα 3 , Παράρτημα Β) είναι όπως παρακάτω:



Σχήμα 28 : Συμμετρικά σφάλματα σε 20 προσομοιώσεις για τις λέξεις αναφοράς "on" και "off" (με θόρυβο)

(4) Οι πληροφορίες για τα αποτελέσματα της τέταρτης προσομοίωσης για το σύστημα 1 έχει ως εξής :

Σήματα αναφοράς : "Door" και "Key" :

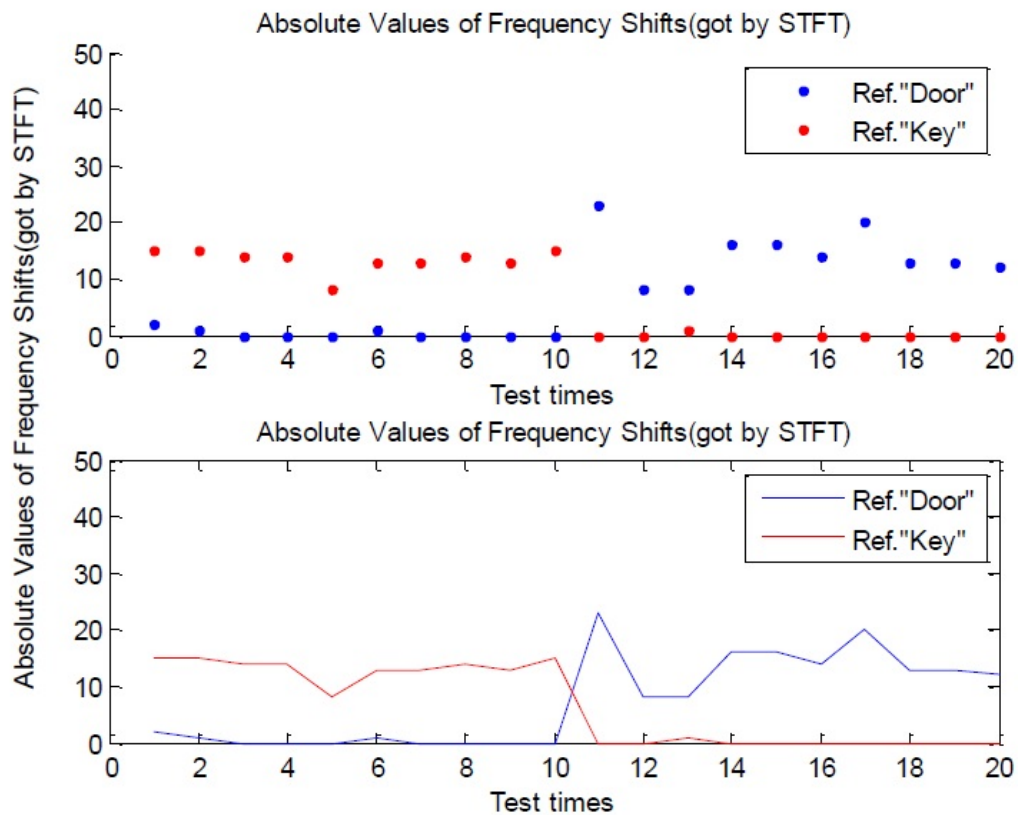
Σήμα στόχος : Από τον χρόνο 1 έως 10 , "Door" .

Από τον χρόνο 11 έως 20 , "Key" .

Ομιλητής : "Γιάννης" για τα δύο σήματα αναφοράς και το σήμα στόχο .

Περιβάλλον: εξακολουθεί να υπάρχει κάποιος θόρυβος

Το απεικονιζόμενο αποτέλεσμα προσομοίωσης (τα δεδομένα δίνονται στον Πίνακα 4, Παράρτημα Β) είναι όπως παρακάτω :



Σχήμα 29 : Η αλλαγή συχνότητας σε 20 προσομοιώσεις για τις λέξεις αναφοράς "Door" και "Key" (με θόρυβο)

(5) Οι πληροφορίες των αποτελεσμάτων της προσομοίωσης για το σύστημα 1 είναι ως εξής :

(6)

Σήματα αναφοράς : Από τον χρόνο 1 έως 10 : "on" και "off "

Από τον χρόνο 11 έως 20 : "Door" και "Key"

Σήμα στόχος: Από τον χρόνο 1 έως 5 , "on"

Από τον χρόνο 6 έως 10 , "off"

Από τον χρόνο 11 έως 15 , "Door"

Από τον χρόνο 16 έως 20 , "Key"

Ομιλητές : “Γιάννης” για τα σήματα αναφοράς

“Δημήτρης” για το σήμα στόχο

Περιβάλλον: σχεδόν κανένας θόρυβος

Σημείωση: Τα σήματα αναφοράς και το σήμα στόχος καταγράφονται από διαφορετικά πρόσωπα.

Όπως αναφέρεται στο τμήμα 3.1.1 , όταν το σύστημα 1 μπορεί να δώσει το αποτέλεσμα με τις μετατοπίσεις συχνότητας, τότε το σύστημα δεν θα υπολογίσει τα συμμετρικά σφάλματα. Όταν το σύστημα 1 , δεν μπορεί να δώσει τα αποτελέσματα με τις μετατοπίσεις συχνότητας, τότε το σύστημα θα υπολογίσει τα συμμετρικά σφάλματα. Σύμφωνα με αυτό, παίρνουμε το αποτέλεσμα της προσομοίωσης, όπως παρακάτω :

Ο Πίνακας 5 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "Door" και "Key" και τις πληροφορίες που δίνονται στην αρχή του παρόντος τμήματος (5).

Δοκιμές	Μετατόπιση συχνότητας on	Μετατόπιση συχνότητας off	Σφάλμα1	Σφάλμα2	Τελικά αποτελέσματα
1	2	8	No need	No need	on

2	7	8	0.2055	0.4324	on
3	8	9	0.2578	0.2573	off
4	9	17	No need	No need	on
5	8	9	0.2304	0.3640	on
6	0	0	0.3268	0.6311	on
7	0	0	0.3193	0.3210	on
8	0	0	2.2153	0.9354	off
9	0	0	0.4603	0.1481	off
10	0	0	0.1189	0.0741	off
11	8	22	No need	No need	Door
12	8	0	No need	No need	Key
13	8	25	No need	No need	Door
14	8	24	No need	No need	Door
15	8	24	No need	No need	Door
16	-15	0	No need	No need	Key
17	-15	0	No need	No need	Key
18	-14	0	No need	No need	Key
19	-14	0	No need	No need	Key
20	-15	0	No need	No need	Key
Σύνολο επιτυχημένων προσπαθειών (για 20 φορές)			80%		

Πίνακας 5 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις ομιλίας "On" , "Off" , "Door" και "Key"

Δεδομένου ότι τα αποτελέσματα της προσομοίωσης δεν είναι καλά, όπως οι αναμενόμενες τιμές. Για αυτό τα αποτελέσματα του πίνακα φαίνονται εδώ σε περίπτωση που τα απεικονιζόμενα αποτελέσματα είναι πολύ συγκεχυμένα.

