Αναγνώριση και καταγραφή ήχων με χρήση μικροελεγκτή

Ηλίας Κατσιούλας

Πτυχιακή Εργασία

Επιβλέπων: Δημόπουλος Δημήτριος

Ιωάννινα, ΙΟΥΝΙΟΣ, 2019

ΤΜΗΜΑ ΜΗΧ. Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ

Πανεπιστημιο Ιωαννινών

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING** 

UNIVERSITY OF IOANNINA

### Εγκρίθηκε από τριμελή εξεταστική επιτροπή

Τόπος: Ημερομηνία:

### ΕΠΙΤΡΟΠΗ ΑΞΙΟΛΟΓΗΣΗΣ

Επιβλέπων Καθηγητής
Όνομα:
Επίθετο:
Τίτλος:
Βαθμίδα:

Μέλος επιτροπής
Όνομα:
Επίθετο:
Τίτλος:
Βαθμίδα:

Μέλος επιτροπής
Όνομα:
Επίθετο:
Τίτλος:
Βαθμίδα:

Ο/Η Προϊστάμενος/η του Τμήματος Όνομα: Επίθετο: Τίτλος: Βαθμίδα: © Κατσιούλας Ηλίας, 2019.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

### Δήλωση μη λογοκλοπής

Δηλώνω υπεύθυνα και γνωρίζοντας τις κυρώσεις του Ν. 2121/1993 περί Πνευματικής Ιδιοκτησίας, ότι η παρούσα πτυχιακή εργασία είναι εξ ολοκλήρου αποτέλεσμα δικής μου ερευνητικής εργασίας, δεν αποτελεί προϊόν αντιγραφής ούτε προέρχεται από ανάθεση σε τρίτους. Όλες οι πηγές που χρησιμοποιήθηκαν (κάθε είδους, μορφής και προέλευσης) για τη συγγραφή της περιλαμβάνονται στη βιβλιογραφία.

Κατσιούλας Ηλίας

### Ευχαριστίες

Πέρα από τις όποιες προσπάθειες και αν είγα καταβάλλει, δεν θα έφθανε η παρούσα μελέτη σε αίσιο τέλος αν δεν είχα την αμέριστη συμπαράσταση και καθοδήγηση από τον Επιβλέποντα της εργασίας μου κ. Δημήτρη Δημόπουλο, που απερίσπαστα, με συνέπεια, υπομονή και κατανόηση, έδινε τις κατάλληλες κάθε φορά κατευθύνσεις στις θεωρητικές και πειραματικές μου προσπάθειες. Ένα μεγάλο μέρος όμως της δύναμης και της διαρκούς προσπάθειας, για την ολοκλήρωση του παρόντος έργου, την οφείλω στους αγαπητούς μου Γονείς, Απόστολο και Αντιγόνη, που από την πρώτη μέρα με στήριζαν ηθικά σε όσες δυσκολίες αντιμετώπιζα. Δεν ξεχνώ όμως και την αμέριστη συμπαράσταση του αγαπητού μου αδελφού Φώτη, που θυσιάζοντας πολλές ώρες από τις καθημερινές του οικογενειακές του υποχρεώσεις, όντας προσφάτως πατήρ ενός απίθανου «μπέμπη», είχε τις δικές μου προσπάθειες σαν πρώτο μέλημα και προτεραιότητα. Μέσα από χρήσιμες συζητήσεις με τον φίλο μου Γιάννη, έγινε εφικτό να αντιμετωπιστούν αρκετά βασικά τεχνικά εμπόδια στην πορεία της μελέτης αυτής, και να δοθούν οι βέλτιστες κάθε φορά τεχνικές προσεγγίσεις. Τελικά, η όλη πορεία έγινε για μένα ένα μεγάλο μάθημα ζωής, ότι τελικά στον «πηγεμό για την Ιθάκη» σημασία έχει ο «δρόμος» που διαβαίνουμε, και ειδικά όταν στον «δρόμο» αυτό συναντούμε την αγάπη και την συνεργασία των συνανθρώπων μας.

### Περίληψη

Η επεξεργασία σημάτων είναι, στα τελευταία χρόνια, ένα πλούσιο πεδίο έρευνας και επιτεύξεων. Το «σήμα», ως ένα φυσικό μέγεθος, όπως είναι το «ηχητικό σήμα», είναι ένα δυναμικά μεταβαλλόμενο μέγεθος που μπορεί να υποστεί επεξεργασία από ένα επεξεργαστικό σύστημα όπως είναι οι σύγχρονοι «μικροελεγκτές» (μC). Οι δυνατότητες τέτοιων συστημάτων τα επιτρέπουν να μπορούν να χρησιμοποιηθούν σε μία μεγάλη ποικιλία εφαρμογών, που περιλαμβάνουν επίσης και τη διασύνδεσή τους με πληθώρα άλλων συσκευών και Η/Υ, δίνοντας αυτονομία και αξιοπιστία στις λειτουργίες τους.

Έτσι, στην παρούσα εργασία γίνεται υλοποίηση της καταγραφής και αναγνώρισης ήχων», με τη χρήση ενός μικρουπολογιστικού συστήματος «αρμονικών (μικροελεγκτή) (μC) σε διασύνδεση μέσω UART με H/Y (pc), και του σχετικού Λογισμικού, που αναπτύχθηκε σε γλώσσα C. Όλα τα βήματα επεξεργασίας γίνονται αποκλειστικά στον μικροελεγκτή. Η «καταγραφή» των ηχητικών αρμονικών σημάτων γίνεται με τη βοήθεια μικροφώνου που συνδέεται κατευθείαν στο ADC του μικροελεγκτή και στη συνέχεια υπολογίζεται το φάσμα συχνοτήτων του αρμονικού σήματος, κατά την μεθοδολογία Fourier, και από κει και πέρα υπολογίζονται οι συντελεστές συγνοτήτων MFCC (Mel frequency cepstral coefficients), με τη χρήση των φίλτρων του τύπου «συχνότητες mel» (mel frequency filterbank), σε ένα μικρό γρονικό παράθυρο μερικών sec, δίνοντας έτσι ένα ψηφιακό ID του αντίστοιγου ήγου. Το ID του ήχου καταγράφεται σε «βάση δεδομένων» στην μόνιμη μνήμη ΕΕΡROM του μικροελεγκτή. Οι αναζητήσεις, οι απαραίτητες για το «ταίριασμα» ήχου, βασίζονται στη σύγκριση της Manhattan μετρικής, μέσω του αθροίσματος των «αποστάσεων» των αντίστοιχων συντελεστών MFCC, ενώ επιλέγεται από τη «βάση δεδομένων» για «ταίριασμα» το στοιχείο που δίνει τη μικρότερη μετρική απόσταση.

Ο σκοπός της μελέτης είναι να αποτελέσει ένα οδηγό για παρόμοιες υλοποιήσεις σε μικροελεγκτές, και για πιο εκτεταμένες εφαρμογές, όπως για παράδειγμα σε «αναγνώριση» και «ταυτοποίηση» ομιλίας, σε ελέγχους με ηχητικά μέσα των κτιριακών και μεταλλικών κατασκευών,κ.α.. Επίσης, παρόμοιες τέτοιες μεθοδολογίες μπορούν να επεκταθούν και στις Τηλεπικοινωνιακές συσκευές όπως για παράδειγμα στην Κινητή τηλεφωνία, όπου η ταυτοποίηση των συνομιλητών είναι ένα επιθυμητό στοιχείο στα πλαίσια των αντίστοιχων παρεχόμενων υπηρεσιών.

Λέξεις - Κλειδιά: Μικροελεγκτής, Ήχοι, ADC, φάσμα, MFCC.

### Abstract

The area of «signal processing» is an upgoing field of research and applications. The concept of «signal» as being a dynamically varying physical entity, such as the «sound waves» are, may prove useful in wide area applications when the processing of such quantities is carried out by microcontrollers ( $\mu$ C). Microcontollers can be used in a variety of applications involving their interconnections with other similar devices and computer machines. In this way, microcontrollers exhibit fast and autonomous operations and also reliable performance in many technological applications.

Thus, in the present study, a system of recording and identifying of acoustic tones has been developed, utilizing a microcontroller system ( $\mu$ C) at an interface with a personal computer via the UART protocol (RS232) and the necessary, for that application, software, that is developed in C# language. All the processing steps are performed solely by the microcontroller. The recording of sounds, using a microphone device connected to microcontroller, is followed by an ADC conversion and the estimation of the Fourier spectrum of the digital signal. Subsequently, the MFCC related cepstral coefficients, for a short time window of few seconds, are calculated and considered to constitute an ID type of the recorded sound. The IDs of some specific for the purposes of the study frequencies are stored in the permanent memory (EEPROM) of the microcontroller, thus forming a kind of stored «database» of IDs of harmonic sounds. The identification of the initially unknown frequency is based on the comparisons of the Manhattan metrics for the distance, by using the sum of the distances of the related MFCC coefficients, for the sounds already stored in the «database». The matcing sound frequency is selected to be that of the recorded element that yields the «minumum» distance among the values.

The purpose of the study is to serve as an example guide to further and more complicated applications in voice identification and recognition, and especially in cases whereas complex sound signals are involved such as the sound tests of civil and mechanical constructions, etc. Also, such or similar techniques can be applied in Mobile telecommunications, where «user identification» and the related information is an essential part of the provided services.

Keywords: Microcontroller, Sounds, ADC, Spectrum, MFCC.

## ΠΕΡΙΕΧΟΜΕΝΑ

Αναγνώριση και καταγραφή ήχων με χρήση μικροελεγκτή1
Ηλίας Κατσιούλας1
Πτυχιακή Εργασία1
Επιβλέπων: Δημόπουλος Δημήτριος1
Ιωάννινα, ΙΟΥΝΙΟΣ, 20191
Ευχαριστίες5
Περίληψη6
Λέξεις – Κλειδιά:6
Abstract7
Keywords:7
REPIEXOMENA
ΕΙΚΟΝΕΣ
ΣΧΗΜΑΤΑ
ΠΙΝΑΚΕΣ
Συντομογραφίες & Ακρωνύμια16
ΚΕΦΑΛΑΙΟ 1. Ήχος και Συχνότητες17
1.1 Εισαγωγή
1.2. Φάσμα σήματος - Μετασχηματισμός Fourier19
1.2.1 Αλγόριθμοι Επεξεργασίας – FFT
Σχήμα 1.1: Φάσμα συχνοτήτων αρμονικού σήματος σε $\pm f0$
Σχήμα 1.2: Σχήμα «πεταλούδας» του DFT23
Εικόνα 1.1: Δειγματοληψία και Κβάντιση σημάτων24
1.3 Οι Συντελεστές MFCC
1.3.1 Γενικά στοιχεία για τους συντελεστές MFCC

Εικόνα 1.2 Γραφική απεικόνιση της κλίμακας mel	26
Εικόνα 1.3: Ομάδα φίλτρων συχνοτήτων mel	27
1.3.2 Μέθοδος υπολογισμού των συντελεστών MFCC	27
Εικόνα 1.4: Διάγραμμα υπολογισμού των συντελεστών MFCC	28
1.3.3 Αναγνώριση ήχου με βάση τους συντελεστές MFCC	31
ΚΕΦΑΛΑΙΟ 2. ΣΤΟΙΧΕΙΑ ΤΟΥ ΥΛΙΚΟΥ ΕΞΟΠΛΙΣΜΟΥ	32
2.1 Υλικός Εξοπλισμός	32
Εικόνα 2.1: Ο XMega256A3U της ATMEL (Πηγές WWW-)	37
Εικόνα 2.2: Εσωτερική αρχιτεκτονική του XMega256	38
Εικόνα 2.3: Οι καταχωρητές γενικού σκοπού του AVR (Πηγές WWW-)	40
Εικόνα 2.4: Οι διπλοί καταχωρητές Χ- , Υ- , Ζ- του ΑVR (Πηγές WWW-)	40
Εικόνα 2.5 Εκτέλεση εντολών από την ALU	41
Εικόνα 2.6: Στάδια εκτέλεσης εντολής από την ALU	41
Εικόνα 2.7: Καταχωρητές θύρας εισόδου – εξόδου του AVR	44
Σχήμα 2.1: Ο χρονιστής στους μικροελεγκτές Atmel	45
Σχήμα 2.2: Ο προκλιμακωτής (prescaler) στους μικροελεγκτές Atmel	45
Η Οθόνη LCD (Liquid Crystal Display)	49
Εικόνα 2.8: Οθόνη LCD 2X16	49
Το αναπτυξιακό STK600	50
Εικόνα 2.9: Το αναπτυξιακό STK 600 (Πηγές WWW-)	51
Το πρωτόκολλο επικοινωνίας UART μεταξύ μικροελεγκτή και PC	51
Σχήμα 2.3: Σειριακή μετάδοση των bits	52
ΚΕΦΑΛΑΙΟ 3. ΜΕΘΟΔΟΛΟΓΙΑ (ΥΛΙΚΟ-ΛΟΓΙΣΜΙΚΟ-ΜΕΤΡΗΣΕΙΣ)	53
3.1 Περιγραφή του υλικού εξοπλισμού-Λογισμικού	53
Εικόνα 3.1:Λογισμικό για UART επικοινωνία μC- Η/Υ	55
Εικόνα 3.2: Γεννήτρια συχνοτήτων	55
Εικόνα 3.3: Παλμογράφος(απεικονιζόμενο σήμα 1KHz)	56

Εικόνα 3.4: Το πειραματικό setup	56
(αναπτυξιακό STK600, μC, microphone, LCD, Παλμογράφος)	56
Σχήμα 3.1: Block διάγραμμα συνδέσεων των συσκευών	57
3.2 Καταγραφές Αρμονικών ήχων	57
Πίνακας 3.1: Αρμονικές συχνότητες	58
3.3 Υπολογισμός συντελεστών MFCC των αρμονικών συχνοτήτων και αποθήκευση σε EEPROM	59
Σχήμα 3.2: Τα στάδια υπολογισμού των συντελεστών MFCC	59
Πίνακας 3.2: Οι συντελεστές MFCC (C <sub>0</sub> έως C <sub>14</sub> )	62
Πίνακας 3.3 Οι συντελεστές MFCC (C <sub>15</sub> έως C <sub>29</sub> )	63
3.4 Ψηφιακό αποτύπωμα ήχων	63
Σχήμα 3.3: Γεωμετρική ερμηνεία της απόστασης «Manhattan»	65
Πίνακας 3.4: Συντελεστές MFCC για αρμ. συχνότητες 680 Hz-2500 Hz	67
ΚΕΦΑΛΑΙΟ 4. ΕΠΕΞΕΡΓΑΣΙΑ ΜΕΤΡΗΣΕΩΝ	68
4.1 Φάσμα σήματος με FFT – Ενδεικτικά γραφήματα	68
Εικόνα 4.1: Στιγμιαίο πλάτος σήματος x(t) με fo=500Hz, MIC	68
Εικόνα 4.2: Πλάτος ισχύος σήματος (fo=500Hz, MIC) με #bin συχνότητας	69
Εικόνα 4.3: Στιγμιαίο πλάτος σήματος x(t) με fo=1000Hz, MIC	69
Εικόνα 4.4: Πλάτος ισχύος σήματος (fo=1000Hz, MIC) με συχνότητα	70
Εικόνα 4.5: Στιγμιαίο πλάτος σήματος x(t) με fo=1500Hz, MIC	70
Εικόνα 4.6: Πλάτος ισχύος σήματος (fo=1500Hz, MIC) με συχνότητα	71
Εικόνα 4.7: Στιγμαίο πλάτος σήματος x(t) με fo=2000Hz, MIC	72
Εικόνα 4.8: Πλάτος ισχύος σήματος (fo=2000Hz, MIC) με #bin συχνότητας	72
Εικόνα 4.9: Πλάτος ισχύος σήματος (fo=2000Hz, MIC) με συχνότητα	73
Εικόνα 4.10: Στιγμαίο πλάτος σήματος x(t) με fo=2500Hz, MIC	73
Εικόνα 4.11:Πλάτος ισχύος σήματος (fo=2000Hz,MIC) με #bin συχνότητας	74
Εικόνα 4.12: Πλάτος ισχύος σήματος (fo=2500Hz,MIC) με συχνότητα	74
Εικόνα 4.13: Συντελεστές MFCC για fo=500Hz	75

(μετρούμενη και θεωρητική καμπύλη)	75
Εικόνα 4.14: Συντελεστές MFCC για fo=1000Hz	76
(μετρούμενη και θεωρητική καμπύλη)	76
Εικόνα 4.15: Συντελεστές MFCC για fo=1500Hz	76
(μετρούμενη και θεωρητική καμπύλη)	76
Εικόνα 4.16: Συντελεστές MFCC για fo=2000Hz	77
(μετρούμενη και θεωρητική καμπύλη)	77
Εικόνα 4.17: Συντελεστές MFCC για fo=2500Hz	77
(μετρούμενη και θεωρητική καμπύλη)	77
Πίνακας 4.1: «Αποστάσεις» της προς ταυτοποίηση συχνότητας,	78
με τους συντελεστές MFCC	78
Πίνακας 4.2: «Αποστάσεις» της προς ταυτοποίηση συχνότητας,	78
με τους συντελεστές MFCC	78
Πίνακας 4.3: «Αποστάσεις» της προς ταυτοποίηση συχνότητας,	79
με τους συντελεστές MFCC	79
Πίνακας 4.4: «Αποστάσεις» της προς ταυτοποίηση συχνότητας,	79
με τους συντελεστές MFCC	79
ΚΕΦΑΛΑΙΟ 5. Συμπεράσματα-Επίλογος	80
5.1 Ανάλυση δεδομένων από τις καταγραφές – ταυτοποιήσεις	80
5.2 Συμπεράσματα – Προτάσεις	81
5.3 ΕΠΙΛΟΓΟΣ	82
ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΠΑΡΑΠΟΜΠΕΣ	84
Πηγές στο Internet	87
В )	93