3.2.2 Αποτελέσματα προσομοίωσης για το σύστημα 2

(1) Τα στοιχεία των πρώτων αποτελεσμάτων της προσομοίωσης για το σύστημα 2 έχει ως εξής :

Σήματα αναφοράς : "on" και "off" :

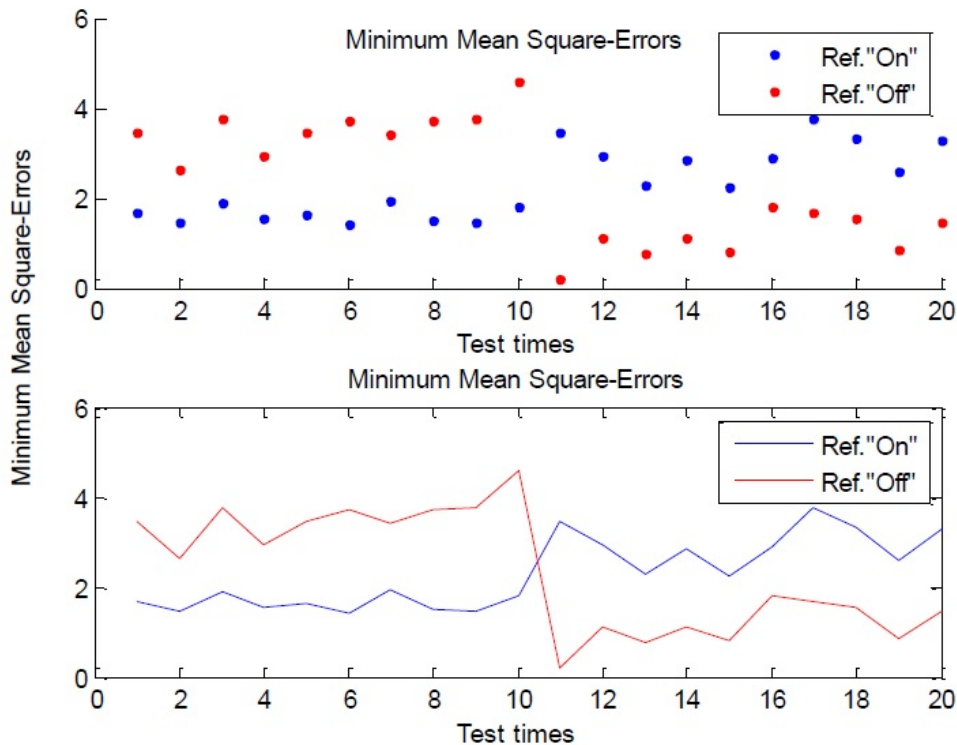
Σήμα στόχος : Από τον χρόνο 1 έως 10 , "on" .

Από τον χρόνο 11 έως 20 , "off" .

Ομιλητής : “Δημήτρης” για τα δύο σήματα αναφοράς και το σήμα στόχο .

Περιβάλλον : σχεδόν κανένας θόρυβος

Ο αριθμός των ελάχιστων μέσων τετραγωνικών σφαλμάτων μεταξύ των σημάτων του στόχου και των σημάτων αναφοράς είναι όπως παρακάτω (τα στοιχεία δίνονται στον Πίνακα 6, Παράρτημα Β) :



Σχήμα 30 : Ελάχιστο μέσο τετραγωνικών σφαλμάτων για τα σήματα του στόχου με τα σήματα αναφοράς

(2) Τα στοιχεία των αποτελεσμάτων της δεύτερης προσομοίωσης για το σύστημα 2 έχει ως εξής :

Σήματα αναφοράς : "Door" και "Key" :

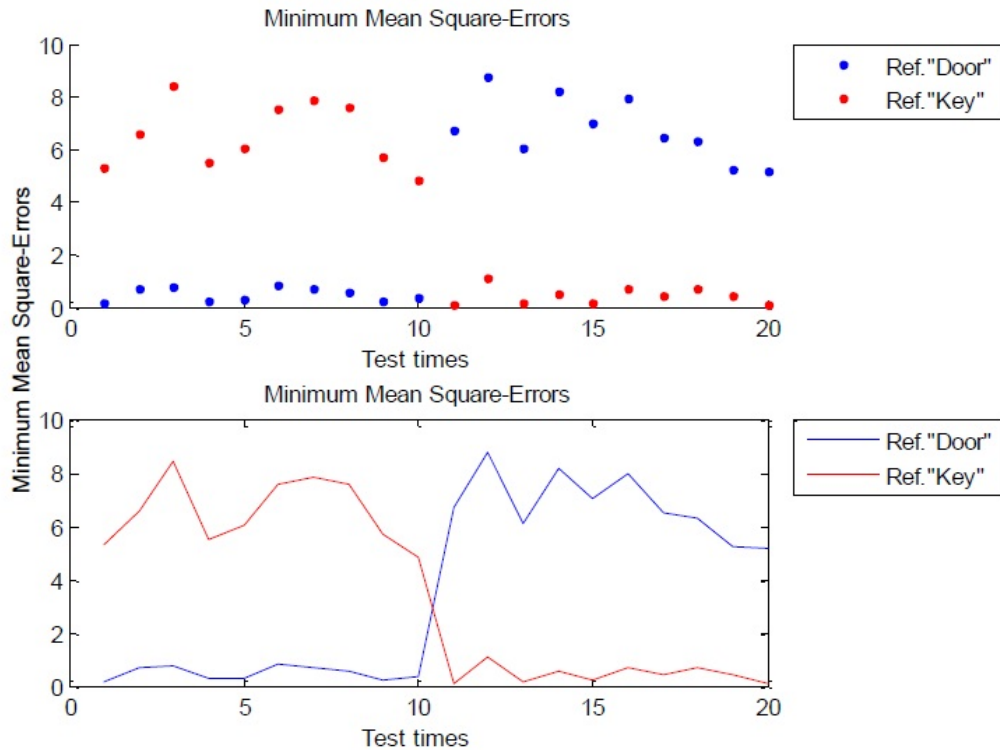
Σήμα-στόχος : Από τον χρόνο 1 έως 10 , " Door"

Από τον χρόνο 11 έως 20 , "Key"

Ομιλητής : "Δημήτρης" για τα δύο σήματα αναφοράς και το σήμα στόχου

Περιβάλλον : σχεδόν κανένας θόρυβος

Ο αριθμός των ελάχιστων μέσων τετραγωνικών σφαλμάτων για τα σήματα στόχου με τα σήματα αναφοράς είναι όπως παρακάτω (τα στοιχεία δίνονται στον Πίνακα 7, Παράρτημα Β):



Σχήμα 31 : Ελάχιστο μέσο τετραγωνικών σφαλμάτων για τα σήματα στόχου με τα σήματα αναφοράς

(3) Τα στοιχεία των αποτελεσμάτων της τρίτης προσομοίωσης για το σύστημα 2 έχει ως εξής:

Σήματα αναφοράς : "on" και "off"

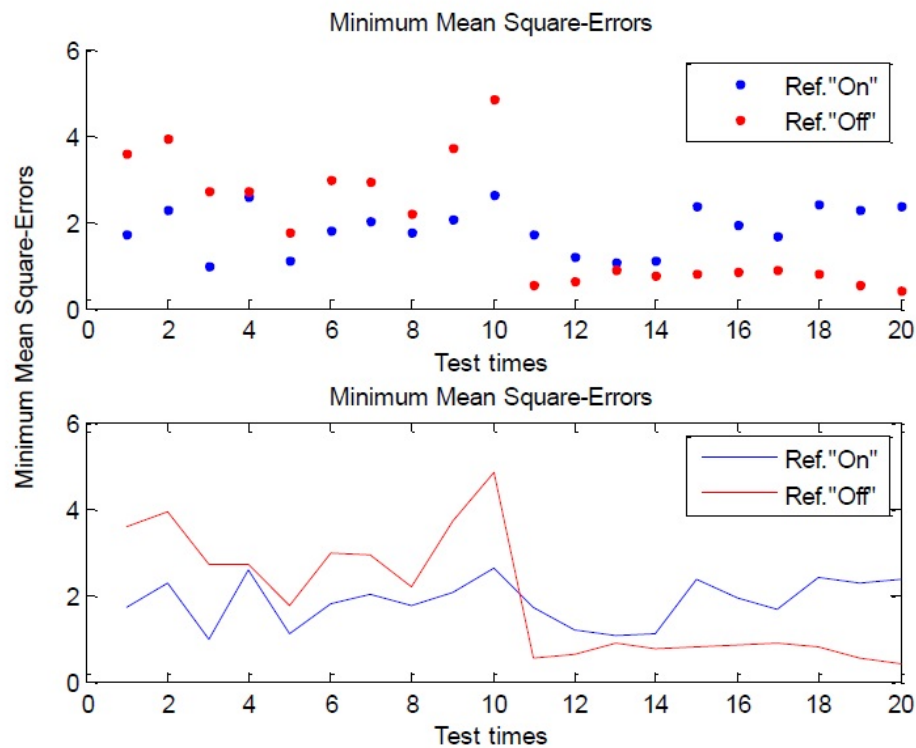
Σήμα-στόχος : Από τον χρόνο 1 έως 10 , "on"

Από τον χρόνο 11 έως 20 , "off "

Ομιλητής : "Γιώργος" για τα δύο σήματα αναφοράς και το σήμα στόχου.

Περιβάλλον : σχεδόν κανένας θόρυβος

Ο αριθμός των ελάχιστων μέσω τετραγωνικών σφαλμάτων για τα σήματα στόχου με τα σήματα αναφοράς είναι όπως παρακάτω (τα στοιχεία δίνονται στον Πίνακα 8 , Παράρτημα Β):



Σχήμα 32 : Ελάχιστο μέσο τετραγωνικών σφαλμάτων για τα σήματα στόχου με τα σήματα αναφοράς

(4) Τα στοιχεία των αποτελεσμάτων της τέταρτης προσομοίωσης για το σύστημα 2 έχει ως εξής :

Σήματα αναφοράς : "Door" και "Key" :

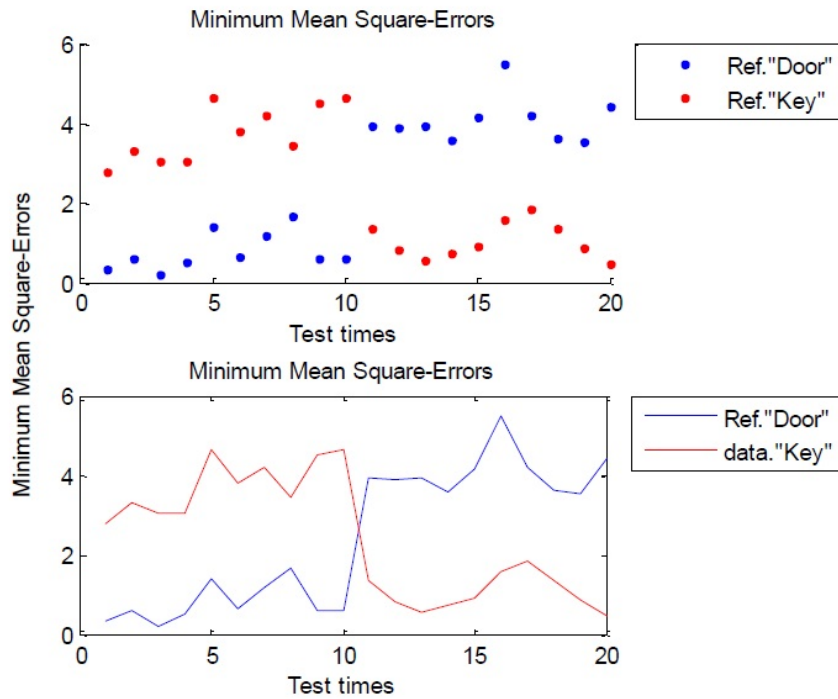
Σήμα-στόχος : Από τον χρόνο 1 έως 10 , "Door" .

Από τον χρόνο 11 έως 20 , "Key" .

Ομιλητής : "Γιώργος" , για τα σήματα αναφοράς και το σήμα στόχο .

Περιβάλλον : σχεδόν κανένας θόρυβος

Ο αριθμός των ελάχιστων μέσων τετραγωνικών σφαλμάτων για τα σήματα στόχου με τα σήματα αναφοράς είναι όπως παρακάτω (τα στοιχεία δίνονται στον Πίνακα 9 , Παράρτημα Β):



Σχήμα 33 : Ελάχιστο μέσο τετραγωνικών σφαλμάτων για τα σήματα στόχου με τα σήματα αναφοράς

(5) Τα στοιχεία των αποτελεσμάτων της πέμπτης προσομοίωσης για το σύστημα 2 έχει ως εξής :

Σήματα αναφοράς : "on" και "off"

Σήμα-στόχος : Από τον χρόνο 1 έως 10 , "on"

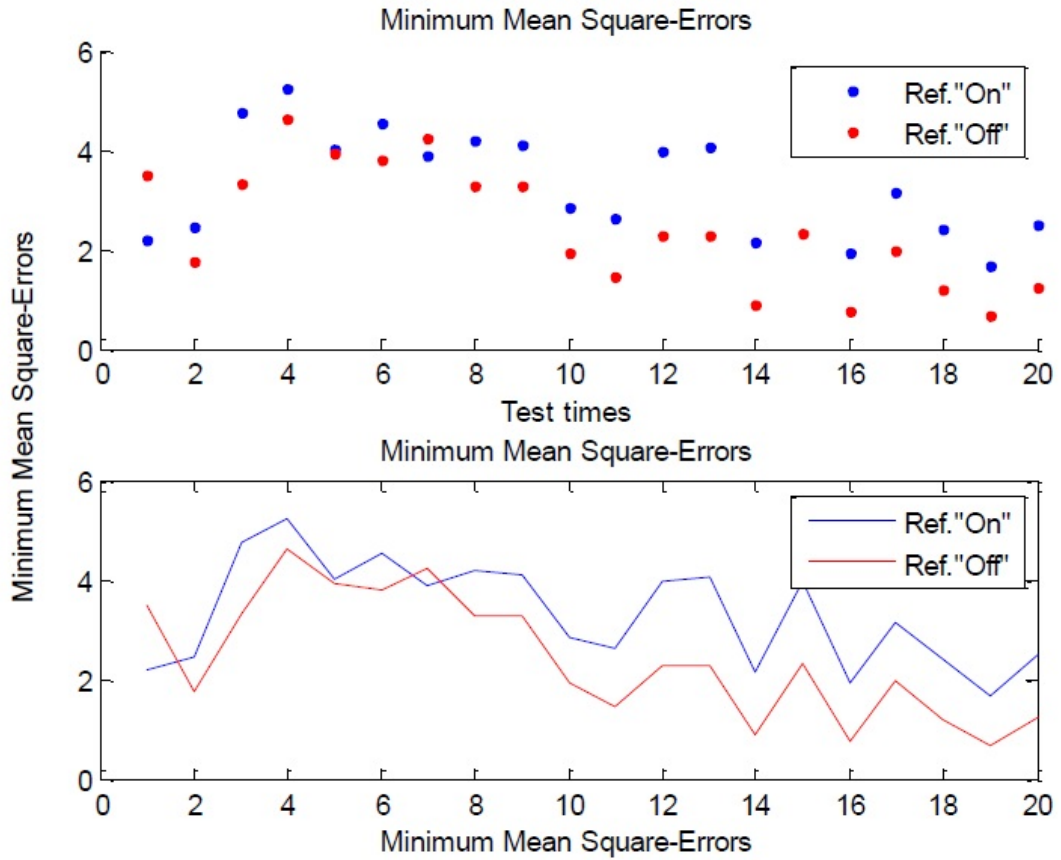
Από τον χρόνο 11 έως 20 , "off "

Ομιλητές : “Γιώργος” για το σήμα στόχο

“Δημήτρης” για τα σήματα αναφοράς

Σημείωση: Τα σήματα αναφοράς και το σήμα στόχος καταγράφονται από διαφορετικά πρόσωπα.