## ΕΙΚΟΝΕΣ

ΕΙΚΟΝΕΣ	12
Εικόνα 1.2 Γραφική απεικόνιση της κλίμακας mel	26
Εικόνα 1.3: Ομάδα φίλτρων συχνοτήτων mel	27
Εικόνα 1.4: Διάγραμμα υπολογισμού των συντελεστών MFCC	28
Εικόνα 2.1: Ο XMega256A3U της ATMEL (Πηγές WWW-)	
Εικόνα 2.2: Εσωτερική αρχιτεκτονική του XMega256	
Εικόνα 2.3: Οι καταχωρητές γενικού σκοπού του ΑVR (Πηγές WWW-)	
Εικόνα 2.4: Οι διπλοί καταχωρητές Χ- , Υ- , Ζ- του ΑVR (Πηγές WWW-)	
Εικόνα 2.5 Εκτέλεση εντολών από την ALU	41
Εικόνα 2.6: Στάδια εκτέλεσης εντολής από την ALU	
Εικόνα 2.7: Καταχωρητές θύρας εισόδου – εξόδου του ΑVR	
Εικόνα 2.8: Οθόνη LCD 2X16	
Εικόνα 2.9: Το αναπτυξιακό STK 600 (Πηγές WWW-)	51
Εικόνα 3.1:Λογισμικό για UART επικοινωνία μC- Η/Υ	55
Εικόνα 3.2: Γεννήτρια συχνοτήτων	55
Εικόνα 3.3: Παλμογράφος(απεικονιζόμενο σήμα 1KHz)	
Εικόνα 3.4: Το πειραματικό setup	
(αναπτυξιακό STK600, μC, microphone, LCD, Παλμογράφος)	56
Εικόνα 4.1: Στιγμιαίο πλάτος σήματος x(t) με fo=500Hz, MIC	
Εικόνα 4.2: Πλάτος ισχύος σήματος (fo=500Hz, MIC) με #bin συχνότητας	
Εικόνα 4.3: Στιγμιαίο πλάτος σήματος x(t) με fo=1000Hz, MIC	69
Εικόνα 4.4: Πλάτος ισχύος σήματος (fo=1000Hz, MIC) με συχνότητα	70
Εικόνα 4.5: Στιγμιαίο πλάτος σήματος x(t) με fo=1500Hz, MIC	70
Εικόνα 4.6: Πλάτος ισχύος σήματος (fo=1500Hz, MIC) με συχνότητα	71
Εικόνα 4.7: Στιγμαίο πλάτος σήματος x(t) με fo=2000Hz, MIC	72

Εικόνα 4.8: Πλάτος ισχύος σήματος (fo=2000Hz, MIC) με #bin συχνότητας72
Εικόνα 4.9: Πλάτος ισχύος σήματος (fo=2000Hz, MIC) με συχνότητα73
Εικόνα 4.10: Στιγμαίο πλάτος σήματος x(t) με fo=2500Hz, MIC73
Εικόνα 4.11:Πλάτος ισχύος σήματος (fo=2000Hz,MIC) με #bin συχνότητας
Εικόνα 4.12: Πλάτος ισχύος σήματος (fo=2500Hz,MIC) με συχνότητα
Εικόνα 4.13: Συντελεστές MFCC για fo=500Hz75
(μετρούμενη και θεωρητική καμπύλη)75
Εικόνα 4.14: Συντελεστές MFCC για fo=1000Hz76
(μετρούμενη και θεωρητική καμπύλη)76
Εικόνα 4.15: Συντελεστές MFCC για fo=1500Hz76
(μετρούμενη και θεωρητική καμπύλη)76
Εικόνα 4.16: Συντελεστές MFCC για fo=2000Hz77
(μετρούμενη και θεωρητική καμπύλη)77
Εικόνα 4.17: Συντελεστές MFCC για fo=2500Hz77
(μετρούμενη και θεωρητική καμπύλη)77

## ΣΧΗΜΑΤΑ

ΣΧΗΜΑΤΑ	. 14
Σχήμα 2.1: Ο χρονιστής στους μικροελεγκτές Atmel	. 45
Σχήμα 2.2: Ο προκλιμακωτής (prescaler) στους μικροελεγκτές Atmel	. 45
Σχήμα 2.3: Σειριακή μετάδοση των bits	. 52
Σχήμα 3.1: Block διάγραμμα συνδέσεων των συσκευών	. 57
Σχήμα 3.2: Τα στάδια υπολογισμού των συντελεστών MFCC	. 59
Σχήμα 3.3: Γεωμετρική ερμηνεία της απόστασης «Manhattan»	. 65

## ΠΙΝΑΚΕΣ

ΠΙΝΑΚΕΣ	15
Πίνακας 3.1: Αρμονικές συχνότητες	58
Πίνακας 3.2: Οι συντελεστές MFCC (C <sub>0</sub> έως C <sub>14</sub> )	62
Πίνακας 3.3 Οι συντελεστές MFCC (C <sub>15</sub> έως C <sub>29</sub> )	63
Πίνακας 3.4: Συντελεστές MFCC για αρμ. συχνότητες 680 Hz-2500 Hz	67
Πίνακας 4.1: «Αποστάσεις» της προς ταυτοποίηση συχνότητας,	78
με τους συντελεστές MFCC	78
Πίνακας 4.2: «Αποστάσεις» της προς ταυτοποίηση συχνότητας,	78
με τους συντελεστές MFCC	78
Πίνακας 4.3: «Αποστάσεις» της προς ταυτοποίηση συχνότητας,	79
με τους συντελεστές MFCC	79
Πίνακας 4.4: «Αποστάσεις» της προς ταυτοποίηση συχνότητας,	79
με τους συντελεστές MFCC	79

## Συντομογραφίες & Ακρωνύμια

μC	μικροελεγκτής (microcontroller)
FFT	Fast Fourier Transform (Ταχύς Μετασχηματισμός Φουριέ)
DCT	Discrete Cosine Transform(Διακριτός Μετασχηματισμός Συνημιτόνου)
DFT	Discrete Fourier Transform (Διακριτός Μετασχηματισμός Φουριέ)
LCD	Liquid Crystal Display (Οθόνη Υγρών Κρυστάλλων)
AC	Alternating Current (Εναλασσόμενο)
KHz	KiloHertz ( x 1000Hz)
MHz	MegaHertz (x 1000 KHz)
GHz	GigaHertz (x 1000 MHz)
РСМ	Pulse Code Modulation
UART	Unified Asyncronous Receive Transmit
USB	Univied Serial Bus
COM Port	Θύρα σειριακής επικοινωνίας
I/0	Είσοδος / Έξοδος
MFCC	Mel Frequency Cepstral Coefficients

## ΚΕΦΑΛΑΙΟ 1. Ήχος και Συχνότητες

### 1.1 Εισαγωγή

Η επεξεργασία σήματος με τη χρήση μικροελεγκτών (μC) είναι ένας τομέας εξελισσόμενος τα τελευταία χρόνια, τόσο στην ανάπτυξη ταχύτατων μικροελεγκτών (2-3 GHz, 2018) όσο και στην ικανότητα εφαρμογής σύγχρονων αλγορίθμων επεξεργασίας αναλογικών και ψηφιακών σημάτων σε πολλαπλά κανάλια επεξεργασίας. Η αναγκαιότητα για τη χρήση τέτοιων συστημάτων προκύπτει από πολλές περιπτώσεις που γίνεται χρήση τέτοιων σημάτων (ραδιόφωνο, αναγνώριση ομιλίας, συστήματα για συμπίεση, μίξη ήχου κ.α.) (Leave 2010).

Σε τυπικές περιπτώσεις χρήσης τέτοιων συστημάτων, το σήμα είναι στην περιοχή των ακουστικών συχνοτήτων (20Hz-20KHz) και οι εφαρμογές που υλοποιούν τέτοια σήματα έχουν τουλάχιστον μια βαθμίδα μετατροπής του αναλογικού σε ψηφιακό σήμα (PCM, Δέλτα). (Oliver, Pierce & Shannon 1948). Σε κάθε τέτοια περίπτωση, σημαντικό ρόλο στις προδιαγραφές του συστήματος παίζει το «φάσμα» συχνοτήτων του σήματος, όπως αυτό αρχικά καθορίζεται από την πηγή του σήματος. Άλλες βαθμίδες επεξεργασίας μπορεί να περιλαμβάνουν κωδικοποίηση του παραγόμενου ψηφιακού σήματος, κρυπτογράφηση, διαμόρφωση φέροντος με την «πληροφορία» να είναι το ακουστικών συχνοτήτων σήμα κλπ.

Στην παρούσα εργασία γίνεται χρήση των αποτελεσμάτων του «φάσματος» συχνοτήτων για ένα ηχητικό σήμα εισόδου, ώστε να αναγνωρίζεται σε πρώτο στάδιο η αρμονική συχνότητα και σε δεύτερο στάδιο να δημιουργείται ένα ψηφιακό αποτύπωμα για ένα ηχητικό σήμα μικρής διάρκειας (<5 sec) για το οποίο γίνεται αναγνώριση φασματικού περιεχομένου με το να υπολογίζονται οι συντελεστές MFCC (Mel Frequency Cepstral Coefficients), (Davis & Mermelstein 1980), για ορισμένες συχνότητες, καταγράφοντας έτσι ένα είδος «ταυτότητας» (ID) στη μόνιμη μνήμη του μικροελεγκτή. Ένας αρμονικός ήχος που εισάγεται στη συνέχεια για αναγνωριστεί κατάλληλα.

Ένα σύστημα αναγνώρισης αρμονικών ήχων μπορεί να είναι ως βάση για ένα πιο περίπλοκο σύστημα επεξεργασίας ήχων. Το σύστημα που περιγράφεται στην παρούσα εργασία διατηρεί τη φιλοσοφία ενός απλοποιημένου στη σχεδίαση κυκλώματος συνδέσεων αλλά όμως χρησιμοποιεί το δυναμικό του μικροελεγκτή (μC) σε λειτουργίες όπως η ελεγχόμενη είσοδος, η μετατροπή αναλογικού σε ψηφιακό σήμα, οι διακοπές, και η έξοδος της πληροφορίας (bits) σε θύρες εξόδου αλλά και τις δυνατότητες σύνδεσης του μC με περιφερειακές συσκευές (UART επικοινωνία με pc). (Grace 2015). Ο σκοπός της υλοποίησης της κατασκευής είναι να γίνεται «αναγνώριση», των ηχητικών αρμονικών συχνοτήτων με βάση ένα «ψηφιακό αποτύπωμα» που έχει εκ των προτέρων καταγραφεί και αποθηκευθεί. Η θεωρία που υποστηρίζει αυτές τις λειτουργίες βασίζεται στη χρήση του FFT (Fast Fouriér Transform) (Γρήγορος Μετασχηματισμός Φουριέ),(Cooley et al. 1965), για την απόκτηση κάθε φορά του φασματικού περιεχομένου του ηχητικού σήματος εισόδου. Η λειτουργία του FFT επιτελείται από κατάλληλο Λογισμικό που «τρέχει» στον μικροελεγκτή. Η κατασκευή χρησιμοποιεί τον μικροελεγτή XMega256A3U της Atmel Co., και το αναπτυξιακό STK600 για τον προγραμματισμό του μC και της λειτουργίες του με τις περιφερειακές συσκευές (Oθόνη LCD και UART θύρα com με H/Y). (Tarafder 2017).

Το Λογισμικό περιλαμβάνει κώδικα σε C που μεταγλωττίζεται με τη βοήθεια του μεταγλωττιστή gcc του περιβάλλοντος ATMEL AVR Studio (7.0) (Πηγές WWW), και «φορτώνεται» στον μικροελεγκτή ως δεκαεξαδικό αρχείο εντολών (elf αρχείο). Ο κώδικας περιλαμβάνει τη χρήση των διακοπών του μικροελεγκτή και της μνήμης SRAM και EEPROM για την «εκτέλεση» του Λογισμικού και την αποθήκευση στοιχείων, αντίστοιχα. (Barnett et al, 2006; Πηγές στο Internet).

Η εργασία διαχωρίζεται σε πέντε (5) κεφάλαια και ένα παράρτημα. Στο πρώτο κεφάλαιο (ΚΕΦΑΛΑΙΟ 1) περιγράφονται τα βασικά στοιχεία της θεωρίας του φάσματος συχνοτήτων σημάτων με μετασχηματισμούς Fouriér, και η μετατροπή αναλογικού σήματος σε ψηφιακό με συστήματα ADC. Στο ίδιο κεφάλαιο περιγράφεται η διαδικασία για την εξαγωγή των συντελεστών MFCC για το «ψηφιακό» αποτύπωμα. Στο επόμενο κεφάλαιο (ΚΕΦΑΛΑΙΟ 2) περιγράφονται τα βασικά χαρακτηριστικά των μικροελεγκτών με έμφαση για τον XMega256A3U, διότι είναι και η κύρια επεξεργαστική μονάδα της κατασκευής. Επίσης, περιγράφονται οι συσκευές της οθόνης υγρών κρυστάλλων και του αναπτυξιακού STK600, καθώς και η διασύνδεση των μικροελεγκτών με Η/Υ μέσω του UART πρωτοκόλλου σειριακής επικοινωνίας. Στο κεφάλαιο αυτό περιγράφεται επίσης το κύκλωμα με σχηματικό διάγραμμα και διάγραμμα ροής και ο τρόπος υλοποίησης της κατασκευής. Στο τρίτο κεφάλαιο (ΚΕΦΑΛΑΙΟ 3) παρατίθεται η μεθοδολογία που ακολουθείται τόσο για το υλικό και το Λογισμικό όσο και για τις μετρήσεις που λαμβάνονται, για του

αρμονικούς ήχους που καταγράφονται. Τα αποτελέσματα από τις «πειραματικές» καταγραφές και μετρήσεις, με γραφικές απεικονίσεις και με το εύρος των αρμονικών συχνοτήτων να κυμαίνονται στα όρια από 500 Hz έως 2500 Hz, περιγράφονται στο τέταρτο κεφάλαιο (ΚΕΦΑΛΑΙΟ 4). Σε αυτό το στάδιο περιγράφονται και ορισμένες ταυτοποιήσεις αρμονικών ήγων, έγοντας αξιοποιήσει τις προηγούμενες καταγραφές και τους σχετικούς υπολογισμούς των συντελεστών MFCC. Στο επόμενο κεφάλαιο (ΚΕΦΑΛΑΙΟ 5) περιγράφεται η ανάλυση των αποτελεσμάτων για τις καταγραφές και τις αναγνωρίσεις με χρήση του «ψηφιακού» αποτυπώματος, όπως επίσης και οι επεκτάσεις που μπορούν να δοθούν σε παρόμοιες πειραματικές διατάξεις. Τέλος, περιγράφεται ένα τμήμα ως επίλογος της μελέτης και παρατίθενται προτάσεις και κατευθύνσεις για τη βελτίωση του εύρους των εφαρμογών παρόμοιων πειραματικών υλοποιήσεων και τεχνικών. Σε συμπλήρωμα όλων των προηγούμενων εδαφίων, παρατίθεται και η βιβλιογραφία της εργασίας. Η τεκμηρίωση του κώδικα για επεξεργασία FFT και υπολογισμό των συντελεστών MFCC παρατίθεται στο Παράρτημα Ι, στο μέρος Α, ενώ ο κώδικας για την UART επικοινωνία του μικροελεγκτή με τον Η/Υ παρατίθεται στο μέρος Β, του Παραρτήματος.

## **1.2. Φάσμα σήματος - Μετασχηματισμός Fourier** 1.2.1 Αλγόριθμοι Επεξεργασίας - FFT

Στην επεξεργασία σήματος ακουστικών συχνοτήτων μπορούν να εφαρμοστούν ποικίλες τεχνικές επεξεργασίας αλλά κάθε μία προσανατολίζεται σε ένα συγκεκριμένο επιθυμητό αποτέλεσμα. Για παράδειγμα, το σήμα μπορεί να μετατραπεί αρχικά σε ψηφιακό (AC σε DC) και στη συνέχεια τα ψηφιακά δείγματα να χρησιμοποιηθούν με πολλού τρόπους ανάλογα με την εκάστοτε αναγκαιότητα. Τέτοιες διαδικασίες χρησιμοποιούν ένα «παράθυρο» χρόνου, δηλαδή ένα πεπερασμένο χρονικό διάστημα, μέσα στο οποίο παίρνουν «δείγματα» του σήματος και στη συνέχεια με τεχνικές όπως η PCM (Stallings 1984), μετατρέπουν τα δείγματα αυτά σε τιμές σε ψηφιακή κλίμακα. Σε άλλες τεχνικές γίνεται αξιοποίηση τέτοιων δειγμάτων για μελέτη ή εξαγωγή του φασματικού περιεχομένου του ακουστικού σήματος. Μια τέτοια κατηγορία τεχνικών είναι αυτή που βασίζεται στον μαθηματικό φορμαλισμό του μετασχηματισμού Fouriér για τη φασματική κατανομή A(f) του σήματος x(t). Σε αυτό τον φορμαλισμό έχει αναπτυχθεί και η μεθοδολογία που αποκαλείται FFT (Fast Fouriér Transform), (Ταχύς μετασχηματισμός Φουριέ), (Cooley & Tukey 1965; Ramirez 1984; Grant et al. 2002; Champeney 1973; Leaver 2010), κατά την οποία απαιτούνται λιγότερα και πιο «συνοπτικά» βήματα στη μετατροπή από x(t) σε A(f).

#### Στοιχεία του Μετασχηματισμού Fouriér

Σύμφωνα με τον γενικό μετασχηματισμό Fouriér, (Robert W. Ramirez, 1984), ένα σήμα x(t) αναλύεται σε ένα φάσμα συχνοτήτων A(f) σύμφωνα με τη σχέση:

$$A(f) = \int_{-\infty}^{+\infty} e^{-j2\pi ft} x(t) dt$$

όπου A(f) παριστάνει το (μιγαδικό) πλάτος του φάσματος στη συχνότητα f.