Ο αριθμός των ελάχιστων μέσων τετραγωνικών σφαλμάτων για τα σήματα στόχου με τα σήματα αναφοράς είναι όπως παρακάτω (τα στοιχεία δίνονται στον Πίνακα 10 , Παράρτημα Β) :



Σχήμα 34 : Ελάχιστο μέσο τετραγωνικά σφάλματα για τα σήματα στόχου με σήματα αναφοράς

4° Κεφάλαιο

Συζήτηση και Συμπεράσματα

Σε αυτή την ενότητα, θα συζητήσουμε τι έχουμε κάνει στο τμήμα αποτελέσματα και θα αναλύσουμε τα αποτελέσματα της προσομοίωσης που εμφανίζονται στο προηγούμενο κεφάλαιο. Στο τέλος, σύμφωνα με τις συζητήσεις των αποτελεσμάτων της προσομοίωσης, θα δώσουμε τα σχετικά συμπεράσματα.

4.1 Συζήτηση

4.1.1 Συζήτηση για τα αποτελέσματα της προσομοίωσης του συστήματος 1

Υπάρχουν 5 πίνακες για τα αποτελέσματα σχημάτων στο προηγούμενο κεφάλαιο για το σχεδιασμένο σύστημα 1. Πίνακας 1 και Πίνακας 2 προσομοιώνονται για να δείτε πώς λειτουργεί το σχεδιασμένο σύστημα για τα διάφορα ζεύγη λέξεων, οι εύκολα αναγνωρισμένες λέξεις και οι δύσκολα αναγνωρισμένες λέξεις. Από τον Πίνακα 1 (Σχήμα 23), μπορεί να παρατηρηθεί ότι η μετατόπιση συχνότητας δεν είναι αξιόπιστη, δεδομένου ότι η προφορά των "on" και "off" είναι παρόμοια. Σε αυτή την κατάσταση, το σχεδιασμένο σύστημα θα δώσει τις αποφάσεις συγκρίνοντας τα λάθη της συμμετρικής ιδιότητας των συσχετίσεων (Σχήμα 24). Η αρχή αυτής της μεθόδου αναγνώρισης αναφέρεται στο τμήμα 2.3.2 της παρούσας πτυχιακής. Αν η προφορά των δύο λέξεων είναι τελείως διαφορετική, όπως "Door" και "Key", το σχεδιασμένο σύστημα 1 θα κάνει τις αποφάσεις άμεσα από τις μεταβολές της συχνότητας, τις οποίες μπορούμε να παρατηρήσουμε από τον Πίνακα 2 (Σχήμα 27). Το σύστημα που σχεδιάστηκε δεν θα υπολογίσει τα σφάλματα πια. Παρατηρώντας μεγάλες ποσότητες των αποτελεσμάτων της προσομοίωσης, έχουμε προγραμματίσει το σύστημα στο MATLAB για να βασιστούμε στις μετατοπίσεις συχνότητας, όταν η διαφορά μεταξύ των απόλυτων τιμών των μετατοπίσεων συχνότητας για τα διαφορετικά σήματα αναφοράς είναι μεγαλύτερη ή ίση με 2. Διαφορετικά, το σχεδιασμένο σύστημα 1, θα υπολογίζει συνεχώς τα σφάλματα της συμμετρίας. Στον Πίνακα 3 (Σχήμα 28)

και στον Πίνακα 4 (Σχήμα 29) έγιναν οι προσομοιώσεις με τους φίλους μας, που προέρχονται από διαφορετικές χώρες . Ο σκοπός αυτών των αποτελεσμάτων είναι να δείξει ότι το σχεδιασμένο σύστημα πραγματικά λειτουργεί για τους διαφορετικούς ανθρώπους που προέρχονται από διαφορετικές χώρες. Και από τον Πίνακα 3 (Σχήμα 28) και τον Πίνακα 4 (Σχήμα 29) , φαίνεται ότι το σχεδιασμένο σύστημα δεν λειτουργεί τόσο καλά όταν υπάρχει θόρυβος γύρω. Αυτό δείχνει ότι το σχεδιασμένο σύστημα είναι εύκολο να διαταραχθεί από το θόρυβο. Τα αποτελέσματα του Πίνακα 5 προσομοιώνονται από δύο άτομα. Ο ένας είναι υπεύθυνος για τις εγγραφές των σημάτων αναφοράς και ο άλλος είναι υπεύθυνος για τις εγγραφές του σήματος στόχου. Μπορεί να παρατηρηθεί ότι το σύστημα που σχεδιάστηκε σχεδόν δεν λειτουργεί για αυτή την κατάσταση.

4.1.2 Συζήτηση για τα αποτελέσματα της προσομοίωσης του συστήματος 2

Υπάρχουν επίσης 5 πίνακες των αποτελεσμάτων της προσομοίωσης για το σχεδιασμένο σύστημα 2. Οι σκοποί αυτών των προσομοιώσεων είναι οι ίδιοι όπως τους σκοπούς του σχεδιασμένου συστήματος 1. Το σύστημα 1 είναι σχεδιασμένο παρατηρώντας μεγάλες ποσότητες των σημάτων των συσχετίσεων. Το σύστημα 2 έχει σχεδιαστεί χρησιμοποιώντας τα σήματα αναφοράς για τη μοντελοποίηση του σήματος στόχου. Με τη σύγκριση των σφαλμάτων μεταξύ του πραγματικού σήματος στόχος και του σήματος στόχος μοντελοποίησης που πήραμε από τα διαφορετικά σήματα αναφοράς , το σύστημα 2 δίνει την κρίση ποιο σήμα αναφοράς είναι παρόμοιο με το σήμα στόχο. Σχεδιάσαμε το φίλτρο για την υλοποίηση της διαδικασίας μοντελοποίησης σήματος. Σε αυτό το σύστημα, τα σήματα αναφοράς χρησιμοποιούνται ως πηγές αυτόματης συσχέτισης, τα οποία είναι οι εισοδοί του φίλτρου. Και το σήμα στόχου χρησιμοποιείται ως το επιθυμητό σήμα. Με την εξίσωση του φίλτρου Wiener $R_{xw} = r_{dx}$, είναι γνωστό ότι κατά την εφαρμογή αυτής της εξίσωσης, δίνει πραγματικά την υπόθεση ότι το σήμα εισόδου $x(n)$ συσχετίζεται με την επιθυμητό σήμα $d(n)$. Με άλλα λόγια, τα σήματα αναφοράς θα πρέπει να συσχετίζονται με το σήμα του στόχου. Αλλά αν ένα άτομο δίνει τα σήματα αναφοράς και το άλλο άτομο δίνει το σήμα στόχου, τότε τα σήματα αναφοράς και το σήμα στόχος δεν συσχετίζονται το ένα με το άλλο. Έτσι, το σχεδιασμένο σύστημα 2 δεν λειτουργεί καλά όταν τα σήματα αναφοράς και το σήμα στόχος καταγράφεται από διαφορετικούς ανθρώπους.

4.2 Συμπεράσματα

Τα γενικά συμπεράσματα είναι ότι τα σχεδιασμένα συστήματα για την αναγνώριση φωνής διαταράσσονται εύκολα από τον θόρυβο, ο οποίος μπορεί να παρατηρηθεί από τον Πίνακα 3 (Σχήμα 28) . Για το σχεδιασμένο σύστημα 1 , όσο καλύτερα είναι τα συμφωνημένα σήματα τόσο έχουν την καλύτερη συμμετρική ιδιότητα συσχέτισης τους. Αυτό το συμπέρασμα φαίνεται από τον Πίνακα 1 (Σχήμα 23 και Σχήμα 24) και στον Πίνακα 2 (Σχήμα 27). Για το σχεδιασμένο σύστημα 2 , αν το σήμα αναφοράς είναι η ίδια λέξη με το σήμα στόχου, η χρήση αυτού του σήματος αναφοράς προς το μοντέλο του στόχου θα έχει λιγότερα σφάλματα. Αυτό το συμπέρασμα μπορεί να αποδειχθεί από όλα τα αποτελέσματα της προσομοίωσης για το σχεδιασμένο σύστημα 2. Όταν και τα δύο σήματα αναφοράς και το σήμα στόχος που καταγράφονται από το ίδιο πρόσωπο, τα δύο συστήματα λειτουργούν καλά για να διακρίνουν τις διάφορες λέξεις , δεν έχει σημασία από ποια χώρα είναι το πρόσωπο αυτό. Αλλά εάν τα σήματα αναφοράς και το σήμα στόχος καταγράφονται από διαφορετικούς ανθρώπους, και τα δύο συστήματα δεν λειτουργούν τόσο καλά . Έτσι, προκειμένου να βελτιωθούν τα συστήματα που έχουν σχεδιαστεί ώστε να λειτουργήσουν καλύτερα, πρέπει οι εργασίες να εξετάσουμε τον θόρυβο και να βρούμε τα κοινά χαρακτηριστικά της φωνής για τους διαφορετικούς ανθρώπους. Διαφορετικά, αν σχεδιάσουμε μερικά αναλογικά και ψηφιακά φίλτρα για την επεξεργασία των σημάτων εισόδου μπορεί να μειώσει τις επιπτώσεις του θορύβου εισόδου και να καθοριστεί η μεγάλη βάση δεδομένων των σημάτων φωνής για διαφορετικές λέξεις. Και μελετώντας πιο προηγμένους αλγόριθμους για το σήμα μοντελοποίησης [9] μπορεί να μας βοηθήσει να πραγματοποιήσουμε την καλύτερη αναγνώριση φωνής.

Αναφορές

- [1] Joseph Picone, “Signal Modeling Techniques In Speech”, Systems and Information Sciences Laboratory, Tsukuba Research and Development Center, Tsukuba,Japan.
- [2] John G.Proakis, Dimitris G.Manolakis, Digital Signal Processing, Principles, Algorithms, and Applications, 4th edition,Pearson Education inc., Upper Saddle River.
- [3] Luis Buera, Antonio Miguel, Eduardo Lleida, Oscar Saz, Alfonso Oretega, “Robust Speech Recognition with On-line Unsupervised Acoustic Feature Compensation”, Communication Technologies Group (GTC),13A, University of Zaragoza, Spain.
- [4] Hartmut Traunmüller , Anders Eriksson, “The frequency range of the voice fundamental in the speech of male and female adults”, Institutionen för lingvistik, Stockholms universitet, S-106 91 Stockholm, Sweden.
- [5] Jian Chen, Jiwan Gupta,“Estimation of shift parameter of headway distributions using crosscorrelation function method”, Department of Civil Engineering, The University of Toledo.
- [6] John Wiley, Sons,Inc. Statistical Signal Processing And Modeling, Monson H Hayes,Georgia I nstitute of Technology.
- [7] Henrik V. Sorensen, C. Sidney Burrus, “Efficient Computation of the Short-Time Fast Fourier Transform”, Electrical and Computer Engineering Department, Rice University, Houston.
- [8] Fredric J. Harris,Member, IEEE, “On the Use of Windows for Harmonic Analysis with the Discrete Fourier transform”, Proceedings of the IEEE, VOL 66, No.1, JANUARY, 1978

[9] Joseph W. Picone, senior member, IEEE “Signal Modeling Techniques in Speech”
Recognition

Παράρτημα Α

Ο κώδικας των συστημάτων στο MATLAB είναι ως εξής:

Κώδικας Συστήματος 1:

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% αρχικές τιμές  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
clear;  
on=0;  
off=0;  
fs=16000; %συχνότητα δειγματοληψίας, 16000 δείγματα ανά δευτερόλεπτο  
test=0;  
duration=2; %χρόνος εγγραφής  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
fprintf('Press any key to start %g seconds of recording.on.\n',duration);  
  
pause;  
  
fprintf('Recording...\n');  
  
r=wavrecord(2*fs,fs);  
%duration*fs: το μήκος των δεδομένων εγγραφής: 2*fs δείγματα σε 2 δευτερόλεπτα  
  
r=r-mean(r);  
  
fprintf('Press any key to start %g seconds of recording. off.\n',duration);  
  
pause;  
  
fprintf('Recording...\n');  
  
y=wavrecord(2*fs,fs);  
  
fprintf('Press any key to start %g seconds of recording. voice.\n',duration);
```



```
y=y-mean(y);  
pause;
```

```
fprintf('Recording...\n');
```

```
voice=wavrecord(2*fs,fs);  
voice=voice-mean(voice);  
fprintf('Finished recording.\n');
```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% χρήση specgram για να πάρουμε τις πληροφορίες των σημάτων
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% στο πεδίο ορισμού συχνοτήτων

```
nfft = min(1023,length(r)); %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%καθορισμός μήκους STFT  
% καθορισμός μήκους παραθύρου και απόφαση βήματος για το παράθυρο
```

```
s=specgram(r, nfft, fs, hanning(511),380);  
s2=specgram(y, nfft, fs, hanning(511),380);  
s3=specgram(voice, nfft, fs, hanning(511),380);
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ΜΕΡΟΣ 1....πληροφορίες συχνότητας  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ΦΑΣΜΑ ΣΥΧΝΟΤΗΤΑΣ%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% s είναι σύνθετος πίνακας  
absolute=transpose(abs(s)); %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%δημιουργία σχήματος  
absolute2=transpose(abs(s2));  
absolute3=transpose(abs(s3));
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ΦΑΣΜΑ ΣΥΧΝΟΤΗΤΑΣ%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% για το φάσμα  
συχνότητας,  
% Αν ο A είναι ένας πίνακας , το sum(A) αντιμετωπίζει τις στήλες του A ως διανύσματα,  
% επιστρέφοντας ένα διάνυσμα γραμμής των αθροισμάτων της κάθε στήλης  
% μετά τη μετατόπιση, οι σειρές και οι στήλες έχουν μετατραπεί,  
% παίρνει το άθροισμα του μεταφερόμενου πίνακα , έχουμε το άθροισμα του άξονα του χρόνου  
% για να επιστρέψει το φάσμα συχνοτήτων  
a4=sum(absolute) %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% παίρνουμε το χρονικό φάσμα συχνοτήτων  
a5=sum(absolute2)  
a6=sum(absolute3)  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ΦΑΣΜΑ ΣΥΧΝΟΤΗΤΑΣ  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% κανονικοποίηση του φάσματος και μείωση επίδρασης θορύβου
```

```
a4_norm=(a4-min(a4))/(max(a4)-min(a4));  
a5_norm=(a5-min(a5))/(max(a5)-min(a5));  
a6_norm=(a6-min(a6))/(max(a6)-min(a6));  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%μεταφορά σειράς για το διάνυσμα  
στήλης%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
F4=transpose(a4_norm); % χρήση για τα δεδομένα διανύσματος
F5=transpose(a5_norm);
F6=transpose(a6_norm);
%%%%%%%%%% ΦΑΣΜΑ ΣΥΧΝΟΤΗΤΑΣ: σύγκριση
συσχέτισης%%%%%%%%%
```

```
[x3,lag3]=xcorr(F6,F4);
[mx3,indice3]=max(x3);
frequency_on_shift=lag3(indice3)
[x4,lag4]=xcorr(F6,F5);
```

```
[mx4,indice4]=max(x4);
frequency_off_shift=lag4(indice4)
```

```
%%%%%%%%%% σχήμα φάσματος στο πεδίο
συχνότητας %%%%%%%%%%
```

```
figure(1) % 2x3 πίνακας , πρώτη γραμμή το αρχικό φάσμα
           % δεύτερη γραμμή άθροισμα πραγματικού φάσματος
```

```
subplot(2,3,1)
plot(abs(s))
subplot (2,3,2)
plot(abs(s2))
subplot (2,3,3)
plot(abs(s3))
subplot(2,3,4)
```

```
plot(F4);
subplot(2,3,5)
plot(F5);
subplot(2,3,6)
plot(F6);
```

```
figure(2)
subplot(1,2,1)
plot(x3)
title('the xcorr of for on summation');
subplot(1,2,2 )
plot(x4)
title('the xcorr of for off summation');
```