Αντίστροφα, με δεδομένα τα A(f), το (μιγαδικό) σήμα x(t) μπορεί να αναπαραχθεί σύμφωνα με τη σχέση (Duhamel, Piron & Etcheto 1988):

$$x(t) = \int_{-\infty}^{+\infty} e^{j2\pi ft} A(f) df$$

Μπορεί κανείς να έχει μια άμεση εικόνα αυτού του μετασχηματισμού, αν θεωρήσει το x(t) να είναι ένα «αρμονικό» σήμα, δηλαδή

$$x(t) = x_0 \cos(2\pi f_0 t)$$

Σε αυτή την περίπτωση το «φάσμα» του σήματος x(t) είναι

$$A(f) = \begin{cases} A_0 & \gamma \iota \alpha & f = \pm f_0 \\ 0 & \gamma \iota \alpha & \alpha \lambda \lambda o \psi \end{cases}$$

Και με γραφική απεικόνιση:



Σχήμα 1.1: Φάσμα συχνοτήτων αρμονικού σήματος σε  $\pm f_0$ 

#### Διακριτός Μετασχηματισμός Fouriér

Σε πολλές περιπτώσεις δεν είναι δυνατό να είναι διαθέσιμες οι τιμές του σήματος x(t) σε άπειρο χρονικό διάστημα. Αντίθετα, είναι γνωστές οι τιμές x(t) σε ένα πεπερασμένο χρονικό διάστημα Τ, που λέγεται και «παράθυρο χρόνου». Πολλά συστήματα που χρησιμοποιούν τέτοια παράθυρα έχουν πρόσβαση στις τιμές x(t) σε ορισμένες στιγμές  $t_0$ ,  $t_1$ ,  $t_2$ ,  $t_3$ ,..., $t_k$  εντός του χρονικού παραθύρου. Έτσι, είναι διαθέσιμες ορισμένες μόνο τιμές  $x_0$ ,  $x_1$ ,  $x_2$ ,  $x_3$ ,..., $x_k$  σε αντίστοιχες χρονικές στιγμές  $t_0$ ,  $t_1$ ,  $t_2$ ,  $t_3$ , ..., $t_k$ . Με τη βοήθεια της θεωρίας του Διακριτού Μετασχηματισμού Fouriér (DFT) (Cooley, Lewis & Welch 1969) (μπορούμε να αποκτήσουμε το «φάσμα» συχνοτήτων του σήματος αυτού σύμφωνα με τη σχέση:

$$X(k) = \sum_{n=0}^{N-1} e^{-j2\pi n/N} X_n$$

όπου οι ω<sub>k</sub> είναι:

$$\omega_k = \left(\frac{2\pi}{N}\right)k, \mu\varepsilon \ k = 0, 1, 2 \dots N - 1$$

Δηλαδή, οι διαδοχικές  $ω_k$  ισαπέχουν διάστημα  $\left(\frac{2\pi}{N}\right)$ , ενώ τα X( $ω_k$ ) είναι τα πλάτη στις αντίστοιχες συχνότητες. Όμοια ορίζεται και ο αντίστροφος Διακριτός Μετασχηματισμός Fouriér (i-DFT) (Duhamel et al. 1988; Mix 2015), σύμφωνα με τη σχέση:

$$x_n = \left(\frac{1}{N}\right) \sum_{k=0}^{N-1} e^{j2\pi kn/N} X(k)$$

Στην περίπτωση ενός «απείρου» σήματος, χρησιμοποιείται ο Διακριτός Μετασχηματισμός Fouriér Διακριτού χρόνου (DFDT), που σε περίπτωση που το

N->
$$\infty$$
 είναι: $X(\omega) = X(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} e^{-j\omega n} x_n$  με  $\omega E(0,2\pi)$ 

και με αντίστροφο μετασχηματισμό τον:

$$x_n = \frac{1}{N} \int_{\omega = -\infty}^{+\infty} e^{j\omega n} X(\omega)$$

Οπότε, για ένα «Περιοδικό» σήμα, με πλήθος Ν σημείων σε μία περίοδο Τ, θα έχουμε για τον Διακριτό Μετασχηματισμό Διακριτού χρόνου την έκφραση:

$$X(\omega) = X(e^{j\omega}) = \sum_{n=0}^{N-1} e^{-j\omega n} x_n$$

και ως αντίστροφο μετασχηματισμό τον:

$$x_{n} = \frac{1}{N} \sum_{k=0}^{N-1} e^{j\omega k} X(e^{j\omega}) = \frac{1}{N} \sum_{k=0}^{N-1} e^{j\omega n} X(\omega)$$

Οι εκφράσεις αυτές είναι ίδιες με αυτές για τον Διακριτό Μετασχηματισμό Fouriér αν ισχύει ότι:

$$\omega = \omega_k = \left(\frac{2\pi}{N}\right)k \quad \mu \in k=0,1,2,\dots,N-1.$$

Συμπερασματικά, για ένα πεπερασμένο χρονικά σήμα με N δείγματα ή για ένα όμοιο αλλά περιοδικό με δείγματα N σημεία εντός της περιόδου του, οι φασματικές γραμμές είναι ένα N το πλήθος σημεία στον άξονα των συχνοτήτων και διαδοχικά ισαπέχουν κατά (2π/N).

#### Ο μετασχηματισμός FFT

Ο «Ταχύς μετασχηματισμός Φουριέ» (FFT) (Proakis, Manolakis & Dimitri 1996; Ramirez 1984; Rao & Hwang 2010), είναι ένας αλγόριθμος που συνδυάζει επιμέρους ανά δύο τους μετασχηματισμούς DFT. Δηλαδή, για να υπολογιστεί ο DFT για N σημεία (έστω N =2r δηλαδή άρτιος). Ο FFT (με αποδεκατισμό στον χρόνο) επιμερίζει τον DFT σε log<sub>2</sub>N στάδια, το καθένα από τα οποία αποτελείται από N/2 υπολογισμούς «πεταλούδας». Κάθε «πεταλούδα» έχει δύο μιγαδικούς αριθμούς p και q ως είσοδο και υπολογίζει από αυτούς δύο άλλους αριθμούς p+aq και p-aq, όπου το a είναι ένας μιγαδικός αριθμός.

#### Σχηματικά,



### Σχήμα 1.2: Σχήμα «πεταλούδας» του DFT

#### Μετατροπή αναλογικού σήματος σε ψηφιακό

Για τη λήψη των πεπερασμένων τιμών x<sub>0</sub>, x<sub>1</sub>, ,,, x<sub>n</sub>, σε ορισμένες χρονικές στιγμές t<sub>0</sub>, t<sub>1</sub>, , ,t<sub>n</sub> του αναλογικού σήματος, χρησιμοποιείται η «Δειγματοληψία». Κατά τη διαδικασία αυτή, μία σειρά παλμών συχνότητας fs και επαρκούς ύψους πολλαπλασιάζεται με το αναλογικό σήμα, δίνοντας μία σειρά από στάθμες V<sub>1</sub>,V<sub>2</sub>,,,V<sub>n</sub> στις χρονικές στιγμές t<sub>1</sub>,t<sub>2</sub>,,,t<sub>n</sub>. Η διάρκεια των παλμών Τρ είναι πολύ μικρή σε σχέση με την περίοδό τους Ts. Για να υπάρχει ευκρινές δείγμα και να μην υπάρχει αλληλοκάλυψη των δειγμάτων, θα πρέπει η συχνότητα δειγματοληψίας fs να πληροί τη σχέση:

### $fs \ge 2f_{max}$ (Κριτήριο Nyquist)

Τα δείγματα V1,V2,,,,Vn στη συνέχεια μπορούν να χρησιμοποιηθούν σε διάφορες τεχνικές για μετατροπή σε ψηφιακά δεδομένα. Μία τέτοια γνωστή διεθνώς τεχνική είναι η PCM (Pulse Code Modulation). (B. M. Oliver, Pierce & Shannon 1948). Η τεχνική αυτή περιλαμβάνει, πέρα από τη δειγματοληψία, τα στάδια της «Κβάντισης» και της «Ψηφιοποίησης». Στο στάδιο της «Κβάντισης» οι τιμές V1,V2,,,,Vn εντάσσονται σε περιοχές τιμών (στάθμες) που καθορίζονται σε πλήθος από το σχεδιασμό των πόσων bits χρειάζονται για την αναπαράσταση των τιμών-σταθμών. Έτσι, πχ για κατάταξη των δειγμάτων σε μία από τις 256 στάθμες γίνεται σχεδιασμός για 8 bit αναπαράσταση των τιμών. Έτσι οι στάθμες θα έχουν Vmin=0 και

Vmax=255. Αν οι στάθμες διευθετούνται ομοιόμορφα, τότε η απόσταση ΔV μεταξύ των σταθμών θα είναι

$$\Delta V = \frac{V_{max} - V_{min}}{(N-1)}$$

Όπου το N αναπαριστάνει το πλήθος των σταθμών που, για το παράδειγμα των 8 bit που αναφέρθηκε, θα είναι

$$N = 256 = 2^8$$

Στο στάδιο της «Ψηφιοποίησης» η κάθε στάθμη Q1,Q2,,,Qn, που προκύπτει ως τιμή από τα δείγματα V1,V2,,,Vn, μετατρέπεται στην αντίστοιχη σειρά από bits, δηλαδή σε ακολουθίες 0..1 ανάλογα με την τιμή. (Champeney 1973; Grant, et al. 2002). Παρακάτω παρατίθεται σχηματικά οι διαδικασίες της «Δειγματοληψίας» και της «Κβάντισης» ενός αναλογικού σήματος:



Εικόνα 1.1: Δειγματοληψία και Κβάντιση σημάτων

### 1.3 Οι Συντελεστές MFCC

### 1.3.1 Γενικά στοιχεία για τους συντελεστές MFCC

Το φάσμα συχνοτήτων ενός σήματος x(t), μπορεί να προκύψει όπως αναφέρθηκε με μετασχηματισμό Fourier του x(t), και να προκύψει έτσι η κατανομή X(f). Το ακουστικό σήμα πχ από την ανθρώπινη ομιλία έχει διάφορες συχνότητες και πλάτη, με κατανομή που εξαρτάται, για ορισμένα χαρακτηριστικά, από την πηγή του ήχου έτσι ώστε η ίδια πηγή, σε διάφορους ήχους να παρουσιάζει στη φασματική κατανομή X(f) σταθερά και ανιχνεύσιμα γνωρίσματα. Μια ανάλυση, με την οποία μπορούν να εξαχθούν τέτοια σχετικά γνωρίσματα, είναι η μεθοδολογία με την οποία, για ένα δείγμα ήχου, υπολογίζεται μια σειρά από συντελεστές (ένας πίνακας συντελεστών) που με κατάλληλους μετασχηματισμούς μπορεί να οδηγήσει σε επανακατασκευή του αρχικού σήματος x(t) με μεγάλο βαθμό ακρίβειας. Οι συντελεστές αυτοί λέγονται mel cepstral coefficients (σε ανάστροφη σειρά των γραμμάτων για το spec της λέξης spectrum – φάσμα).(Davis & Mermeilstein 1980).

Η κλίμακα mel είναι μια «ακουστική» κλίμακα υποδεικνύοντας την ευαισθησία της ακοής στις συχνότητες της φωνής, δηλαδή από 0-16000 Hz. Το ανθρώπινο αυτί αντιλαμβάνεται σχεδόν γραμμικά τις συχνότητες από 100-1000 Hz, ενώ η ευαισθησία μειώνεται μη γραμμικά για μεγαλύτερες από 1000 Hz συχνότητες. Έτσι, η κλίμακα mel μπορεί να παρασταθεί από μία συνάρτηση, όπως παρακάτω:

Η συχνότητα mel = 2595 \* 
$$\log(1 + \frac{f(Hz)}{700})$$

Οι συντελεστές MFCC χρησιμοποιούνται τα τελευταία χρόνια ευρέως σε περιπτώσεις αυτοματοποιημένης αναγνώρισης ομιλίας (Automatic Speech Recognition), (Brown 1974; Morgan, Bourlard & Hermansky 2004), μιας και τα χαρακτηριστικά της ομιλίας που αναδεικνύονται από τις μεθοδολογίες με MFCC είναι απαλλαγμένα από επιδράσεις του θορύβου υποβάθρου ή συναισθηματικών αποχρώσεων στην ομιλία, αλλά σε αντίθεση, καταγράφονται τα βασικά στοιχεία που έχουν σχέση με τη φωνητική της λέξης που ακούγεται όπως τα σημεία-συχνότητες με οξεία ακουστική (pitch) ή και οι βασικές συχνότητες που οι εντάσεις της ομιλίας μεγιστοποιούνται (formants).(Rabiner & Juang 1993). Η συνάρτηση των συχνοτήτων της κλίμακας mel είναι, σε γραφική απεικόνιση, όπως παρακάτω:



Εικόνα 1.2 Γραφική απεικόνιση της κλίμακας mel

Οι συντελεστές MFCC εξάγονται με τη χρήση μιας ομάδας φίλτρων (mel filterbank), με τα οποία προβάλλεται η φασματική κατανομή |X(f)|, για κάθε πλαίσιο ή στιγμιότυπο του σήματος x(t) (. Η συνάρτηση των φίλτρων διέλευσης h(m,k) μπορεί να περιγραφεί όπως:

$$h(m,k) = \begin{cases} 0 \text{ av } k < f(m-1) \\ \frac{k - f(m-1)}{f(m) - f(m-1)} \text{ av } f(m-1) \le k < f(m) \\ \frac{f(m+1) - k}{f(m+1) - f(m)} \text{ av } f(m) \le k < f(m+1) \\ 0 \text{ av } k \ge f(m+1) \end{cases}$$

Η γραφική απεικόνιση των φίλτρων h(m,k) δείχνεται παρακάτω, για την περιοχή συχνοτήτων με αντιστοιχία για το k που αναφέρεται στην τελευταία σχέση.



Εικόνα 1.3: Ομάδα φίλτρων συχνοτήτων mel

Οι συντελεστές Cn του φάσματος |X(f)| αποκτιούνται με μετασχηματισμό DCT II, στα δεδομένα που προκύπτουν από τη σχέση:

$$E(m) = \log(|X(f)| * M_H(m))$$

και από εκεί, οι συντελεστές Cn είναι:

$$Cn = \sum_{k=1}^{M} \log(E(k)) e^{\frac{j2\pi kn}{N}} \quad \text{ó}\pi ov \ n = 0, 1, 2 \dots M - 1$$

### 1.3.2 Μέθοδος υπολογισμού των συντελεστών MFCC

Έχοντας κάποιος στη διάθεσή του ένα ηχητικό απόσπασμα, πχ ένα αρχείο WAV μικρής διάρκειας (~ 1-3 sec), μπορεί να υπολογίσει τους συντελεστές MFCC, σύμφωνα με την παρακάτω μεθοδολογία (Young 1993; Ashish 2004), όπως δείχνεται και στο παρακάτω διάγραμμα:



Εικόνα 1.4: Διάγραμμα υπολογισμού των συντελεστών MFCC

Έτσι, τα στάδια υπολογισμού των συντελεστών MFCC μπορούν να περιγραφούν όπως παρακάτω:

### 1º Στάδιο: Προέμφαση

Στο στάδιο αυτό το εισερχόμενο σήμα x(t) μπορεί να υποστεί μια προέμφαση, σύμφωνα με τη σχέση:

$$y(t) = x(t) - ax(t-1)$$

όπου το α είναι συνήθως α=0.95. Κατά την προέμφαση δημιουργείται μία «λείανση» (smoothing out) του εισερχόμενου σήματος.

#### 2º Στάδιο: Πλαίσια

Στο στάδιο αυτό, ο χρόνος ως διάρκεια του σήματος x(t) διαιρείται σε χρονικά παράθυρα με επικάλυψη μεταξύ τους, διάρκειας συνήθως 25-30 ms και με 5ms επικάλυψη, ώστε κάθε πλαίσιο (frame) να περιέχει ένα μέρος του συνολικού χρόνου του σήματος.

### 3º Στάδιο: Παράθυρο

Στο σήμα που λαμβάνεται από ένα πλαίσιο, μπορεί να χρησιμοποιηθεί επιπλέον ένα παράθυρο διέλευσης, για καλύτερη εξομάλυνση του φάσματος που θα προκύψει. Τέτοια παράθυρα είναι συνήθως τα παράθυρα Hamming και Hanning. Το σήμα y(n) που προκύπτει από το παράθυρο διέλευσης W(n), εύρους N σημείων, για τα χρονικά σημεία t<sub>1</sub>, t<sub>2</sub>,,, t<sub>n</sub> του σήματος x(n) θα είναι:

y(n) = W(n)x(n) με n = 0, 1, 2, ..., N - 1 γιά τα χρονικά σημεία

Για παράδειγμα, το παράθυρο Hanning έχει τη διέλευση

W(n) = 0.5 \* 
$$(1 - \cos\left(\frac{2\pi n}{N-1}\right))$$

### 4º Στάδιο: Φιλτράρισμα με ομάδα φίλτρων mel.

Η ομάδα φίλτρων mel, όπως περιγράφηκε στο προηγούμενο εδάφιο, (1.3.1), είναι ένα σύνολο από φίλτρα διέλευσης που ορίζονται με βάση την «ακουστική» συχνότητα κλίμακας Mel, που προκύπτει με τη συνάρτηση:

$$F(mel) = 2595 * \log(1 + \frac{f}{700})$$

Το σήμα που προκύπτει, με φιλτράρισμα των φίλτρων πάνω στο «φάσμα» |Y(f)|, θα είναι,

$$S(f) = |Y(f)|M(fmel)$$

και έτσι, για κάθε πλαίσιο, υπολογίζονται οι «ενεργειακοί» όροι,

$$E(m) = \sum_{k=1}^{N-1} |Y(k)| * M(m, k)$$

Οι όροι E(m), στη συνέχεια χρησιμοποιούνται στο επόμενο στάδιο, για την εξαγωγή των συντελεστών MFCC. Το πλήθος των φίλτρων τύπου Mel, μπορεί να είναι 10-40 συνήθως, ενώ οι σημαντικοί και καθοριστικοί συντελεστές MFCC είναι οι πρώτοι 10-15, από το σύνολο των 10-40.

#### 5° Στάδιο: Συντελεστές MFCC

Τα στοιχεία E(m), για m=0,1,2,,M-1, όπου M είναι το πλήθος της ομάδας φίλτρων τύπου Mel που επιλέγεται, χρησιμοποιούνται στον μετασχηματισμό DCT II, για την ανάκτηση των συντελεστών MFCC, ως ένα διάνυσμα γραμμή (ή στήλη) με M στοιχεία, για κάθε επιλεγόμενο χρονικό πλαίσιο του σήματος x(t). Στον μετασχηματισμό DCT II χρησιμοποιείται η λογαριθμική εξάρτηση των E(m), δηλαδή το log(E(m)), ως στοιχείο για μετατροπή σε «φάσμα» συχνοτήτων που λέγονται qufrencies. Ουσιαστικά, η μετατροπή DCT παρέχει τους συντελεστές MFCC, σύμφωνα και με τη σχέση (Anusuya & Katti 2011),

$$Cn = \sum_{m=1}^{M} \log(E(m)) \cos(n\left(m - \frac{1}{2}\right)\frac{\pi}{M})$$

Επιπλέον, μπορούν να υπολογιστούν και οι συντελεστές delta, που αφορούν τη δυναμική εξέλιξη των Cn συντελεστών από το ένα χρονικό πλαίσιο στο άλλο, σύμφωνα με τη σχέση:

$$dt = \frac{\sum_{n=1}^{N} n(C_{t+n} - C_{t-n})}{2\sum_{n=1}^{N} n^2}$$

Όπου το dt είναι ο συντελεστής delta στο πλαίσιο t, υπολογισμένος με βάση τους στατικούς συντελεστές  $C_{t+N}$  έως  $C_{t-N}$ . Συνηθισμένη τιμή για το N είναι N=2. Ακόμη, με παρόμοια σχέση μπορούν να υπολογιστούν οι συντελεστές delta-delta (επιτάχυνση), όπου στη θέση των Cj συντελεστών θα χρησιμοποιούνταν οι υπολογισμένοι d<sub>i</sub> ( delta στο πλαίσιο j), συντελεστές.

### 1.3.3 Αναγνώριση ήχου με βάση τους συντελεστές MFCC

Το «ταίριασμα» ενός αποσπάσματος ήχου με ένα αντίστοιχο καταγεγραμμένο σε μία βάση δεδομένων μπορεί να γίνει με τη βοήθεια των συντελεστών MFCC. (Τ. Ganchev, Fakotakis & Kokkinakis 2005; Xiong 2011). Το σήμα X(t), που εξετάζεται για «ταυτοποίηση», είναι συνήθως σε άλλα χρονικά διαστήματα καταγεγραμμένο και δεν είναι «συντονισμένο» με τα διαθέσιμα καταγεγραμμένα σε βάση δεδομένων σήματα. Παρόλα αυτά ένας επαρκής για τους σκοπούς του «ταιριάσματος» συντονισμός μπορεί να επιτευχθεί με τεχνικές όπως ο αλγόριθμος DWT. Στη συνέχεια υπολογίζονται οι Manhattan αποστάσεις των συντελεστών MFCC του προς εξέταση σήματος x(t), με τους συντελεστές MFCC των καταγεγραμμένων σημάτων rec(t). Η Manhtatta μετρική για δύο «διανύσματα»  $V=(V_1, V_2, V_n)$  και  $R=(R_1, R_2, ..., R_n)$ ορίζεται ως η «απόστασή» τους στο διανυσματικό χώρο  $R^n$ , με βάση τις αποστάσεις ανά ζεύγος συντεταγμένων τους.

Δηλαδή,

$$d_{\mathsf{M}} = \left| \vec{\mathsf{V}} - \ \vec{\mathsf{R}} \right| = \sum_{i=1}^{n} |\mathsf{V}_i - \mathsf{R}_i|$$

Αυτές οι αποστάσεις υπολογίζονται για τα πλαίσια των αποσπασμάτων x(t) και rec(t), και στον ρόλο των  $V_1$ ,  $V_n$  και  $R_1$ ,  $R_n$  να είναι, ανά πλαίσιο, οι MFCC συντελεστές, των συντονισμένων κατά DWT αντίστοιχων πλαισίων. Το απόσπασμα ήχου «ταυτοποιείται» με το καταγεγραμμένο απόσπασμα που, σύμφωνα με τα προηγούμενα αναφερθέντα, δίνει τη μικρότερη Manhattan απόσταση, στους αντίστοιχους συντελεστές MFCC, στα «συντονισμένα» χρονικά, κατά DWT, πλαίσια.

# κεφαλαίο 2. Στοιχεία του υλικού εξοπλισμού

### 2.1 Υλικός Εξοπλισμός

Για την πειραματική διαδικασία χρησιμοποιούνται οι παρακάτω βαθμίδες, με τη σειρά εισόδου και επεξεργασίας του σήματος:

Γεννήτρια AC σήματος, για την παραγωγή αρμονικών συχνοτήτων στην περιοχή 200Hz έως 13KHz με ρυθμιζόμενο πλάτος, Παλμογράφος διπλού καναλιού, για τις πειραματικές μετρήσεις και για την εποπτεία των AC σημάτων, Μικρόφωνο, για τη λήψη του ηχητικού σήματος και σύνδεση στον μικροελεγκτή για τη μετατροπή σε ψηφιακό σήμα. Ο μικροελεγκτής, (AVR XMega256A3U), διαθέτει ADC διαφορικό κανάλι ώστε να μετατρέπει το αναλογικό σήμα με μέγιστη συχνότητα 20Khz σε ψηφιακό, και το οδηγεί σε επεξεργασία για FFT ανάλυση. Για την μετατροπή του αναλογικού σε ψηφιακό σήμα χρησιμοποιούνται οι «διακοπές» (Interrupts) του μικροελεγκτή. Για την απόκτηση του φασματικού περιεχομένου του ηχητικού σήματος χρησιμοποιείται λογισμικό, που αναπτύχθηκε σε γλώσσα C και «τρέχει» στην μνήμη του μC. Το λογισμικό επιτελεί μετασχηματισμό FFT, ενώ για την εποπτεία των μεγεθών χρησιμοποείται αποτύπωση σε οθόνη LCD και σε Η/Υ (pc) με επικοινωνία μέσω σειριακής θύρας COM με πρωτόκολλο UART και κατάλληλου Λογισμικού. Επίσης, τα δεδομένα από το μετασχηματισμό FFT χρησιμοποιούνται περαιτέρω για την απόκτηση των συντελεστών MFCC και του σχετικού ψηφιακού αποτυπώματος του ηχητικού σήματος. Παρακάτω περιγράφονται σχετικά στοιχεία για τις βαθμίδες που αναφέρθηκαν, λόγω του ρόλου τους στην πειραματική διαδικασία που χρησιμοποιήθηκε.

### Μικροελεγκτές

Για την επεξεργασία σημάτων (εισόδου – εξόδου) με κατάλληλο Λογισμικό χρησιμοποιούνται ευρέως διάφορες οικογένειες μικροελεγκτών – μικροεπεξεργαστών (μC – mcu). (Leaver 2010), Οι υπολογιστικές αυτές διατάξεις έχουν τη δυνατότητα να συνδεθούν με διάφορες μονάδες Εισόδου/Εξόδου (I/O Units) και να επεξεργαστούν ψηφιακά δεδομένα, με κατάλληλα «Προγράμματα» (Λογισμικό). Διαθέτουν μνήμη για προσπέλαση και αποθήκευση δεδομένων (Flash memory, SRAM, EEPROM) και οι λειτουργίες τους μπορούν να προγραμματιστούν σε χαμηλό επίπεδο, σε επίπεδο «Καταχωρητών» (Registers). Οι ταχύτητες επεξεργασίας ποικίλουν από μερικά MHz σε μερικά GHz. Οι διαφορές των μικροελεγκτών από τους μικροεπεξεργαστές εντοπίζονται στην αρχιτεκτονική τους, στο ρεπερτόριο εντολών (Instruction Set), στη μνήμη, και στις ταχύτητες λειτουργίας. (Grace 2015). Συνήθως, οι μικροεπεξεργαστές είναι πιο αυξημένων δυνατοτήτων υπολογιστικές διατάξεις από τους μικροελεγκτές. Από την άλλη μεριά, οι μικροελεγκτές είναι πιο αποτελεσματικοί στο να «τρέχουν» ένα συγκεκριμένο πρόγραμμα που το επαναλαμβάνουν σε κύκλους επανάληψης, κάτι που τους εξειδικεύει σε κατηγορίες εφαρμογών που οι λειτουργίες τους προκαθορίζονται και επαναλαμβάνονται για μεγάλο χρονικό διάστημα (εκατοντάδες ή χιλιάδες ώρες συνεχούς λειτουργίας). Επίσης οι μικροελεγκτές έχουν μικρότερη κατανάλωση ισχύος, μικρότερες τάσεις λειτουργίας (3-12 V) και μπορούν να «προγραμματιστούν» πιο άμεσα σε επίπεδο «Καταχωρητών» από ότι οι μικροεπεξεργαστές.

Οι μικροελεγκτές, (Grace 2015), κατά τη λειτουργία τους, παράγουν σχετικά μικρό Ηλεκτρομαγνητικό Θόρυβο (ΕΜΙ) αφού λειτουργούν με βασικές απλές συνδεσμολογίες (καλωδιώσεις), αρκετά απλούστερες σε σγέση με αυτές των μικροεπεξεργαστών. Επίσης, ο λιγότερος θόρυβος ΕΜΙ οφείλεται στο ότι η συχνότητα λειτουργίας (clock) είναι αρκετά μικρότερη (μερικές δεκάδες MHz) σε σχέση με εκείνη των μικροεπεξεργαστών, με αποτέλεσμα τα επαγόμενα ΗΜ πεδία σε αγωγούς να είναι αρκετά μικρότερα. Ακόμη, το σχετικά μικρό πλήθος ηλεκτρικών διασυνδέσεων και οι κοντινές καλωδιώσεις παρέχουν ένα επιπλέον πλεονέκτημα στην εξοικονόμηση χώρου, ειδικά σε περιπτώσεις που πρέπει να τοποθετούνται και άλλες συσκευές. Η αρχιτεκτονική της CPU και της λειτουργίας των μικροελεγκτών είναι ίδια στη βάση με αυτή των μικροεπεξεργαστών. Ενώ όμως οι κοινοί μικροεπεξεργαστές βασίζονται σε αρχιτεκτονική τύπου Neumann, οι μικροελεγκτές χρησιμοποιούν, (σε μεγάλη πλειονότητα των γνωστών διεθνών Firms), αρχιτεκτονική τύπου Harvard, στη οποία η μνήμη συνδέεται με διάφορες αρτηρίες σύνδεσης (buses) με τη μνήμη προγράμματος και τη μνήμη δεδομένων. Για παράδειγμα τέτοια αρχιτεκτονική ακολουθούν οι μικροελεγκτές της σειράς AVR της Atmel co., καθώς και οι μικροελεγκτές τύπου PIC της Microchip co.

#### Οι βασικές δομικές μονάδες των μικροελεγτών

Ένας μικροελεγκτής διαθέτει απαραίτητα ένα ολοκληρωμένο κύκλωμα (IC) για την υλοποίηση της ALU μονάδας (Αριθμητικής και Λογικής), έτσι ώστε να είναι σε θέση να εκτελέσει βασικούς υπολογισμούς (Tarafder, 2017). Τα δεδομένα που παίρνουν μέρος στους υπολογισμούς μπορούν να αποθηκεύονται σε μνήμη που διαθέτει ο μικροελεγκτής ως αναπόσπαστο τμήμα για τις λειτουργίες του. Οι υπολογισμοί, για τους μικροελεγκτές γίνονται σε ένα ελεγχόμενο αλλά επαναλαμβανόμενο κύκλο εντολών (program loop), έτσι ώστε να ανταποκρίνεται στους προσδιοριζόμενους για την εφαρμογή ρόλους του. Το «πρόγραμμα» αυτό, ως σύνολο εντολών που επαναλαμβάνονται, «φορτώνεται» (loaded) στη μνήμη που χαρακτηρίζεται ως SRAM για τους μικροελεγκτές, ή σε μνήμη Flash, ανάλογα με τις απαιτήσεις.( Barnett, Cox & O'Cull 2006). Ο προγραμματισμός της μνήμης FLASH γίνεται με εγγραφή πάνω στο κύκλωμα της ενσωματωμένης εφαρμογής (ISP Programming).( In circuit serial programming). Σε ορισμένους μικροελεγκτές υποστηρίζεται και ο τρόπος εγγραφής με υψηλή τάση (High Voltage Programming). Η κατηγορία αυτών των μικροελεγκτών είναι η σύγγρονη εξέλιξη στην παλιά μέθοδο διαγραφής της μνήμης (συνήθως EPROM) με τη χρήση υπεριώδους ακτινοβολίας.

Επίσης, τα δεδομένα μπορούν μόνιμα ή προσωρινά να αποθηκεύονται στη μνήμη ΕΕΡROM που διαθέτουν οι μικροελεγκτές (μόνιμη επανεγγράψιμη μνήμη). Τέτοιες μνήμες ποικίλλουν σε μέγεθος ανάλογα με την «οικογένεια» και τη διαβάθμισή τους, την Εταιρία-Κατασκευαστή, και επίσης και τηη εποχή που διατίθενται στο κοινό για χρήση. Έτσι, οι μνήμες SRAM κυμαίνονται από 8K και άνω, με δυνατότητα εξωτερικών επεκτάσεων. Οι μνήμες Flash επίσης κυμαίνονται από 4K και άνω, ενώ σε κάθε περίπτωση, οι ΕΕΡROM είναι αρκετά μικρότερες (μέχρι μερικές δεκάδες των 256 bytes). Το πρόγραμμα που «τρέχει» ο μC φορτώνεται με ειδικές διατάξεις (bootloaders) έτσι ώστε να ελέγχεται η διαδικασία φόρτωσης και αποθήκευσης του προγράμματος. Η «φόρτωση» μπορεί να γίνεται μέσω των «αναπτυξιακών» των μC (όπως στην περίπτωση του STK600 για τους μικροελεγκτές AVR της Atmel), και με διασύνδεση με σειριακή USB ή COM θύρα (Tarafder 2017). Σε ορισμένα κατηγορίες των μC, είναι δυνατό να «κλειδωθεί» αυτή μνήμη μετά την εγγραφή της και να προστατεύεται έτσι ότι έχει γραφεί από διάφορες απόπειρες υποκλοπής-αντιγραφής του Λογισμικού που εγγράφηκε. (Gadre 2009).

### Άλλες λειτουργίες των μικροελεγκτών

Εκτός από τις βασικές τους λειτουργίες, οι μικροελεγκτές, διαθέτοντας δυνατότητες hardware, μπορούν να υποστηρίζουν λειτουργίες τύπου DSP, λειτουργίες επικοινωνιών με Ιντερνετ, λειτουργίες μετατροπών από αναλογικά σε ψηφιακά (ADC), και αντίστροφα, από ψηφιακό σε αναλογικό σήμα (DAC). (Grace 2015). Ακόμη, με τη χρήση περιφερειακών μονάδων όπως οθόνες, πληκτρολόγια, κλπ, μπορούν να εφαρμοστούν σε πληθώρα εφαρμογών, με δυνατότητες μεταφοράς δεδομένων τύπων όπως I2C, SPI, UART μεταξύ του μικροελεγκτή και της περιφερειακής συσκευής.

### Ο Προγραμματισμός της μνήμης FLASH

Όπως αναφέρθηκε, η μνήμη FLASH μπορεί να επαναπρογραμμαστιστεί πάνων στο ίδιο κύκλωμα της εφαρμογής (ISP) ή με χρήση συνδέσεων για παράλληλο υψηλής τάσης προγραμματισμό (High Voltage Programming). Αυτό μπορεί να γίνει με χρήση Η/Υ και σύνδεση του προγραμματιστή του μικροελεγκτή με μία σειριακή ή USB θύρα. Επίσης απαιτείται η χρήση κατάλληλου λογισμικού. (Barnett, Cox & O'Cull 2006). Στον μικροελεγκτή είναι εγκατεστημένο ένα στοιχειώδες λογισμικό που επιτρέπει τα παραπάνω και λέγεται λογισμικό υποδοχής (bootstrap). Μία άλλη περίπτωση προγραμματισμού χωρίς τη χρήση bootstrap είναι η μέθοδος JTAG. Σε τέτοιες περιπτώσεις απαιτείται έλεγχος του λογισμικού πριν τη φόρτωση για αποφυγή δημιουργίας προβλημάτων στη λειτουργία του μικροελεγκτή.

### Γλώσσες προγραμματισμού και εργαλεία ανάπτυξης λογισμικού

Κάθε μορφής λογισμικό είτε αναπτύσσεται σε γλώσσα μηχανής είτε σε ανώτερη γλώσσα, θα πρέπει να διαθέτει την δυνατότητα προγραμματισμού της εσωτερικής μνήμης, και τη δυνατότητα αποσφαλμάτωσης. Συνήθως οι κατασκευαστές των μικροελεγκτών δίνουν και τα εργαλεία λογισμικού που είναι απαραίτητα για τον προγραμματισμό του μικροελεγκτή συνήθως από ένα PC καθώς και διάφορους μεταγλωττιστές για αποσφαλμάτωση του λογισμικού. (Barr & Massa 2006). Το τελικό αρχείο που φορτώνεται στον μικροελγκτή είναι σε δεκαεξαδική μορφή (.hex αρχείο). Το αρχείο αυτό μπορεί να δημιουργηθεί από σύγχρονους μεταγλωττιστές που υποστηρίζουν τη γλώσσα C ή και C++. Για τη χρήση τέτοιων εργαλείων διατίθενται μεταγλωττιστές με περιβάλλον ανάπτυξης όπως είναι το MPLAB, το AVR Studio, το CodeVision κλπ.

### Ο μικροελεγκτής AVR (ATXMega256A3U)

Στην υλοποίηση της παρούσας εργασίας έγινε χρήση του μικροελεγκτή XMega256A3U της εταιρίας ATMEL. (Πηγές Ιντερνετ[1]). Ο μικροελεγκτής αυτός ανήκει στην οικογένεια των 8 bit AVR μικροελεκτών της ATMEL Corp., και έχει επαυξημένες δυνατότητες σε σχέση με άλλους 8 bit μικροελεγκτές. Για τους σκοπούς της εργασίας έγινε χρήση του αναπτυξιακού STK600, το οποίο έχει τη δυνατότητα να χρησιμοποιεί τον μC με προκαθορισμένες ψηφιακές εισόδους / εξόδους (Digital Ports) και να μπορεί να ελέγχεται έτσι άμεσα η εκτέλεση ενός προγράμματος από τον μC. Ο προγραμματισμός του μC γίνεται στη γλώσσα C μέσω του AVR Studio 7, και το «φόρτωμα» του προγράμματος από το PC γίνεται μέσω του USB interface του αναπτυξιακού STK600. Ο μC XMega256A3U διαθέτει 256KB μνήμης για προγράμματα (Flash Memory), 8KB τμήμα για boot προγράμματος, 16KB SRAM, και 4096 bytes EEPROM, 4 κανάλια για DMA ελεγκτή, 8 κανάλια συστήματος ανίγνευσης γεγονότων (event channels). Επίσης, περιλαμβάνει ένα διπλό ADC σύστημα των 12bit, δύο 12 bit DAC κανάλια, 4 αναλογικούς συγκριτές, και ενσωματωμένες γραμμές για κρυπτογράφηση AES και DES. Ακόμη, διαθέτει 7 χρονιστές / απαριθμητές, 6 USART, 2 SPI, και 2 TWI και RTC σύστημα backup μπαταρίας. (Huang 2013). Η συσκευή μπορεί να χρησιμοποιηθεί σε ποικιλία κατασκευών: σε οικοδομικές κατασκευές, στη βιομηχανία, σε εφαρμογές μοτέρ, αυτοματισμούς αυτοκινήτων, οικιακές συσκευές, ως μετρητές, σε δίκτυα Η/Υ, καθώς και σε οπτικές και Ιατρικές εφαρμογές. Ο XMega256A3U ανήκει στην κατηγορία των ATxmega με 64 pin στο chip package. Η οικογένεια των AVR XMega έγει μC χαμηλής ισχύος, υψηλής απόδοσης, και με πλήθος δυνατοτήτων περιφερειακών συνδέσεων, με βάση πάντα της ενισχυμένη αρχιτεκτονική RISC. Με την εκτέλεση εντολών σε ένα απλό κύκλο ρολογιού, οι XMega μC πετυχαίνουν αποδόσεις σε χρήση της CPU της τάξης του ενός εκατομμυρίου εντολές ανά δευτερόλεπτο (MIPS) ανά MHz, επιτρέποντας έτσι την βελτιστοποίηση της απόδοσης έναντι της κατανάλωσης ισχύος. (Grace 2015). Συνδυάζουν επίσης ένα πλήθος από «οδηγίες» (Instruction Set) διαθέτοντας 32 καταχωρητές γενικού σκοπού (General Purpose Registers). Οι καταχωρητές αυτοί είναι απευθείας συνδεδεμένοι με την Αριθμητική και Λογική Μονάδα (ALU) έτσι ώστε να μπορούν να προσπελαστούν ανεξάρτητα δύο καταχωρητές με μία μόνο οδηγία, σε ένα κύκλο ρολογιού. Έτσι, με τέτοια
αρχιτεκτονική, η απόδοση σε υπολογισμούς είναι κατά πολύ καλύτερη από τις αντίστοιχες των απλού συσσωρευτή (Single Accumulator) ή CISC βασισμένων μικροελεγκτών.