```
%%%%%%%%%% δοκιμή 6 %%%%%%%%%%%
%%%%%%%%%% σύγκριση φάσματος συχνότητας
% πρώτον έλεγχος για το θέμα της αλλαγής συχνότητας
% Έχω εμπιστοσύνη στην κρίση μόνο όταν η διαφορά μετατόπισης συχνότητας είναι
```

```

% μεγαλύτερη ή ίση του 2.
% όσο μικρότερη είναι η μετατόπιση συχνότητα, τόσο καλύτερα τα ταιριαστά σήματα
if abs(abs(frequency_on_shift)-abs(frequency_off_shift))>=2
    if abs(frequency_on_shift)>abs(frequency_off_shift)
        off=off+1;
        test=0;
    else
        on=on+1;
        test=1;
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%αποτέλεσμα στην οθόνη
frequency_on_shift

```

```

frequency_off_shift
test
if on>off
display('answer is on')
else
display('answer is off')
end

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Όταν δεν μπορεί να δώσει την απόφαση με τη μετατόπιση συχνότητας, μπορούμε να
% κάνουμε άμεσα την απόφαση από την συμμετρική ιδιότητα συσχέτισης για το
% ανάλογο σχήμα του σήματος.

```

```

else

```

```

if indice3<length(x3)/2
q=1:(indice3-1);
p=indice3+length(q):-1:indice3+1
length(p);
length(q);
x3_left=x3(q);
min(x3_left);
x3_right=x3(p);
min(x3_right);
error1=mean((abs(x3_right-x3_left)).^2);
else
q=1+frequency_on_shift*2:indice3-1
p=length(x3):-1:indice3+1
length(q);
length(p);
x3_left=x3(q);
x3_right=x3(p);
error1=mean((abs(x3_right-x3_left)).^2);
end

```

```

if indice4 < length(x4)/2
q2=1:indice4-1
p2=indice4+length(q2):-1:indice4+1
x4_left=x4(q2)
length(q2);
length(p2);
min(x4_left);

x4_right=x4(p2)
min(x4_right);
error2=mean((abs(x4_right-x4_left)).^2)
else
q2=1+frequency_off_shift*2:indice4-1
p2=length(x4):-1:indice4+1
length(q2);
length(p2);
x4_left=x4(q2);
x4_right=x4(p2);
error2=mean((abs(x4_right-x4_left)).^2)
end
if error1 > error2
off=off+1;
test=0;
else
on=on+1;
test=1;
end

```

```

%%%%%%%%%%

```

% όταν δεν μπορεί να δώσει την κρίση με τη μετατόπιση συχνότητας, μπορούμε να κάνουμε
%άμεσα την κρίση από τη συμμετρική ιδιότητα συσχέτισης για το αντίστοιχο σχήμα του
%σήματος.

```

else
if indice3 < length(x3)/2
q=1:(indice3-1);
p=indice3+length(q):-1:indice3+1
length(p);
length(q);
x3_left=x3(q);
min(x3_left);
x3_right=x3(p);
min(x3_right);
error1=mean((abs(x3_right-x3_left)).^2);
else
q=1+frequency_on_shift*2:indice3-1
p=length(x3):-1:indice3+1

```

```

length(q);
length(p);
x3_left=x3(q);
x3_right=x3(p);
error1=mean((abs(x3_right-x3_left)).^2);
end
if indice4<length(x4)/2
q2=1:indice4-1
p2=indice4+length(q2):-1:indice4+1
x4_left=x4(q2)
length(q2);
length(p2);
min(x4_left);
x4_right=x4(p2)
min(x4_right);
error2=mean((abs(x4_right-x4_left)).^2)
else
q2=1+frequency_off_shift*2:indice4-1
p2=length(x4):-1:indice4+1
length(q2);

length(p2);
x4_left=x4(q2);
x4_right=x4(p2);
error2=mean((abs(x4_right-x4_left)).^2)
end
if error1>error2
off=off+1;
test=0;
else
on=on+1;
test=1;
end

%%%%%%%%%%%% αποτελέσματα %%%%%%%%%%%%%
on
off
error1
error2
frequency_on_shift
frequency_off_shift
test
if on>off
display('answer is on')
if on==off
display('donno what u have said')
end
else

```

```
display('answer is off')
end
end
```

Κώδικας Συστήματος 2:

```
%%%%%%%%%%wiener filter αναγνώριση φωνής φίλτρο του φάσματος συχνοτήτων
%%%%%%%%%%wiener filter μέθοδος για να πραγματοποιήσουμε την αναγνώριση φωνής
%%% γραμμική πρόβλεψη
%%% σήμα φωνής d(n)
% χρήση σήματος αναφοράς 1 και σήματος 2 για την εκτίμηση d(n)_hat
%%%%%%%% εύρεση σφάλματος e=mean(abs(d(n)-d(hat))^2)
%%%%%%%%%% αρχικές τιμές

clear;
on=0;
off=0;
fs=16000; %δειγματοληψία συχνότητας, 16000 δείγματα ανά δευτερόλεπτο
duration=2; %χρόνος εγγραφής
%%%%%%%%%%
fprintf('Press any key to start %g seconds of recording.on.\n',duration);
pause;
fprintf('Recording...\n');
r=wavrecord(2*fs,fs);
%διάρκεια*fs συνολικά δείγματα : 2*fs δείγματα χρειάζονται 2 δευτερόλεπτα

r=r-mean(r);
fprintf('Press any key to start %g seconds of recording. off.\n',duration);
pause;
fprintf('Recording...\n');

y=wavrecord(2*fs,fs); %διάρκεια*fs : 2*fs δείγματα χρειάζονται 2 δευτερόλεπτα

%%%%%%%%%% έλεγχος πόσες φορές θέλουμε να γίνει η δοκιμή

for T=1:20
fprintf('Press any key to start %g seconds of recording.voice.\n',duration);
```

```

y=y-mean(y);
pause;
fprintf('Recording...\n');
voice=wavrecord(2*fs,fs);
voice=voice-mean(voice);
fprintf('Finished recording.\n');

```

χρήση *specgram* για να πάρουμε τις πληροφορίες των σημάτων στο πεδίο συχνοτήτων και στο πεδίο του χρόνου

```

nfft = min(1023,length(r)); %ορισμός μήκους του STFT
s=specgram(r, nfft, fs, hanning(511),380);
s2=specgram(y, nfft, fs, hanning(511),380);
s3=specgram(voice, nfft, fs, hanning(511),380);

```

ΜΕΡΟΣ 1...πληροφορίες συχνότητας
ΦΑΣΜΑ ΣΥΧΝΟΤΗΤΑΣ
s είναι σύνθετος πίνακας
absolute=transpose(abs(s)); %παίρνουμε το abs για να δημιουργήσουμε το σχήμα
absolute2=transpose(abs(s2));
absolute3=transpose(abs(s3));

ΦΑΣΜΑ ΣΥΧΝΟΤΗΤΑΣ
για το φάσμα συχνότητας,
a4=sum(absolute) %παίρνουμε το χρόνο συχνότητας σχετικά με το φάσμα
a5=sum(absolute2)
a6=sum(absolute3)

ΦΑΣΜΑ ΣΥΧΝΟΤΗΤΑΣ
κανονικοποίηση του φάσματος και μείωση επίδρασης θορύβου
a4_norm=(a4-min(a4))/(max(a4)-min(a4));
a5_norm=(a5-min(a5))/(max(a5)-min(a5));
a6_norm=(a6-min(a6))/(max(a6)-min(a6));
μετατόπιση γραμμής σε στήλη
F4=transpose(a4_norm);
F5=transpose(a5_norm);
F6=transpose(a6_norm);
F4 είναι x1
F5 είναι x2
F6 είναι d(n)
N=length(F6);
rx1=xcorr(F4,F4);
rx2=xcorr(F5,F5);
d=F6;
rdx1=xcorr(F6,F4);
rdx2=xcorr(F6,F5);
rd=xcorr(d,d);
χρήση x1 για τον υπολογισμό d(n) =====