#### Εικόνα 2.1: Ο XMega256A3U της ATMEL (Πηγές WWW-)

#### Η Αρχικτεκτονική της CPU του XMega256A3

Στην εκτέλεση των εντολών ο AVR χρησιμοποιεί μία ειδικά σχεδιασμένη παραλλαγή της αρχιτεκτονικής Harvard – με ξεχωριστή μνήμη και buses για το πρόγραμμα και τα δεδομένα. Ο AVR χρησιμοποιεί pipeline ενός σταδίου. Ενώ μία εντολή εκτελείται η επόμενη ανακαλείται από την μνήμη προγράμματος. Αυτό οδηγεί στην εκτέλεση των εντολών σε ένα κύκλο μηχανής (οι περισσότερες από αυτές). Η μνήμη του προγράμματος είναι In-System Reprogrammable Flash. Το Register File περιέχει 32 x 8 bit γενικής χρήσης καταχωρητές με χρόνο πρόσβασης ένα κύκλο ρολογιού. Οι καταχωρητές αυτοί είναι άμεσα συνδεδεμένοι με την αριθμητική λογική μονάδα επιτρέποντας την λειτουργία της ALU σε μία μονάδα χρόνου (ένα κύκλο ρολογιού). Έξι από τους 32 καταχωρητές μπορούν να χρησιμοποιηθούν σαν 3 16–bit έμμεσοι καταχωρητές-δείκτες της μνήμης δεδομένων επιτρέποντας έτσι αποδοτικούς υπολογισμούς διευθύνσεων μνήμης. (Thomas Grace, 2015). Ένας από αυτούς μπορεί να χρησιμοποιηθεί σαν δείκτης σε look-up πίνακες στην μνήμη του προγράμματος. Αυτοί οι καταχωρητές είναι οι X-, Y-, Z-.

#### Σχηματικά,



Εικόνα 2.2: Εσωτερική αρχιτεκτονική του XMega256

Η αριθμητική λογική μονάδα επιτρέπει αριθμητικές και λογικές πράξεις μεταξύ δύο καταχωρητών ή μεταξύ ενός καταχωρητή και μίας σταθεράς. Ακόμα λειτουργίες ενός μόνο καταχωρητή εκτελούνται από τη ALU. (Tarafder 2017). Έπειτα από μία εκτέλεση λειτουργίας από την αριθμητική λογική μονάδα ο status-καταχωρητής ανανεώνεται για να απεικονίζει τα αποτελέσματα της λειτουργίας. Η ροή του προγράμματος διευκολύνεται από με και χωρίς προϋποθέσεις εντολές άλματος και κλήσης που μπορούν άμεσα να μεταφέρουν την εκτέλεση του προγράμματος σε οποιοδήποτε σημείο της μνήμης προγράμματος. Η μνήμη προγράμματος είναι χωρισμένη σε δύο μέρη: την μνήμη εκκίνησης και την μνήμη για τις εφαρμογές. Όλες οι μνήμες στον AVR είναι γραμμικές. Κατά την διάρκεια διακοπών και υπορουτινών η διεύθυνση επιστροφής αποθηκεύεται στην στοίβα. Η στοίβα

SRAM που διαθέτει ο κάθε επεξεργαστής και την χρησιμοποίηση της για αποθήκευση μεταβλητών. Ο δείκτης της στοίβας πρέπει να αρχικοποιείται κάθε φορά που αρχίζει ένα πρόγραμμα. Ο δείκτης της στοίβας είναι γενικά προσβάσιμος στην περιοχή Ι/Ο του μικροελεγκτή.

Η μνήμη SRAM είναι προσβάσιμη μέσω πέντε τρόπων διευθυνσιοδότησης στην αρχιτεκτονική του AVR. Ακόμα αξιοσημείωτα στην αρχιτεκτονική του AVR είναι το ευέλικτο σύστημα για την διαχείριση των διακοπών και οι 64 διευθύνσεις για περιφερειακές συσκευές.

#### Αριθμητική Λογική Μονάδα (ALU)

Η υψηλής απόδοσης ALU λειτουργεί σε απευθείας σύνδεση με τους 32 γενικής χρήσης καταχωρητές. (Huang 2013). Μέσα σ' ένα κύκλο ρολογιού εκτελούνται οι αριθμητικές πράξεις μεταξύ δύο καταχωρητών ή μεταξύ ενός καταχωρητή και ενός ορίσματος από την μνήμη. Οι λειτουργίες της ALU διακρίνονται σε αριθμητικές, λογικές και λειτουργίες σε bit επίπεδο. Ακόμα υποστηρίζεται δυναμική χρήση πολλαπλασιασμού τόσο σε αριθμούς χωρίς πρόσημο όσο και σε προσημασμένους αριθμούς.

#### Οι καταχωρητές γενικής χρήσης του ΑVR

Οι καταχωρητές του AVR είναι κατασκευασμένοι έτσι ώστε να υποστηρίζουν το ενισχυμένο RISC ρεπερτόριο εντολών. (Grace 2015). Έτσι υποστηρίζονται οι εξής λειτουργίες από τους καταχωρητές:

- 1. Ένα 8-bit όρισμα έξοδος και ένα 8-bit αποτέλεσμα είσοδος
- 2. Δύο 8-bit ορίσματα έξοδος και ένα 8-bit αποτέλεσμα είσοδος
- 3. Δύο 8-bit ορίσματα έξοδος και ένα 16-bit αποτέλεσμα είσοδος
- Ένα 16-bit όρισμα έξοδος και ένα 16-bit αποτέλεσμα είσοδος
- Η διάρθρωση των καταχωρητών, σχηματικά, είναι όπως παρακάτω:

	7	0	Addr.	
	R0		\$00	
	R1		\$01	
	R2		\$02	
			1	
	R13		\$0D	
General	R14		\$0E	
Purpose	R15		\$0F	
Working	R16		\$10	
Registers	R17		\$11	
	R26		\$1A	X-register Low Byte
	R27		\$1B	X-register High Byte
	R28		\$1C	Y-register Low Byte
	R29		\$1D	Y-register High Byte
	R30		\$1E	Z-register Low Byte
	R31		\$1F	Z-register High Byte

#### Εικόνα 2.3: Οι καταχωρητές γενικού σκοπού του AVR (Πηγές WWW-)

Οι περισσότερες από τις εντολές που λειτουργούν στο αρχείο καταχωρητών έχουν άμεση πρόσβαση σε όλους τους καταχωρητές, και οι περισσότερες από αυτές εκτελούνται σε ένα κύκλο. Σε κάθε καταχωρητή ορίζεται επίσης μια διεύθυνση μνήμης δεδομένων, χαρτογραφώντας τους άμεσα στις πρώτες 32 θέσεις του διαστήματος δεδομένων. Αν και όχι φυσικά ως θέσεις SRAM, αυτή η οργάνωση μνήμης παρέχει τη μεγάλη ευελιξία στην πρόσβαση των καταχωρητών, όπως το X -, Y -, και οι καταχωρητές Ζ-δεικτών μπορούν να τεθούν ως δείκτης για να συντάξουν ευρετήριο σε οποιοδήποτε καταχωρητή στο αρχείο.



Εικόνα 2.4: Οι διπλοί καταχωρητές Χ-, Υ-, Ζ- του AVR (Πηγές WWW-)

#### Χρόνοι εκτέλεσης εντολών

Η αναβαθμισμένη αρχιτεκτονική τύπου Harvard, που διαθέτει ο AVR (Leaver 2010), οδηγεί σε ένα μεγάλο παραλληλισμό εκτέλεσης εντολών κάτι που τον κάνει να φθάνει σχεδόν σε απόδοση το 1 MIPS ανά MHz ρολογιού. (Huang 2013). Στα παρακάτω διαγράμματα φαίνεται η εκτέλεση συνεχόμενων εντολών από την ALU καθώς και τα στάδια εκτέλεσης μίας εντολής.







Εικόνα 2.6: Στάδια εκτέλεσης εντολής από την ALU

#### Μνήμες του ΑVR

Τα τμήματα μνήμης του AVR είναι βασισμένα στο μοντέλο αρχιτεκτονικής Harvard στο οποίο τα σημαντικά τμήματα της μνήμης είναι διαχωρισμένα για να επιτυχαίνεται ταχύτερη πρόσβαση σε αυτά και αυξημένη χωρητικότητα. (Grace 2015). Η CPU έχει ξεχωριστό interface πρόσβασης στην μνήμη προγράμματος FLASH, την μνήμη δεδομένων και στην ΕΕΡROM (αν υπάρχει).

#### Η μνήμη προγράμματος FLASH

Η μνήμη προγράμματος FLASH είναι ένα block μνήμης FLASH η οποία αρχίζει από την διεύθυνση \$0000 και το μέγεθος της εξαρτάται από το μοντέλο του AVR. (Ειδικότερα το όνομα του AVR προδίδει το μέγεθος της μνήμης. Ο ATmega32 παραδείγματος χάριν που χρησιμοποιήθηκε έχει μνήμη FLASH 32KB. Παρόμοια ο ATmega128 έχει 128KB FLASH). Η μνήμη FLASH είναι non-volatile (δηλαδή διατηρεί τα δεδομένα της μετά την απομάκρυνση της τροφοδοσίας) και έχει διάρκεια ζωής 10,000 κύκλους εγγραφής διαγραφής. Χρησιμοποιείται για την αποθήκευση του κώδικα του προγράμματος και σταθερών. Η διευθυσιοδότηση της μνήμης είναι της τάξης των 16 bits. Παρόλο που η μνήμη είναι επανεγγράψιμη δεν μπορεί να εγγραφεί μέσω ενός εκτελέσιμου προγράμματος, πρέπει να προγραμματιστεί από εξωτερικά μέσα. Συνεπώς είναι μία read-only μνήμη και εκεί μπορούν να αποθηκευτούν μόνο σταθερές μαζί με τον εκτελέσιμο κώδικα. Οι σταθερές όταν αποθηκεύονται στην μνήμη FLASH αυτόματα προβιβάζονται σε ακεραίους λόγω του μεγέθους της μονάδας αποθήκευσης.

#### Η μνήμη δεδομένων SRAM

Η μνήμη δεδομένων του AVR περιέχει τρεις περιοχές μνήμης ανάγνωσης / εγγραφής. Το χαμηλότερο σε διεύθυνση τμήμα της μνήμης περιέχει τους 32 γενικής χρήσης καταχωρητές ακολουθούμενο από τους 64 καταχωρητές εισόδου εξόδου (το νούμερο εξαρτάται από το μοντέλο του επεξεργαστή) και την εσωτερική SRAM. (Leaver 2010). Οι γενικής χρήσης καταχωρητές χρησιμοποιούνται για την αποθήκευση global μεταβλητών καθώς και άλλων προσωρινών δεδομένων κατά την εκτέλεση του προγράμματος. Οι 64 καταχωρητές εισόδου – εξόδου χρησιμοποιούνται για την για την επικοινωνία με τις συσκευές εισόδου – εξόδου και των άλλων περιφερειακών.

#### Διακοπές του AVR – Reset

Οι διακοπές είναι ουσιαστικά κλήσεις συναρτήσεων από το Hardware. (Huang 2013). Όπως προδίδει και το όνομα τους οι διακοπές διακόπτουν την ροή του κυρίου προγράμματος και τον κάνουν να συνεχίσει την εκτέλεση του από ένα σημείο που είναι η ρουτίνα εξυπηρέτησης της διακοπής. Οι διακοπές είναι χρήσιμες για αυτές τις περιπτώσεις που ο επεξεργαστής πρέπει να ανταποκριθεί αμέσως σ' ένα γεγονός ή σε περιπτώσεις που αντί να περιμένει ένα αργό γεγονός να συμβεί αφήνει την διακοπή να τον ειδοποιήσει. Ο AVR παρέχει διάφορες διακοπών που όλες αυτές μαζί με το ξεγωριστό διάνυσμα διακοπής Reset έγουν ένα ξεγωριστό διάνυσμα στην μνήμη προγράμματος. Όλες οι διακοπές έχουν ξεχωριστά bits ενεργοποίησης τα οποία για να ενεργοποιήσουν μία διακοπή πρέπει να γραφούν με λογικό 1 και ταυτόχρονα να είναι ενεργοποιημένες οι διακοπές στον Status καταχωρητή. (Huang 2013). Στο χαμηλότερο κομμάτι του προγράμματος μνήμης έχει by default ανατεθεί ο χώρος στον οποίο βρίσκονται τα διανύσματα των διακοπών. Η σειρά αυτή καθορίζει και την προτεραιότητα των διακοπών. Όσο πιο μικρή είναι η διεύθυνση της διακοπής τόσο πιο μεγάλη είναι η προτεραιότητα της. Την μεγαλύτερη προτεραιότητα έχει η διακοπή Reset και αμέσως επόμενη είναι η εξωτερική διακοπή Int0. Αυτό δεν σημαίνει ότι μία διακοπή μπορεί να σταματήσει την ρουτίνα εξυπηρέτησης μίας διακοπής (καθώς όταν συμβαίνει μία διακοπή απενεργοποιούνται οι διακοπές) με μικρότερη προτεραιότητα καθώς αυτό δεν μπορεί να γίνει αλλά ότι αν ταυτόχρονα περιμένουν να γίνουν δύο διακοπές πρώτη θα εκτελεστεί η διακοπή με την μεγαλύτερη προτεραιότητα.

Βασικά υπάρχουν δύο είδη διακοπών. Το πρώτο είδος συμβαίνει από γεγονότα που θέτουν την σημαία της διακοπής. Τότε ο επεξεργαστής πηγαίνει στην διεύθυνση που δείχνει το διάνυσμα διακοπής για να εκτελέσει την ρουτίνα εξυπηρέτησης της διακοπής. Έπειτα επιστρέφει στην εκτέλεση του κυρίως προγράμματος και καθαρίζει την σημαία της διακοπής. Το δεύτερο είδος προέρχεται από γεγονότα τα οποία προκαλούν διακοπές όσο είναι ενεργά και δεν έχουν απαραίτητα σημαίες διακοπής. (Leaver 2010).

Ο χρόνος ανταπόκρισης σε μία διακοπή είναι τουλάχιστον τέσσερις κύκλοι ρολογιού. Μετά από τέσσερις κύκλους ρολογιού η ρουτίνα εξυπηρέτησης της διακοπής εκτελείται. Κατά την διάρκεια του τέταρτου κύκλου ο PC

αποθηκεύεται στην στοίβα . Έπειτα εκτελείται ένα jump για την ρουτίνα εξυπηρέτησης της διακοπής που παίρνει τρεις κύκλους. Όταν επιστρέφει το πρόγραμμα από μία ISR ο PC ανακτάται από την στοίβα, ο SP αυξάνεται κατά δύο και ενεργοποιούνται οι διακοπές στον Status καταχωρητή. Αυτό παίρνει τέσσερις κύκλους ρολογιού. Ο πίνακας των διακοπών είναι όπως παρακάτω:

#### Θύρες Εισόδου/Εξόδου (I/O ports)

Κάθε θύρα του AVR έχει για τον έλεγχο της 3 καταχωρητές, της PORTx, PINx και DDRx. (Huang 2013). Ανάλογα με την τιμή του DDRx η πόρτα μπορεί να χρησιμοποιηθεί σαν είσοδος ή έξοδος. Ο PORTx χρησιμοποιείται για την θέση της τιμής της πόρτας και PINx για την ανάγνωση της τιμής της πόρτας.

Port A Data Register – PORTA										
	Bit	7	6	5	4	3	2	1	0	
		PORTA7	PORTAS	PORTAS	PORTA4	PORTA3	PORTA2	PORTA1	PORTAD	PORTA
	Read/Write	RW	RW	R/W	RW	RW	R/W	R/W	RW	
	Initial Value	0	0	0	0	0	0	0	0	
Port A Data Direction Register										
– DDRA	Bit	7	6	5	4	3	2	1	0	
		DDA7	DDA6	DDA5	DDA4	DOA3	DDA2	DDA1	DDA0	DDRA
	Read/Write	RW	RW	R/W	RW	RW	R/W	R/W	RW	
	Initial Value	0	0	0	0	0	0	0	0	
Port A Input Pins Address –										
PINA	Bit	7	6	5	4	3	2	1	0	
		PINA7	PINA6	PINA5	PINA4	PINA3	PINA2	PINA1	PINAO	PINA
	Read/Write	R	R	R	R	R	R	R	R	
	Initial Value	N/A								

#### Εικόνα 2.7: Καταχωρητές θύρας εισόδου - εξόδου του ΑVR

Εκτός από θύρες γενικής χρήσης οι πόρτες του AVR μπορούν να ενεργοποιηθούν για να επιτελέσουν λειτουργίες εισόδου εξόδου άλλων περιφερειακών του μικροελεγκτή.

#### Οι Επιλογές του Ρολογιού

Η μονάδα του ρολογιού στους ATMEL χρονιστές αποτελείται από ένα προκλιμακωτή (prescaler) συνδεδεμένο με πολυπλέκτη. Ο προκλιμακωτής μπορεί να θεωρηθεί σαν διαιρέτης ρολογιού.

Σχηματικά,



Επίσης μπορούμε να έχουμε συνδεδεμένο εξωτερικό pin για τον χρονιστή. Οι χρονιστές TIMER0 και TIMER1 χρησιμοποιούν το ρολόι της MCU για τον βασικό χρονισμό. Σε περιπτώσεις που χρειαζόμαστε υψηλή ακρίβεια πρέπει να θέσουμε αρχική τιμή 0 στους μετρητές κτύπων των χρονιστών σε κατάλληλες στιγμές. Ο προκλιμακωτής σχηματικά:



Σχήμα 2.2: Ο προκλιμακωτής (prescaler) στους μικροελεγκτές Atmel

#### Χρονισμός από το ρολόι του συστήματος (της MCU)

Ο χρονιστής TIMER0 είναι συγχρονισμένος με το ρολόι της MCU. Η υπερχείλιση του χρονιστή μπορεί να γίνει ανεξάρτητα από την εκτέλεση του προγράμματος. Λόγω της μεγάλης συχνότητας λειτουργίας (<=16 MHz) μπορούμε να παρακολουθήσουμε παλμούς μικρής διάρκειας με σύνδεση της εισερχόμενης τάσης σε ένα pin.

#### Παράδειγμα ρύθμισης χρονισμού με προκλιμακωτή.

Ας υποθέσουμε ότι έχουμε ως Fck τη συχνότητα λειτουργίας της MCU. Για χρονιστή με ανάλυση 16 bit η μέγιστη τιμή είναι MaxVal=16535. Η συχνότητα υπερχείλισης είναι

```
Fovfl=Fck/(MaxVal*Presc)
```

Για τιμή του πρκλιμακωτή 1024

έχουμε

Για Fck=3690000Hz=3.69MHz είναι

Fovfl=3.69/(255\*1024)=225/16 υπερχειλίσεις /sec = ~14υπερχειλίσεις/sec

Δηλαδή, μία υπερχείλιση / 71 msec. (71msec=1/14 sec)

#### Εξωτερικός χρονισμός

Ο εξωτερικός χρονισμός υποστηρίζεται για τους χρονιστές Timer0 και Timer1. Επιτρέπεται έτσι η χρήση μεγάλου εύρους συχνοτήτων εξωτερικού ρολογιού. Αυτό λέγεται συγχρονισμένος χρονισμός, που σημαίνει ότι και η MCU συγχρονίζεται στον εξωτερικό ρυθμό. Η μέγιστη εξωτερική συχνότητα είναι (Poλόι συστήματος/2) (System Clock/2). Μπορεί να χρησιμοποιηθεί το ανερχόμενο ή κατερχόμενο του παλμού στους T0/T1 για να δειχτεί ένα γεγονός.

#### Χρονιστές του ΑVR

Ο XMega έχει 4 χρονιστές / μετρητές των 16 bit για την επιτέλεση διάφορων διεργασιών. (Huang 2013). Οι timers είναι πιθανότατα το πιο ευρέως χρησιμοποιημένο περιφερειακό σε ένα μικροελεγκτή. Αυτό συμβαίνει λόγω των πολλών εφαρμογών που μπορούν να επιτελέσουν. Μέσω ενός timer μπορούν να

μετρηθούν περίοδοι χρόνου, να καθοριστεί το εύρος παλμών, να μετρηθεί η ταχύτητα και η συχνότητα ή να δημιουργηθούν διαφόρων ειδών σήματα.

Παρόλο που υπάρχουν δύο διαχωρισμένες μορφές, χρονομέτρηση και μέτρηση, οι timers είναι ουσιαστικά απλοί δυαδικοί μετρητές προς τα πάνω. Όταν χρησιμοποιούνται σε timing mode, οι δυαδικοί μετρητές μετρούν περιόδους ρολογιού που εφαρμόζονται στην είσοδο τους ενώ στην timing mode μετρούν παλμούς στην είσοδο τους. Οι timers μπορούν να χρησιμοποιήσουν ως είσοδο είτε υποδιαιρέσεις του ρολογιού του μικροελεγκτή είτε εξωτερική πηγή ρολογιού. (Leaver 2010).

#### Διακοπές και χρονιστές

Οι βασικοί καταχωρητές που παίζουν ρόλο στις διακοπές του χρονιστή TIMER0 είναι όπως παρακάτω:



TCN0 Timer Counter 8 bit

OCR0 Timer Output Compare Register

#### Υπερχείλιση

Μία μέτρηση χρόνου μπορεί να συσχετιστεί με τη διακοπή υπερχείλισης. Η υπερχείλιση σημαίνει ότι ο χρονιστής έχει μετρήσει μέχρι την μέγιστη τιμή και ξαναχρίζει από το 0 στον επόμενο χτύπο του ρολογιού εισόδου. Η ανάλυση σε bit δείχνει την μέγιστη τιμή μέτρησης MAXVAL σύμφωνα με τη σχέση MaxVal=2<sup>Res</sup>-1 Π.χ. για Res=8 bit είναι MaxVal=255. Το γεγονός της υπερχείλισης (overflow event) προκαλεί να τεθεί το bit TOVx στον καταχωρητή TIFR.

#### Καταχωρητής ελέγχου του Timer 0 απαριθμητή

Timer/Counter Control Register – TCCR0	BI	7	6	5	4	3	2	1	0	
		FOC0	WGM00	COM01	COM00	WGM01	C\$02	CS01	C\$00	TCCR0
	Read/Write	W	R/W	RW	R/W	R/W	RW	RW	RW	
	Initial Value	0	0	0	0	0	0	0	0	

Καταχωρητής ελέγχου του Timer1 απαριθμητή



Καταχωρητής Μάσκας Διακοπών

Timer/Counter Interrupt Flag Register – TIFR <sup>(1)</sup>	Bit	7	6	5	4	з	2	1	0	
		TOV1	OCF1A	OC1FB	-	ICF1	-	TOVO	OCF0	TIFR
	Read/Write	R/W	R/W	R/W	R/W	R/W	R/W	R/W	R/W	
	Initial Value	0	0	0	0	0	0	0	0	

Καταχωρητής σηματοδοτήσεων διακοπών

Timer/Counter Interrupt Flag Register – TIFR <sup>(1)</sup>	Bit	7	6	5	4	3	2	1	0	
		TOV1	OCF1A	OC1FB	-	ICF1	-	TOVO	OCF0	TIFR
	Read/Write	R/W	R/W	R/W	R/W	R/W	R/W	R/W	R/W	
	Initial Value	0	0	0	0	0	0	0	0	

Τέλος χρονισμού

Μία τιμή 0 στον προκλιμακωτή τερματίζει τον χρονισμό.
Διαμόρφωση καταχωρητών στον ΤΙΜΕR0 για χρονισμό
Αρχικοποίση prescaler στο 1024 TCCR0 |= (1<<CS02) | (1<<CS00)</li>
Καθάρισμα εκκρεμών διακοπών TIFR |= 1<<TOV0</li>
Ενεργοποίση διακοπής χρονιστή TIMSK |=1<<TOIE0</li>
Αρχικοποίηση μετρητή TCNT0=0
Το διάνυσμα εξυπηρέτησης διακοπής TIMER0\_OVF0\_vect *Η ρουτίνα εξυπηρέτησης διακοπής*ISR (TIMER0\_OVF0\_vect)
{
.....code
};
Η αρχικοποίηση των χρονιστών TIMER0 και TIMER1 γίνεται με την TCNT0=0
TCNT1=0

### Η Οθόνη LCD (Liquid Crystal Display)

Το LCD είναι μία διάταξη που μπορεί να απεικονίσει σε γραμμές ή σε γραφική μορφή χαρακτήρες. Το LCD του αυτόνομου συστήματος που περιγράφεται αποτελείται από οθόνη με 2 γραμμές των 16 χαρακτήρων. (Yang & Wu 2014). Μία απεικόνιση του LCD που χρησιμοποιήθηκε φαίνεται παρακάτω:



#### **Εικόνα 2.8: Οθόνη LCD 2X16**

Τα τεχνικά χαρακτηριστικά του LCD περιγράφονται ως εξής:

Διασύνδεση με 4 ή 8 bit mcu των μικροελεγκτών, απεικόνιση σε 16 ή 64 χαρακτήρες (Χ 2 γραμμές, ή Χ4 γραμμές), χρώμα προβολής με διαφορετικό χρώμα για υπόβαθρο,

τρόπος πόλωσης θετικός ή αρνητικός, μνήμη ROM με αποθηκευμένους χαρακτήρες (CGROM), καθώς και μνήμη RAM και γεννήτρια χαρακτήρων (CGRAM).

#### **Βασικές λειτουργίες του LCD**

Αρχικά, για τη λειτουργία του LCD πρέπει να γίνει αρχικοποίηση του LCD. Στη συνέχεια επιλέγεται ο καταχωρητής (Καταχωρητής εντολών IR-Instruction Register, ή Καταχωρητής δεομένων DR Data Register) με την τοποθέτηση είτε εντολής (RS bit σε 0) είτε δεδομένων (RS bit σε 1) αντίστοιχα. Για εγγγραφή δεδομένων στο LCD θέτουμε το R/W bit σε 0, ενώ για ανάγνωση θέτουμε το R/W bit σε 1. (Deng-Ke Yang & Wu 2014).

Το LCD έχει μνήμη ROM που λέγεται CGRAM και στην οποία είναι καταχωρημένοι οι χαρακτήρες που μπορούν να αποτυπωθούν στην οθόνη. Υπάρχουν δύο μορφές, οι χαρακτήρες των 5x7 pixel και των 5x10 pixel. Η αρχικοποίηση του LCD γίνεται είτε με εσωτερικό Reset είτε στέλνοντας εντολές από την MCU. (Cristaldi, Pennisi & Pulvirenti 2009). Για την αποστολή χαρακτήρα στο LCD για αποτύπωση στην οθόνη, με κατάλληλο κώδικα από τον μικροελεγκτή, στέλνουμε ένα χαρακτήρα, αρχικά επιλέγουμε τον καταχωρητή δεδομένων του LCD με το RS->1, και εγγραφή στο LCD με το R/W σε 0 και το Enable σε 1. Τον χαρακτήρα τον έχουμε δηλώσει ως uint8\_t δηλαδή ακέραιο του ενός Byte. Έτσι στέλνουμε τον ASCII κωδικό του, που για τους λατινικούς χαρακτήρες είναι ίδιος με τον κωδικό που έχει η CGRAM. Για παράδειγμα, το P από το "PIESTE… που γράφουμε στο LCD έχει κωδικό ASCII 01010000 (80 στο δεκαδικό) και το LCD αποτυπώνει στη γραμμή που καθορίζουμε με σχετική διεύθυνση, αρχίζοντας από την πρώτη γραμμή, σε ποια θέση θα τυπωθεί ο χαρακτήρας.

#### Το αναπτυξιακό STK600

Το STK600 είναι ένα ένα αναπτυξιακό σύστημα για πολλά μοντέλα μικροελεγκτών που κατασκευάζονται από την εταιρεία Atmel. (Papazoglou 2018). Το σύστημα αυτό διαθέτει δυνατότητα επικοινωνίας με PC μέσω της σειριακής θύρας RS232 και τροφοδοτείται με DC τάση από 8 έως 13 V. Έχει έτοιμες βάσεις υποδοχής για μικροελεγκτές με 8, 20, 28 και 40 ακροδέκτες. Ταυτόχρονα διαθέτει κύκλωμα προγραμματιστή για σειριακό προγραμματισμό (ISP) όπως και για προγραμματισμό με υψηλή τάση (High Voltage Programming). (Mazidi et al. 2010). Επίσης διαθέτει 8 κουμπιά δαικοπτών (switch buttons) και 8 Leds γενικής χρήσης. Διαθέτει εξωτερικές συνδέσεις των 10 pins για τις θύρες Α-Ε των μικροελεγκτών. Διαθέτει επίσης και μια δεύτερη σειριακή θύρα βοηθητική γενικής χρήσης καθώς και 2 Mbits μνήμης δεδομένων Flash.



Εικόνα 2.9: Το αναπτυξιακό STK 600 (Πηγές WWW-)

## Το πρωτόκολλο επικοινωνίας UART μεταξύ μικροελεγκτή και PC

Ο μικροελεγκτής XMega256A3U διαθέτει σειριακή θύρα COM για επικοινωνία με συσκευές μέσω του σειριακού πρωτοκόλλου RS232 (UART για τον μC). Το πρωτόκολλο αυτό υποστηρίζει σειριακή μετάδοση δεδομένων ανά bit, και σε ομάδες των 7 ή 8 bit, με αντίστοιχο bit ελέγχου ορθότητας (parity bit) και bit τερματισμού του συρμού των bits. Είναι ιδανικό για μετάδοση χαρακτήρων με αναπαράσταση σε 7 ή 8 bits, (ASCII), αλλά ο ρυθμός μετάδοσης είναι σχετικά μικρός (baud rate από 1200-119200), σε σχέση με τις εκατοντάδες Mbps που πετυχαίνονται στις σύγχρονες τηλεπικοινωνίες. (Willi 2014). Για την ενεργοποίηση του πρωτοκόλλου χρειάζεται να ρυθμιστούν, σε ότι αφορά τον μC, οι αντίστοιχοι καταχωρητές για την UART μετάδοση και λήψη. (Axelson 2007). Το αναπτυξιακό STK600 παρέχει διασύνδεση του σειριακού σήματος που διασυνδέεται μέσω της θύρας COM με τις «θύρες» (ports), του μC, με ένα ειδικό ζευγάρι pins, τα Tx, και Rx, που βρίσκονται στην πλακέττα του STK600. Η μετάδοση δεδομένων με το σειριακό πρωτόκολλο UART (RS232), (Campbell 1993), μπορεί να αναπαρασταθεί με γραφικό τρόπο όπως παρακάτω: Μορφή μετάδοσης δεδομένων με UART



Σχήμα 2.3: Σειριακή μετάδοση των bits

# ΚΕΦΑΛΑΙΟ 3. ΜΕΘΟΔΟΛΟΓΙΑ (ΥΛΙΚΟ-ΛΟΓΙΣΜΙΚΟ-ΜΕΤΡΗΣΕΙΣ)

## 3.1 Περιγραφή του υλικού εξοπλισμού-Λογισμικού

Για την πειραματική διαδικασία χρησιμοποιούνται οι παρακάτω βαθμίδες, με τη σειρά εισόδου και επεξεργασίας του σήματος:

**Γεννήτρια AC σήματος:** Χρησιμοποιείται για την παραγωγή αρμονικών συχνοτήτων στην περιοχή 200Hz έως 13KHz με ρυθμιζόμενο πλάτος Vpp ~ (1-5) V. Για τους σκοπούς της παρούσας μελέτης χρησιμοποιείται γεννήτρια AC τύπου YB16200 Function Generator με 0,2 -200 KHz. Η γεννήτρια διαθέτει για έξοδο αρμονικό, τριγωνικό και τετραγωνικό σήμα, με ρυθμιζόμενο πλάτος μέχρι 20 V και ρυθμιζόμενο DC offset μέχρι 20V.

Παλμογράφος διπλού καναλιού: Για τις πειραματικές μετρήσεις και για την εποπτεία των AC σημάτων χρησιμοποιείται παλμογράφος με εύρος 0-20MHz, με 2 κανάλια και διευκρίνιση τάσεων στην περιοχή 5 mV-5V /DIV, ενώ η χρονική κλίμακα είναι στην περιοχή από 0.5 sec – 0.2 μsec/DIV. Ο τύπος παλμογράφου είναι της σειράς GW Oscilloscope GOS-622 στα 20 MHz.

**Μικρόφωνο**: Λαμβάνει το ηχητικό σήμα και το οδηγεί στη βαθμίδα ADC του μικροελεγκτή. Ο μικροελεγκτής διαθέτει ADC διαφορικό κανάλι στα 12 bits και γίνεται έτσι χρήση των A1,A2 pins εισόδου, ενώ οι καταχωρητές του ρυθμίζονται ανάλογα. Οι στάθμες του σήματος (1024 στάθμες) περιλαμβάνονται σε εύρος τάσεων 0-1,5 Volt, με ανάλυση στα 1,5 V/1024= 1,46 mV.

**Βαθμίδα ADC στον μικροελεγκτή:** Η βαθμίδα μετατρέπει το αναλογικό σήμα με μέγιστη συχνότητα 20Khz σε ψηφιακό, και το οδηγεί σε επεξεργασία για FFT ανάλυση. Τα δείγματα λαμβάνονται με συχνότητα ρολογιού ADC πάνω από 100 Khz,

**Λογισμικό ADC** : Για την μετατροπή ADC του σήματος, χρησιμοποιούνται οι «διακοπές» (Interrupts) του μικροελεγκτή XMega256A3U. Ρυθμίζεται η ενεργοποίηση των διακοπών να γίνεται με την εκάστοτε χρησιμοποιούμενη συχνότητα δειγματολειψίας για την ADC μετατροπή, fs, η οποία ανατίθεται στην «διεθνώς» χρησιμοποιούμενη τιμή των 44100Hz, όσον αφορά τη δειγματολειψία των

ηχητικών σημάτων. Το κύκλωμα ADC του XMega256A3U ρυθμίζεται για δειγματολειψία στα 12 bit.

**Λογισμικό FFT και MFCC:** Για την απόκτηση του φασματικού περιεχομένου X[f], του ηχητικού σήματος x(t), χρησιμοποιείται λογισμικό, που αναπτύχθηκε σε γλώσσα C και «τρέχει» στην SRAM του μC, με είσοδο για FFT τις τιμές Xn (n=0,1,2,...) που αποκτούνται από την δειγματοληψεία και μετατροπή ADC του σήματος, από την προηγουμένως αναφερθείσα βαθμίδα του μC. Εξετάζονται τρεις διαφορετικές περιπτώσεις για τα αποτελέσματα του FFT, δηλαδή εξετάζονται τρεις διαφορετικές περίπτωση καταγραφής πόματος, του φάσμα |X(f)| που υπολογίζεται είναι για τον προσδιορισμό των κύριων συχνοτήτων του σήματος. Το ψηφιακό αποτύπωμα υπολογίζεται μέσω των συντελεστών MFCC, που προκύπτουν με νέους μετασχηματισμούς στο αποκτηθένα φάσμα συχνοτήτων, όπως περιγράφηκε και στο Kεφάλαιο 1. Ένα μέρος του κώδικα FFT και του υπολογισμού των συντελεστών MFCC παρατίθεται στο Παράρτημα Ι.