```

for p=1:20;
rx_1=rx1(N:N+p-1);
rdx_1=rdx1(N:N+p-1);
Rx1=toeplitz(rx_1, rx_1)
det(Rx1);%%%%%%%%%%%%%% έλεγχος Rx ότι δεν είναι μοναδικό
I=inv(Rx1);
w=I*rdx_1;
e1(p)=rd(N)-transpose(w)*rdx_1
end
%%%%%%%%%%%%%% χρήση x2 για τον υπολογισμό d(n) =====
for p=1:20;

rx_2=rx2(N:N+p-1);
rdx_2=rdx2(N:N+p-1);
Rx2=toeplitz(rx_2, rx_2)
det(Rx2);%%%%%%%%%%%%%% έλεγχος Rx ότι δεν είναι μοναδικό
I=inv(Rx2);
w=I*rdx_2;

e2(p)=rd(N)-transpose(w)*rdx_2
end
figure(1)
subplot(211)
title('reference on')
plot(e1);grid;
subplot(212)
title('reference off')
plot(e2);grid;
%%%%%%%%%%%%%% σχήμα φάσματος στο πεδίο συχνότητας
figure(2)
subplot(2,3,1)
plot(abs(s))

subplot (2,3,2)
plot(abs(s2))
subplot (2,3,3)
plot(abs(s3))
subplot(2,3,4)
plot(F4);
subplot(2,3,5)
plot(F5);
subplot(2,3,6)
plot(F6);

m1=mean(e1)
m2=mean(e2)