Η/Υ (pc): Για την άμεση λήψη και καταχώρηση των αριθμητικών αποτελεσμάτων, χρησιμοποιείται Λογισμικό σειριακής επικοινωνίας του μC με Η/Υ (pc), που αναπτύχθηκε ειδικά για τους σκοπούς της παρούσας μελέτης, σε γλώσσα C# (Visual Studio 2017), παρέχοντας έτσι τη δυνατότητα αμφίδρομης επικοινωνίας του μC με τον Η/Υ μέσω του «σειριακού» πρωτοκόλλου RS232. Το πρωτόκολλο αυτό υποστηρίζεται, από το μέρος του μC, με τη βαθμίδα UART για σειριακή επικοινωνία. Από τη μεριά του p.c. χρησιμοποιείται εικονική θύρα COM με σύνδεση με USB θύρα και λογισμικό USB to serial για διασφάλιση επικοινωνίας με τύπου RS232 com to com σειριακό πρωτόκολλο. Οι παράμετροι της σειριακής επικοινωνίας ρυθμίζονται ως εξής, τόσο στον μC, με τις κατάλληλες εντολές για τους UART καταχωρητές του μC, όσο και μέσω του Λογισμικού για το pc.: Θύρα com: Virtual com Baud rate: 9600

Αρτιότητα (parity): none

Stop bits: one (1)

Flow control: none

Μέρος του κώδικα του «σειριακού» Λογισμικού παρατίθεται στο Παράρτημα Ι.

Παρατίθεται μία σχετική εικόνα από το αναπτυχθέν σε γλώσσα C# Λογισμικό, για τους σκοπούς της αποτύπωσης αποτελεσμάτων στον Η/Υ μέσω σειριακής επικοινωνίας με μC.

(Απεικονιζόμενα ληφθέντα δεδομένα φασματικής ισχύος για ημιτονικό σήμα, με Vpp=1,34V, και με N=512. Ο αριθμός 11 είναι το frequency bin και αντιστοιχεί σε υπολογισμό μιάς f=(11/511)\*44100=950 Hz, για εφαρμοζόμενη συχνότητα 1000 Hz)

	Εισερχόμενα δεδομένα			
ΘΥΡΕΣ- αρχικοποίηση	11 13437 2594	Θύρα:	COM4	
	2595 2563 1922	Baud rate:	9600	
IsClosed	3471 3882 2683	Bits ανά frame:	8	
Μετάδοση αρχείου data	2689 2597 11366	Αρτιότητα:	None	
	25869 18438	Στon bits:	One	
	7565 4817 5729	Έλεγχος Ροής:	None	
Clear	5259 3113 1064 824 2097 3561		ΕΔΟΜΕΝΑ	

Εικόνα 3.1:Λογισμικό για UART επικοινωνία μC- Η/Υ



Εικόνα 3.2: Γεννήτρια συχνοτήτων



Εικόνα 3.3: Παλμογράφος (απεικονιζόμενο σήμα 1KHz)

ΦΩΤΟ ΚΥΚΛΩΜΑΤΟΣ



Εικόνα 3.4: Το πειραματικό setup (αναπτυξιακό STK600, μC, microphone, LCD, Παλμογράφος..)



Σχήμα 3.1: Block διάγραμμα συνδέσεων των συσκευών

## 3.2 Καταγραφές Αρμονικών ήχων

Στο αρχικό στάδιο καταγραφής συχνοτήτων, χρησιμοποείται γεννήτρια ακουστικών συχνοτήτων με τάσεις εξόδου Vpp<=2,4 V για να μην δημιουργηθεί υπερφόρτωση στο διαφορικό ζεύγος A1,A2 pins του μικροελεγκτή, το οποίο μεταφέρει την διαφορά δυναμικού μεταξύ των pins στο εσωτερικό κύκλωμα ADC του μικροελεγκτή. Οι μετρήσεις διεξάγονται για συχνότητες από 0,250 KHz έως 11KHz, με βήμα ανά 150-250 Hz. Σε σχέση με τον μετασχηματισμό FFT, χρησιμοποιούνται για όλες τις συχνότητες N=512 σημεία. Το φάσμα |X(f)| υπολογίζεται για τη συχνότητα fπ, δηλαδή της πηγής σήματος, με βάση το bin m (όπου m=1, 2,..(N/2)) γιά το οποίο η ισχύς |X(f)| είναι μέγιστη. Οπότε η fπ προσδιορίζεται από τη σχέση:

$$f\pi = m\Delta f = m(\frac{fs}{N})$$

όπου fs είναι η συχνότητα δειγματοληψίας (44138Hz), και N το πλήθος των σημείων για τον μετασχηματισμό FFT. Οι μετρήσεις και τα υπολογισθέντα στοιχεία συνοψίζονται στον πίνακα που ακολουθεί:

fπ(H	n(#bi	fm(H	$ \Delta f (H$	$ \Delta f/f $	fπ(H	n(#bi	fm(H	$ \Delta f (H$	$ \Delta f/f $
Z)	n)	Z)	Z)	%	Z)	n)	Z)	Z)	%
400	5	431	31	7,7%	2500	28	2413	87	3,4%
500	6	517	17	3,4%	3000	35	3017	17	0,5%
600	7	603	3	0,5%	3500	40	3448	52	1,4%
700	8	689	11	1,5%	4000	46	3965	35	0,8%
800	9	778	21	2,7%	4500	52	4482	18	0,4%
900	10	862	38	4,2%	5000	58	5000	0	0%
1000	11	948	52	5,2%	5500	63	5431	69	1,2%
1100	12	1034	66	6%	6000	69	5948	52	0,8%
1200	13	1120	80	6,6%	6500	75	6465	35	0,5%
1240	14	1206	34	2,7%	7000	81	6982	18	0,2%
1300	15	1293	7	0,5%	7500	87	7500	0	0
1400	16	1379	21	1,5%	8000	92	7931	69	0,8%
1500	17	1465	35	2,3%	8500	98	8448	52	0,6%
1600	18	1551	49	3%	9000	104	8965	35	0,4%
1700	19	1637	63	3,7%	1000	116	10000	0	0
					0				
1800	20	1724	76	4,2%	1050	121	10431	69	0,6%
					0				
1900	22	1896	4	0,2%	1100	127	10948	52	0,5%
					0				
2000	24	2068		3,4%	1150	134	11551	1	0,008
					0				%

Πίνακας	3.1:	Αρμονικές	συχνότητες
---------	------	-----------	------------

# 3.3 Υπολογισμός συντελεστών MFCC των αρμονικών συχνοτήτων και αποθήκευση σε EEPROM

Για τις αναφερθείσες στο προηγούμενο κεφάλαιο (ΚΕΦ. 3) συχνότητες, δηλαδή των 500,1000,1500,2000 και 2500Hz, υπολογίζονται στον μικροελεγκτή, μετά την εξαγωγή της φασματικής κατανομής, οι συντελεστές MFCC με τα βασικά στάδια που περιγράφονται στο Κεφ1, 1.3. Σε σύντομη αναφορά, ο υπολογισμός και η αποθήκευση των συντελεστών MFCC γίνεται στα εξής στάδια:



#### Σχήμα 3.2: Τα στάδια υπολογισμού των συντελεστών MFCC

**Στάδιο 1**° : Γίνεται καταγραφή των x(t), σε ορισμένες στιγμές t<sub>0</sub>,t<sub>1</sub>,t<sub>2</sub>..t<sub>N-1</sub> με βάση τη συχνότητα δειγματοληψίας των 16 KHz. Η καταγραφή γίνεται μέσω μικροφώνου που συνδέεται κατευθείαν στα ADC pins (A1,A2) της θύρας «A» του μικροελεγκτή. Το μικρόφωνο παράγει αναλογικό ηλεκτρικό σήμα (ηλεκτρική τάση) με πλάτος 0,8 V σε μέγιστη τιμή, όπως αυτό επαληθεύεται με «παρακολούθηση» της παραγόμενης τάσης σε παλμογράφο. Η διάρκεια καταγραφής είναι 1-55 sec, και το σήμα x(t) λαμβάνεται, μέσω των διακοπών (interrupts) του μC, που «δειγματοληπτεί» με ADC, σε

«χρονικά» πλαίσια ή αλλιώς «παράθυρα», διάρκειας 512\* (1/16000)=0,032 sec ή 32 msec.

**Στάδιο 2°** : Στο κάθε χρονικό παράθυρο εφαρμόζεται FFT 512 σημείων και αποκτούνται τα πλάτη X(fn), σε συχνότητες fn=nfo, με fo την συχνότητα fo=fs/N.

Στάδιο 3°: Υπολογίζονται οι «ενεργειακοί» όροι log(E(mel)M(mel)) όπου E(mel) είναι οι ενεργειακοί όροι στις συχνότητες mel frequencies, με βάση την ομάδα φίλτρων (filterbank) τύπου mel frequencies, όπως περιγράφεται αναλυτικά στο Κεφάλαιο 1. Οι «ενεργειακοί» όροι είναι  $E(mel f)=|X(f)|^2$ , όπου X(f) είναι το (μιγαδικό) πλάτος των σημείων συχνότητας που προκύπτει από το μετασχηματισμό FFT του κάθε χρονικού «παραθύρου». Στον υπολογισμό των συντελεστών MFCC, οι «ενεργειακοί» όροι «προβάλλοντα» στα φίλτρα τύπου Mel, και υπολογίζεται, όπως περιγράφεται παρακάτω, 0 μετασχηματισμός DCT για τους όρους log10([X(f)]\*Mel(f)) που είναι ισοδύναμοι, πλην μιάς πολλαπλασιαστικής σταθεράς, με τους όρους:

$$\log_{10}\{|X(f)|^2 * Mel(f)\} = 2 * \log_{10}\{|X(f)| * Mel(f)\}$$

Στάδιο 4° : Υπολογίζεται ο DCT II (Discrete Fourier Trasfmorm) ή ο μετασχηματισμός Φουριέ συνημιτόνου όπως αλλιώς λέγεται ο μετασχηματισμός, για τους «ενεργειακούς» όρους από το προηγούμενο στάδιο. Τα πλάτη Xmfcc του μετασχηματισμού DCT είναι οι νέοι συντελεστές MFCC, (Mel Frequency Cepstral Coefficients), και ακολουθείται στη συνέχεια η αποθήκευσή τους στη μνήμη EEPROM του μικροελεγκτή, σε συνεχόμενες για αυτή την ομάδα συντελεστών, θέσεις μνήμης με μέγεθος 2 bytes /θέση, αφού οι τιμές των MFCC προκύπτει ότι είναι, κατά απόλυτη τιμή, μικρότερες των 32768, αριθμός που «καλύπτεται» με την απαίτηση των 2 bytes, για θετικές ή αρνητικές τιμές.

Χρησιμοποιούνται 32 συνολικά σημεία για 31 συνολικά φίλτρα, από συχνότητες fmin=100Hz, έως fmax=fs/2=8000Hz. Οι συντελεστές MFCC αθροίζονται σε μέσους όρους για όλα τα χρονικά «παράθυρα», δίνοντας έτσι μια τελική ομάδα 30 συντελεστών MFCC για κάθε αρμονική συχνότητα. Οι συντελεστές καταγράφονται, με αρίθμηση 1-29, (χωρίς τον συντελεστή MFFC<sub>0</sub>, μιάς και αυτός φέρει πληροφορία για την DC ενέργεια του σήματος x(t)), οπότε το σύνολο αυτό των 29 συντελεστών

MFCC είναι η «ψηφιακή υπογραφή» της αρμονικής συχνότητας. Διαφορετικά όρια συχνοτήτων δεν επηρεάζουν το τελικό αποτέλεσμα για τους συντελεστές MFCC, εφόσον βέβαια οι καταγραφόμενες συχνότητες δεν είναι εκτός των ορίων των φίλτρων της ομάδας τύπου φίλτρων mel. Δεν εφαρμόζεται προέμφαση στο σήμα μιάς και το σήμα περιέχει μία βασική συχνότητα, οπότε η προέμφαση θα αλλοιώσει την ημιτονική κατανομή ισχύος. Ως παράθυρο του σήματος λαμβάνεται το απλό ορθογώνιο παράθυρο, με W(n)=1, για n=0,1,2..N-1, ενώ ένας αρχικός πειραματισμός με παράθυρα Hamming και Hanning δίνει αποτελέσματα για πολλαπλάσια συχνότητα (αρμονικές) και δεν υιοθετείται τέτοιο είδος παραθύρων. Έτσι, η διαδικασία για τον υπολογισμό των MFCC συντελεστών ακολουθείται επακριβώς για όλα στάδια που έχουν περιγραφεί στο Κεφάλαιο 1.3. Ο πίνακας των εξαγθέντων συντελεστών MFCC, για τις συχνότητες των οποίων έχουν παρατεθεί οι φασματικές κατανομές τους στα προηγούμενα, παρατίθεται να είναι όπως παρακάτω, έχοντας καταγράψει τους πρώτους 30 συντελεστές MFCC και έχοντας «καταχωρήσει» στην EEPROM του μικροελεγκτή τους 1-29 συντελεστές, για κάθε περίπτωση συχνότητας που εξετάζεται.

Ci∖f	500Hz	1000Hz	1500Hz	2000Hz	2500Hz
C0	0,007	0,007	0,006	0,006	0,005
C1	-0,231	-0,216	-0,200	-0,185	-0,169
C2	0,222	0,208	0,193	0,178	0,163
С3	-0,208	-0,194	-0,180	-0,167	-0,153
C4	0,189	0,176	0,164	0,151	0,139
C5	-0,165	-0,154	-0,143	-0,132	-0,121
C6	0,137	0,128	0,119	0,110	0,101
C7	-0,106	-0,099	-0,092	-0,085	-0,078
<b>C8</b>	0,072	0,067	0,062	0,057	0,053
С9	-0,036	-0,034	-0,031	-0,029	-0,026
C10	1,44E- 16	1,34E- 16	1,25E- 16	1,15E- 16	1,05E- 16
C11	0,036	0,034	0,031	0,029	0,026
C12	-0,072	-0,067	-0,062	-0,057	-0,053
C13	0,106	0,099	0,092	0,085	0,078
C14	-0,137	-0,128	-0,119	-0,112	- 0,10103

Πίνακας 3.2: Οι συντελεστές MFCC (C<sub>0</sub> έως C<sub>14</sub>)

C15	0,165	0,154	0,143	0,132	0,121
C16	-0,189	-0,176	-0,164	-0,151	-0,139
C17	0,208	0,194	0,180	0,167	0,153
C18	-0,222	-0,208	-0,193	-0,178	-0,163
C19	0,231	0,216	0,200	0,185	0,169
C20	-0,234	-0,218	-0,203	-0,187	-0,171
C21	0,231	0,216	0,200	0,185	0,169
C22	-0,222	-0,208	-0,193	-0,178	-0,163
C23	0,208	0,194	0,180	0,167	0,153
C24	-0,189	-0,176	-0,164	-0,151	- 0,13905
C25	0,165	0,154	0,143	0,132	0,123
C26	-0,137	-0,128	-0,119	-0,110	-0,101
C27	0,106	0,099	0,092	0,085	0,078
C28	-0,072	-0,067	-0,062	-0,057	-0,053
C29	0,036	0,034	0,031	0,029	0,0268

Πίνακας 3.3 Οι συντελεστές MFCC ( $C_{15}$  έως  $C_{29}$ )

## 3.4 Ψηφιακό αποτύπωμα ήχων

Για τις αναφερθείσες αρμονικές συχνότητες, των 500,1000,1500,2000 και 2500 Hz, γίνεται καταγραφή και υπολογισμός των συντελεστών MFCC για κάθε συχνότητα. Στη συνέχεια, οι 29 συντελεστές, μετά τον MFCC<sub>0</sub>, δηλαδή οι

MFCC<sub>1</sub>,MFCC<sub>2</sub>,..,MFCC<sub>29</sub> καταγράφονται στην EEPROM του μικροελεγκτή, με την εντολή:

eeprom\_write\_word((uint16t)..address,..value)

η οποία ορίζεται στο αρχείο – επικεφαλίδα « avr/eeprom.h » Οι διευθύνσεις στην EEPROM, οι οποίες επιλέγονται για την καταγραφή, ξεκινούν από 0xA00 και τελειώνουν στη διεύθυνση 0xAF0. Η τιμή του κάθε συντελεστή τύπου MFCC είναι μεταξύ -0,9 και +1,0 και για την αποθήκευση, για τις δεκαδικές αυτές τιμές, αποθηκεύεται ο 10<sup>3</sup> x Ci, δηλαδή αποθηκεύονται ως ακέραιοι, αφού έχουν πολλαπλασιαστεί με 1000. Αυτό γίνεται για εξοικονόμηση μνήμης σε περίπτωση που θα αποθηκεύονταν με δεκαδικές τιμές με μη σημαντικά δεκαδικά ψηφία (ειδικά μετά το 3° δεκαδικό). Έτσι, ο ακέραιος των 16 bit είναι, όπως αναφέρθηκε στην τελευταία παράγραφο, επαρκής για την αναπαράσταση των συντελεστών τύπου MFCC. Επίσης, για κάθε τέτοια ομάδα, αφού υπολογιστούν με DCT ΙΙ μετασχηματισμό οι τελικές τιμές των συντελεστών MFCC, γίνεται «πρόβλεψη» για το διάστημα εντός του οποίου ανήκει η «μη ταυτοποιημένη» αρμονική συγνότητα. Έτσι, έγοντας δημιουγηθεί, μέσω των συντελεστών MFCC, ένα «ψηφιακό» αποτύπωμα των ήχων που αναφέρθηκαν, μπορεί να γίνει πλέον μια «ταυτοποίηση» μιάς νέας αρμονικής συχνότητα, με άγνωστη τιμή f<sub>Hz</sub>, με βάση το κριτήριο που περιγράφηκε στα προηγούμενα (Κεφάλαιο 1.2), δηλαδή όπως παρακάτω:

**Βήμα 1°** : Αποκτούνται οι συντελεστές MFCC(x)i, (i=1,2,...,30) της «άγνωστης» συχνότητας, σύμφωνα με τα περιγραφέντα στάδια της προηγούμενης παραγράφου (2Γ).