if m1<m2

```



```

display('answer is on')
if m1==m2
    display ('donno what to say')
end
else
    display ('what u said is off')
end
end

```

Παράρτημα Β

Οι ορισμοί του "Error1" και "Error2" στον παρακάτω πίνακα μπορείτε να τα βρείτε στο τμήμα 2.3.2 της παρούσας διατριβής. Τα "frequency_on_shift" και "frequency_off_shift" στον ακόλουθο πίνακα είναι οι μετατοπίσεις συχνότητας, τα οποία πήραμε συγκρίνοντας τα φάσματα, όπως φαίνεται στο Σχήμα 21. Το "No need" σημαίνει ότι το σχεδιασμένο σύστημα μπορεί να δώσει την κρίση μόνο με τη μετατόπιση συχνότητας, χωρίς τον υπολογισμό των καθορισμένων σφαλμάτων.

Ο Πίνακας 1 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "on" και "off" σύμφωνα με τις πληροφορίες που δίνονται στο μέρος Τα Αποτελέσματα της Προσομοίωσης 3.2.1(1) .

Δοκιμές	Μετατόπιση συχνότητας on	Μετατόπιση συχνότητας off	Σφάλμα1	Σφάλμα2	Τελικά αποτελέσματα
1	1	1	0.0207	0.2858	on
2	-1	0	0.0434	0.1564	on
3	-1	2	0.0164	0.0664	on
4	-1	0	0.0400	0.1665	on
5	-1	0	0.0446	0.4071	on
6	-1	0	0.0280	0.3758	on
7	-1	0	0.0932	0.3273	on
8	0	0	0.0357	0.2381	on
9	-1	0	0.0483	0.5041	on
10	-1	-2	0.0465	0.1386	on
11	-1	-1	0.2493	0.0202	off
12	0	0	0.2002	0.0343	off

13	0	0	0.4141	0.0271	off
14	0	1	0.1310	0.0146	off
15	0	0	0.3697	0.0133	off
16	0	0	0.2942	0.0322	off
17	0	0	0.2096	0.0731	off
18	-1	0	0.4953	0.0164	off
19	-1	0	0.5644	0.0141	off
20	-1	0	0.2353	0.0194	off
Σύνολο επιτυχημένων προσπαθειών (on)			100%		
Σύνολο επιτυχημένων προσπαθειών (off)			100%		

Πίνακας 1 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "on" και "off "

Ο Πίνακας 2 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "Door" και "Key" σύμφωνα με τις πληροφορίες που δίνονται στο μέρος 3 Τα αποτελέσματα της προσομοίωσης 3.2.1 (2).

Δοκιμές	Μετατόπιση συχνότητας door	Μετατόπιση συχνότητας key	Σφάλμα1	Σφάλμα2	Τελικά αποτελέσματα
1	-2	28	No need	No need	Door
2	-1	30	No need	No need	Door
3	0	29	No need	No need	Door
4	0	29	No need	No need	Door
5	-1	34	No need	No need	Door
6	-1	30	No need	No need	Door
7	0	29	No need	No need	Door
8	-1	27	No need	No need	Door
9	-2	29	No need	No need	Door
10	-1	27	No need	No need	Door
11	-19	0	No need	No need	Key
12	-21	0	No need	No need	Key
13	-21	-1	No need	No need	Key
14	-20	0	No need	No need	Key
15	-22	0	No need	No need	Key
16	-20	0	No need	No need	Key
17	-20	0	No need	No need	Key
18	-21	0	No need	No need	Key
19	-20	0	No need	No need	Key
20	-20	0	No need	No need	Key
Σύνολο επιτυχημένων προσπαθειών (door)			100%		
Σύνολο επιτυχημένων προσπαθειών (key)			100%		

Πίνακας 2 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "Door" και "Key "

Ο Πίνακας 3 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "on" και "off" σύμφωνα με τις πληροφορίες που δίνονται στην αρχή του τμήματος Αποτελέσματα της Προσομοίωσης 3.2.1 (3) .

Δοκιμές	Μετατόπιση συχνότητας on	Μετατόπιση συχνότητας off	Σφάλμα1	Σφάλμα2	Τελικά αποτελέσματα
1	0	0	0.0888	0.2858	on
2	0	0	0.0979	0.2645	on
3	0	0	0.1073	0.3327	on
4	0	0	0.0430	0.1958	on
5	0	0	0.0075	0.0476	on
6	0	0	0.0885	0.1834	on
7	0	0	0.1121	0.0390	off
8	0	0	0.0281	0.1699	on
9	0	0	0.0755	0.0286	off
10	0	0	0.0389	0.3312	on
11	0	0	0.2289	0.0075	off
12	0	0	0.2316	0.1493	off
13	0	0	0.1519	0.0228	off
14	0	0	0.1123	0.0072	off
15	0	0	0.0240	0.0360	on
16	-1	0	0.2900	0.0245	off
17	0	0	0.1984	0.0162	off
18	0	0	0.2414	0.0526	off
19	0	-1	0.4284	0.0246	off
20	-1	0	0.1334	0.0269	off
Σύνολο επιτυχημένων προσπαθειών (on)			80%		
Σύνολο επιτυχημένων προσπαθειών (off)			90%		

Πίνακας 3 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "on" και "off "

Ο Πίνακας 4 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "Door" και "Key", σύμφωνα με τις πληροφορίες που δόθηκαν κατά την έναρξη του μέρους Αποτελέσματα της Προσομοίωσης 3.2.1 (4) .

Δοκιμές	Μετατόπιση συχνότητας door	Μετατόπιση συχνότητας key	Σφάλμα 1	Σφάλμα2	Τελικά αποτελέσματα
1	-2	15	No need	No need	Door
2	-1	15	No need	No need	Door

3	0	14	No need	No need	Door
4	0	14	No need	No need	Door
5	0	8	No need	No need	Door
6	-1	13	No need	No need	Door
7	0	13	No need	No need	Door
8	0	14	No need	No need	Door
9	0	13	No need	No need	Door
10	0	15	No need	No need	Door
11	-23	0	No need	No need	Key
12	-8	0	No need	No need	Key
13	-8	-1	No need	No need	Key
14	-16	0	No need	No need	Key
15	-16	0	No need	No need	Key
16	-14	0	No need	No need	Key
17	-20	0	No need	No need	Key
18	-13	0	No need	No need	Key
19	-13	0	No need	No need	Key
20	-12	0	No need	No need	Key
Σύνολο επιτυχημένων προσπαθειών (door)			100%		
Σύνολο επιτυχημένων προσπαθειών (key)			100%		

Πίνακας 4 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "Door" και "Key"

Ο Πίνακας 5 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "Door" και "Key", σύμφωνα με τις πληροφορίες που δόθηκαν κατά την έναρξη του μέρους Αποτελέσματα της Προσομοίωσης 3.2.1 (5) .

Δοκιμές	Μετατόπιση συχνότητας on	Μετατόπιση συχνότητας of	Σφάλμα1	Σφάλμα2	Τελικά αποτελέσματα
1	2	8	No need	No need	on
2	7	8	0.2055	0.4324	on
3	8	9	0.2578	0.2573	off
4	9	17	No need	No need	on
5	8	9	0.2304	0.3640	on
6	0	0	0.3268	0.6311	on
7	0	0	0.3193	0.3210	on
8	0	0	2.2153	0.9354	off
9	0	0	0.4603	0.1481	off
10	0	0	0.1189	0.0741	off
11	8	22	No need	No need	Door
12	8	0	No need	No need	Key
13	8	25	No need	No need	Door
14	8	24	No need	No need	Door

15	8	24	No need	No need	Door
16	-15	0	No need	No need	Key
17	-15	0	No need	No need	Key
18	-14	0	No need	No need	Key
19	-14	0	No need	No need	Key
20	-15	0	No need	No need	Key
Σύνολο επιτυχημένων προσπαθειών (για 20 φορές)			80%		

Πίνακας 5 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "on", "off", "Door" και "Key"

Στους πίνακες που ακολουθούν, το "m1" είναι η μέση τιμή του ελάχιστου μέσου τετραγωνικού σφάλματος για το σήμα αναφοράς 1. Το "m2" είναι η μέση τιμή του ελάχιστου μέσου τετραγωνικού σφάλματος για το σήμα αναφοράς 2. Ο ορισμός της μέσης τιμής των ελάχιστων μέσων τετραγωνικών σφαλμάτων βρίσκεται στο τμήμα 2.3.4 της παρούσας διατριβής.

Ο Πίνακας 6 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "on" και "off" σύμφωνα με τις πληροφορίες που δίνονται στην αρχή του παρόντος τμήματος (1) .

Δοκιμές	m1	m2	Τελικά αποτελέσματα
1	1.6740	3.4707	on
2	1.4442	2.6448	on
3	1.9087	3.7704	on
4	1.5448	2.9563	on
5	1.6103	3.4758	on
6	1.3971	3.7114	on
7	1.9205	3.3964	on
8	1.4944	3.7154	on
9	1.4716	3.7707	on
10	1.7879	4.5664	on
11	3.4775	0.1948	off
12	2.9213	1.0938	off
13	2.3013	0.7820	off
14	2.8370	1.1304	off
15	2.2277	0.7933	off
16	2.8922	1.8087	off
17	3.7633	1.6842	off
18	3.3211	1.5603	off
19	2.5852	0.8402	off
20	3.2708	1.4720	off
Σύνολο επιτυχημένων προσπαθειών (on)			100%

Σύνολο επιτυχημένων προσπαθειών (off)	100%
----------------------------------------------	-------------

Πίνακας 6 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "on" και "off"

Ο πίνακας 7 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "Door" και "Key" σύμφωνα με τις πληροφορίες που δίνονται στην αρχή του παρόντος τμήματος (2)

Δοκιμές	m1	m2	Τελικά αποτελέσματα
1	0.1837	5.2746	Door
2	0.7070	6.5936	Door
3	0.7565	8.4193	Door
4	0.2680	5.5084	Door
5	0.2973	6.0271	Door
6	0.8471	7.5534	Door
7	0.7039	7.8490	Door
8	0.5523	7.5952	Door
9	0.2467	5.7104	Door
10	0.3792	4.8257	Door
11	6.7231	0.0990	Key
12	8.7500	1.0829	Key
13	6.0756	0.1670	Key
14	8.1771	0.5392	Key
15	7.0094	0.2012	Key
16	7.9720	0.7285	Key
17	6.4771	0.4326	Key
18	6.3291	0.6853	Key
19	5.2181	0.4563	Key
20	5.1473	0.1231	Key
Σύνολο επιτυχημένων προσπαθειών (Door)			100%
Σύνολο επιτυχημένων προσπαθειών (Key)			100%

Πίνακας 7 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "Door" και "Key"

Ο πίνακας 8 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "on" και "off" σύμφωνα με τις πληροφορίες που δίνονται στην αρχή του παρόντος τμήματος (3) .

Δοκιμές	m1	m2	Τελικά αποτελέσματα
1	1.7351	3.5948	on
2	2.2663	3.9466	on
3	0.9809	2.7017	on
4	2.6011	2.7017	on
5	1.1137	1.7504	on
6	1.8093	2.9836	on
7	2.0433	2.9250	on
8	1.7423	2.1840	on
9	2.0782	3.7172	on
10	2.6252	4.8561	on
11	1.7214	0.5358	off
12	1.2078	0.6105	off
13	1.0685	0.8707	off
14	1.1108	0.7572	off
15	2.3774	0.7926	off
16	1.9559	0.8530	off
17	1.6756	0.8925	off
18	2.4116	0.8229	off
19	2.2664	0.5321	off
20	2.3713	0.4023	off
Σύνολο επιτυχημένων προσπαθειών (on)		100%	
Σύνολο επιτυχημένων προσπαθειών (off)		100%	

Πίνακας 8 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "on" και "off"

Ο Πίνακας 9 παρουσιάζει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "Door" και "Key" σύμφωνα με τις πληροφορίες που δίνονται στην αρχή του παρόντος τμήματος (4) .

Δοκιμές	m1	m2	Τελικά αποτελέσματα
1	0.3277	2.7586	Door
2	0.5835	3.3163	Door
3	0.2032	3.0460	Door
4	0.4941	3.0519	Door
5	1.4109	4.6609	Door
6	0.6575	3.7882	Door
7	1.1808	4.2020	Door
8	1.6448	3.4522	Door
9	0.5781	4.5291	Door
10	0.5794	4.6320	Door
11	3.9298	1.3690	Key

12	3.8828	0.8265	Key
13	3.9405	0.5655	Key
14	3.5953	0.7455	Key
15	4.1770	0.9070	Key
16	5.4789	1.5823	Key
17	4.1875	1.8230	Key
18	3.6388	1.3420	Key
19	3.5240	0.8551	Key
20	4.4077	0.4852	Key
Σύνολο επιτυχημένων προσπαθειών (Door)			100%
Σύνολο επιτυχημένων προσπαθειών (Key)			100%

Πίνακας 9 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "Door" και "Key"

Ο Πίνακας 10 δείχνει τα αποτελέσματα της προσομοίωσης για τα σήματα αναφοράς "on" και "off" σύμφωνα με τις πληροφορίες που δίνονται στην αρχή του παρόντος τμήματος (5) .

Δοκιμές	m1	m2	Τελικά αποτελέσματα
1	2.1943	3.5103	on
2	2.4698	1.7392	on
3	4.7760	3.3472	off
4	5.2285	4.6510	off
5	4.0235	3.9339	off
6	4.5611	3.8235	off
7	3.8988	4.2199	on
8	4.2026	3.2715	off
9	4.1196	3.2929	off
10	2.8686	1.9450	off
11	2.6504	1.4479	off
12	3.9870	2.2611	off
13	4.0748	2.2867	off
14	2.1320	0.8903	off
15	3.9950	2.3133	off
16	1.9331	0.7774	off
17	3.1565	1.9995	off
18	2.4163	1.1888	off
19	1.6946	0.6655	off
20	2.5083	1.2284	off
Σύνολο επιτυχημένων προσπαθειών (on)			30%
Σύνολο επιτυχημένων προσπαθειών (off)			100%

Πίνακας 10 : Τα αποτελέσματα της προσομοίωσης για τις λέξεις "on" και "off"