**Βήμα 2°** : Υπολογίζεται, για τους συντελεστές MFCC, από 1-29, η «απόσταση Manhattan», για το διάνυσμα Fx, σε σχέση με τους καταγεγραμμένους, στην EEPROM, συντελεστές MFCC των συγκεκριμμένων συχνοτήτων που αναφέρθηκαν. Η ανάγνωση των δεδομένων από την EEPROM γίνεται με την εντολή

eeprom\_read\_word((uint16t)address,..value)

που ορίζεται στο αρχείο «avr/eeprom.h» όπως και η αντίστοιχη της εγγραφής σε ΕΕΡROM. Η «απόσταση», για τα 29-διάστατα διανύσματα MFFCr και MFFCx, (όπου το r αντιπροσωπεύει ένα από τα καταγεγραμμένα σύνολα των 29 συντελεστών τύπου MFCC, για κάποια καταγεγραμμένη γνωστή συχνότητα, ενώ το MFFCx αντιπροσωπεύει το διάνυσμα με 29 στοιχεία για «ταυτοποίηση» συχνότητας, αρχικά άγνωστης), είναι ως εξής:

Ένας αριθμός μικρός δείχνει την «γειτνίαση» των «διανυσμάτων», ενώ ένας μεγάλος αριθμός δείχνει την μεγάλη απόσταση στις συχνότητες των δύο σημάτων. Το επιθυμητό αποτέλεσμα, στην παρούσα περίπτωση, είναι για την «ταυτοποίηση» της «άγνωστης» συχνότητας, να υπάρχουν ζευγάρια, (Mr, Mx), πού θα έχουν «απόσταση Manhattan» όσο τον δυνατόν μικρότερη, (με ιδανική τιμή το μηδέν), ή ισοδύναμα να απέχουν όσο το δυνατόν λιγότερο, στο «διανυσματικό χώρο» των 29 συντελεστών MFCC. Στο παρακάτω σχήμα, αναδεικνύεται η «φυσική» σημασία της απόστασης, ως «μέτρο γειτνίασης» των εξεταζόμενων συντελεστών Mi, για την απλούστερη περίπτωση των 2D διανυσμάτων:





Τα  $\vec{\alpha}$  και  $\vec{\beta}$  είναι «μη παράλληλα» μεταξύ τους και με κοινή αρχή προς χάριν γεωμετρικής εποπτείας. Η «απόσταση Manhattan» εκφράζεται από την τιμή d<sub>M</sub> ( $\alpha$ , $\beta$ )=|X<sub>α</sub>-X<sub>β</sub>|+|Y<sub>α</sub>-Y<sub>β</sub>|, μιάς και έχουν κοινή αρχή. Το «άγνωστο» διάνυσμα  $\vec{w}$  θα έχει αποστάσεις τύπου Manhattan από τα  $\vec{\alpha}$  και  $\vec{\beta}$  με την μορφή:

 $d_{M}(w,\alpha) = |X_{w}-X_{\alpha}| + |Y_{w}-Y_{\alpha}| \text{ kan } d_{M}(w,\beta) = |X_{w}-X_{\beta}| + |Y_{w}-Y_{\beta}|$ 

Από το παραπάνω σχήμα μπορεί άμεσα να συναχθεί ότι το διάνυσμα  $\vec{w}$  έχει τις συντεταγμένες (Xw,Yw) πιο κοντά σε αυτές του διανύσματος  $\vec{a}$  σε σχέση με αυτές

	680	1190	1850	2340	500	1000	1500	2000	2500
C1	-0,225	-0,209	-0,189	-0,174	0,053	-0,216	-0,200	-0,185	-0,169
C2	0,216	0,202	0,182	0,168	0,049	0,208	0,193	0,178	0,163
C3	-0,203	-0,189	-0,170	-0,157	0,043	-0,194	-0,180	-0,167	-0,153
C4	0,184	0,171	0,155	0,142	0,035	0,176	0,164	0,151	0,139
С5	-0,161	-0,150	-0,135	-0,124	0,027	-0,154	-0,143	-0,132	-0,121
C6	0,134	0,124	0,112	0,103	0,018	0,128	0,119	0,110	0,101
C7	-0,103	-0,096	-0,087	-0,080	0,011	-0,099	-0,092	-0,085	-0,078
C8	0,070	0,065	0,059	0,054	0,005	0,067	0,062	0,057	0,053
С9	-0,035	-0,033	-0,029	-0,027	0,001	-0,034	-0,031	-0,029	-0,026
C10	1,40E- 16	1,30E- 16	1,17E- 16	1,08E- 16	2,064	1,34E- 16	1,25E- 16	1,15E- 16	1,05E- 16
C11	0,035	0,033	0,029	0,027	0,001	0,034	0,031	0,029	0,026
C12	-0,070	-0,065	-0,059	-0,054	0,005	-0,067	-0,062	-0,057	-0,053
C13	0,103	0,096	0,087	0,080	0,011	0,099	0,092	0,085	0,078
C14	-0,134	-0,124	-0,112	-0,103	0,018	-0,128	-0,119	-0,110	-0,101
	680	1190	1850	2340	500	1000	1500	2000	2500
C15	0,161	0,150	0,135	0,124	0,027	0,154	0,143	0,132	0,121

του διανύσματος  $\vec{\beta}$ . Έτσι, οι διαφορές |Xw-Xa| και |Yw-Ya| θα είναι επι μέρους μικρότερες (άρα και το άθροισμά τους), από αυτές του  $\vec{w}$  με το  $\vec{\beta}$ , και έτσι τελικά η απόσταση d<sub>M</sub>(w,a) θα είναι μικρότερη της απόστασης d<sub>M</sub>(w,β).

C16	-0,184	-0,171	-0,155	-0,142	0,035	-0,176	-0,164	-0,151	-0,139
C17	0,203	0,189	0,170	0,157	0,043	0,194	0,180	0,167	0,153
C18	-0,216	-0,202	-0,182	-0,168	0,049	-0,208	-0,193	-0,178	-0,163
C19	0,225	0,209	0,189	0,174	0,053	0,216	0,200	0,185	0,169
C20	-0,228	-0,212	-0,191	-0,176	0,054	-0,218	-0,203	-0,187	-0,171
C21	0,225	0,209	0,189	0,174	0,053	0,216	0,200	0,185	0,169
C22	-0,216	-0,202	-0,182	-0,168	0,049	-0,208	-0,193	-0,178	-0,163
C23	0,203	0,189	0,170	0,157	0,043	0,194	0,180	0,167	0,153
C24	-0,184	-0,171	-0,155	-0,142	0,035	-0,176	-0,164	-0,151	-0,139
C25	0,161	0,150	0,135	0,124	0,027	0,154	0,143	0,132	0,121
C26	-0,134	-0,124	-0,112	-0,103	0,018	-0,128	-0,119	-0,110	-0,101
C27	0,103	0,096	0,087	0,080	0,011	0,099	0,092	0,085	0,078
C28	-0,070	-0,065	-0,059	-0,054	0,005	-0,067	-0,062	-0,057	-0,053
C29	0,035	0,033	0,029	0,027	0,001	0,034	0,031	0,029	0,026

## Πίνακας 3.4: Συντελεστές MFCC για αρμ. συχνότητες 680 Hz-2500 Hz

## ΚΕΦΑΛΑΙΟ 4. ΕΠΕΞΕΡΓΑΣΙΑ ΜΕΤΡΗΣΕΩΝ

# 4.1 Φάσμα σήματος με FFT - Ενδεικτικά γραφήματα

Έχοντας ως πηγή σήματος την γεννήτρια AC σήματος που περιγράφηκε πρότερα, λαμβάνονται στον Η/Υ από τον μικροελεγκτή μέσω UART επικοινωνίας, για ενδεικτικές συχνότητες, αναλυτικά οι τιμές των φασματικών κατανομών [X(f)]. Το σήμα της γεννήτριας μετατρέπεται στον μικροελεγκτή σε ψηφιακό, με τη βαθμίδα ADC του μικροελεγκτή, και στη συνέχεια, το λογισμικό που «τρέχει» στον μικροελεγκτή, τα δείγματα x1,x2,,,xN, μετατρέπονται, με FFT κώδικα με N=512 σημεία, σε φασματική κατανομή [X(fn)], όπου τα fn προσδιορίζονται με τη σχέση fn=nfs/N, (n=0,1,2,,,N-1), με συχνότητα δειγματολειψίας fs=44138 Hz. Από τις καταγεγραμμένες φασματικές τιμές, παρατίθενται ορισμένες φασματικές τιμές για συχνότητες 500Hz, 1000Hz, 1500Hz, 2000Hz, και 2500Hz. Από τις κατανομές μπορεί άμεσα να προσδιοριστεί η κεντρική συχνότητα των ημιτονικών σημάτων της πηγής των AC σημάτων.



i) Συχνότητα fo=500Hz

Εικόνα 4.1: Στιγμιαίο πλάτος σήματος x(t) με fo=500Hz, MIC

Φασματική κατανομή ισχύος με #bin συχνότητας, ακουστικό σήμα, μέτρηση με Microphone (fo=500Hz)







ii) Συχνότητα fo=1000Hz

Εικόνα 4.3: Στιγμιαίο πλάτος σήματος x(t) με fo=1000Hz, MIC

Φασματική κατανομή ισχύος, ακουστικό σήμα, μέτρηση με Microphone (fo=1000Hz)







iii) Συχνότητα fo=1500Hz

Εικόνα 4.5: Στιγμιαίο πλάτος σήματος x(t) με fo=1500Hz, MIC

Φασματική κατανομή ισχύος, ακουστικό σήμα, μέτρηση με Microphone (fo=1500Hz)



Εικόνα 4.6: Πλάτος ισχύος σήματος (fo=1500Hz, MIC) με συχνότητα

Ο αριθμός συχνότητας που δίνει ο FFT είναι 11, οπότε η κεντρική συχνότητα στη μεγαλύτερη ισχύ είναι fo=(m/N)fs=(11/512)44138 Hz= 948 Hz, που είναι εντός του ορίου των 86 Hz σφάλματος. Στο λαμβανόμενο φάσμα μπορούν να παρατηρηθούν και ορισμένες συχνότητες με ισχύ στο 10% της συνολικής, με κάποιες από αυτές να εμφανίζονται σε τιμές ως αρμονικές ανώτερης τάξεως. Αυτό μπορεί να αποδοθεί ως ένα βαθμό στη γεννήτρια συχνοτήτων που δίνει και ένα μικρό (<0.5 V) DC σήμα που στο ADC κύκλωμα αυτό οδηγεί σε εμφάνιση ενός μικρού ποσοστού ισχύος και σε άλλες συχνότητες.

#### iv) Συχνότητα fo=2000 Hz



Φασματική κατανομή ισχύος με #bin συχνότητας, ακουστικό σήμα, μέτρηση με Microphone (fo=2000Hz)



Εικόνα 4.8: Πλάτος ισχύος σήματος (fo=2000Hz, MIC) με #bin συχνότητας
Φασματική κατανομή ισχύος με συχνότητα, ακουστικό σήμα, μέτρηση με Microphone (fo=2000Hz)



### Εικόνα 4.9: Πλάτος ισχύος σήματος (fo=2000Hz, MIC) με συχνότητα



v) Συχνότητα fo=2500Hz



Φασματική κατανομή ισχύος με #bin συχνότητας, ακουστικό σήμα, μέτρηση με Microphone (fo=2500Hz)



Εικόνα 4.11:Πλάτος ισχύος σήματος (fo=2000Hz,MIC) με #bin συχνότητας

Φασματική κατανομή ισχύος με συχνότητα, ακουστικό σήμα, μέτρηση με Microphone (fo=2500Hz)



Εικόνα 4.12: Πλάτος ισχύος σήματος (fo=2500Hz,MIC) με συχνότητα

Στο παραπάνω γράφημα, η αρμονική συχνότητα προσδιορόζεται να είναι, με αριθμό bin συχνότητας =24, οπότε fo=(24/512)\*fo=2069, που είναι ικανοποιητικά εντός των ορίων σφάλματος.

#### 4.2 Γραφικές απεικονίσεις μετρήσεων - υπολογισμών

Σε γραφική απεικόνιση, οι συντελεστές MFCC, σε σχέση με την αρμονική συχνότητα, που καταγράφεται κάθε φορά, είναι ως εξής:

#### i)Συχνότητα fo=500Hz



Εικόνα 4.13: Συντελεστές MFCC για fo=500Hz

(μετρούμενη και θεωρητική καμπύλη)

#### ii)Συχνότητα fo=1000Hz



(μετρούμενη και θεωρητική καμπύλη)

#### iii)Συχνότητα fo=1500Hz



Εικόνα 4.15: Συντελεστές MFCC για fo=1500Hz (μετρούμενη και θεωρητική καμπύλη)

### iv)Συχνότητα 2000Hz







### v)Συχνότητα 2500 Hz

Εικόνα 4.17: Συντελεστές MFCC για fo=2500Hz

(μετρούμενη και θεωρητική καμπύλη)

14

Με τη διαδικασία που περιγράφηκε παραπάνω, λαμβάνονται τα παρακάτω δεδομένα για τους συντελεστές MFCC της «άγνωστης» συχνότητας, και συγκρίνονται οι «αποστάσεις» των 29 διαστάσεων με τους καταγεγραμμένους συντελεστές MFCC στην EEPROM, των συχνοτήτων σύγκρισης που έχουν ήδη αναφερθεί, δηλαδή των 500, 1000, 1500, 2000, 2500 Hz.

1. Ταυτοποίηση «άγνωστης» αρμονικής συχνότητας fx (680 Hz).

fx∖f	500	1000	1500	2000	2500
<b>d</b> <sub>M</sub>	0,117	0,173	0,463	0,753	1,043

## Πίνακας 4.1: «Αποστάσεις» της προς ταυτοποίηση συχνότητας, με τους συντελεστές MFCC

Σύμφωνα με τα παραπάνω δεδομένα, τη μικρότερη d<sub>M</sub> δίνει το ζευγάρι (fx,500), (d<sub>M</sub>=0,117) οπότε η ζητούμενη συχνότητα είναι αυτή από τη λίστα των εξεταζόμενων που είναι πιο κοντά στη συχνότητα των 500Hz. Πράγματι, η αναζητούμενη αρμονική έχει τιμή 680Hz, κάτι που είναι αρχικά γνωστό αλλά σε αυτό το σημείο *επαληθεύεται* η «ταυτοποίηση» με το «ψηφιακό» αποτύπωμα της αρμονικής συχνότητας με τους συντελεστές MFCC.

#### 2. Ταυτοποίηση «άγνωστης» αρμονικής συχνότητας fx (1190 Hz).

Η «άγνωστη» συχνότητα καταγράφεται και υπολογίζονται οι συντελεστές MFCC. Οι αποστάσεις d<sub>M</sub> για τους συντελεστές MFCC στις συχνότητες των 500Hz,...,2500 Hz δίνονται από τον παρακάτω πίνακα:

fx∖f	500	1000	1500	2000	2500
$\mathbf{d}_{\mathbf{M}}$	0,564	0,174	0,276	0,800	1,421

## Πίνακας 4.2: «Αποστάσεις» της προς ταυτοποίηση συχνότητας, με τους συντελεστές MFCC

Από τον παραπάνω πίνακα εύκολα διακρίνεται ότι η μικρότερη απόσταση d<sub>M</sub> είναι για την συχνότητα των 1000 Hz (d<sub>M</sub>=0,174). Από τη λίστα των εξεταζόμενων

συχνοτήτων η συχνότητα αυτή είναι η των 1190 Hz. Αυτό επαληθεύει την επιλογή της πλησιέστερης συχνότητας με τη βοήθεια της μετρικής Manhattan για τους συντελεστές MFCC.

#### 3. Ταυτοποίηση «άγνωστης» αρμονικής συχνότητας fx (1850Hz)

Από τις καταγραφές και τον υπολογισμό των MFCC για τη συχνότητα fx, λαμβάνονται τα παρακάτω δεδομένα:

fx∖f	500	1000	1500	2000	2500
d <sub>M</sub>	0,792	0,502	0,212	0,077	0,368

## Πίνακας 4.3: «Αποστάσεις» της προς ταυτοποίηση συχνότητας, με τους συντελεστές MFCC

Από τον παραπάνω πίνακα συνάγεται ότι η μικρότερη απόσταση είναι η d<sub>M</sub>=0,077. Έτσι, το «ζευγάρι» (fx,2000Hz) επιλέγεται ανάμεσα στα άλλα, μιάς και δίνει τη μικρότερη απόσταση τύπου Manhattan. Πράγματι, η εφαρμοσθείσα συχνότητα ήτο η των 1850 Hz, που είναι πιο κοντά σε αυτή των 2000Hz.

#### 4. Ταυτοποίηση «άγνωστης» αρμονικής συχνότητας fx (2340 Hz)

Σε αυτή την περίπτωση, από τις καταγραφές και τον υπολογισμό των MFCC για τη συχνότητα fx, υπολογίζονται στον πίνακα:

fx∖f	500	1000	1500	2000	2500
d <sub>M</sub>	1,070	0,785	0,490	0,200	0,090

## Πίνακας 4.4: «Αποστάσεις» της προς ταυτοποίηση συχνότητας, με τους συντελεστές MFCC

Από τον παραπάνω πίνακα είναι φανερό ότι η μικρότερη απόσταση είναι η  $d_M=0,090$ . Έτσι, το «ζευγάρι» (fx,2500Hz) επιλέγεται ανάμεσα στα άλλα, μιάς και δίνει τη μικρότερη απόσταση τύπου Manhattan. Πράγματι, η εφαρμοσθείσα συχνότητα ήτο η των 2340 Hz, που είναι αυτή στην εξεταζόμενη λίστα συχνοτήτων που είναι πιο κοντά στην συχνότητα 2500Hz.

# ΚΕΦΑΛΑΙΟ 5. Συμπεράσματα-Επίλογος

## 5.1 Ανάλυση δεδομένων από τις καταγραφές – ταυτοποιήσεις

Στα προηγούμενα σημεία που περιγράφηκαν για το πειραματικό setup, τις καταγραφές αρμονικών συχνοτήτων και τις μετρήσεις συχνοτήτων για αναγνώριση των αρμονικών τιμών, χρησιμοποιήθηκε η ανάλυση Fourier για το φάσμα των σημάτων, και η μεθοδολογία των συντελεστών MFCC για την καταγραφή και την αναγνώριση αρμονικών ήχων. Με βάση τις ληφθείσες μετρήσεις, η αναγνώριση των αρμονικών ήχων γίνεται με ακρίβεια 100%, όσον αφορά τιμές που διαφέρουν τουλάγιστον κατά ένα frequency bin, που για δειγματοληψία στα 16000Hz και για 512 σημεία είναι 16000/512=31.2 Hz. Αυτό είναι το όριο «διακριτότητας» στις αναγνωρίσεις ήχων, στο σύστημα που χρησιμοποιήθηκε. Η δυνατότητα επεξεργασίας σε πραγματικό χρόνο και η δυνατότητα αποθήκευσης στοιχείων του μικροελεγκτή, έκανε εφικτή την «ενσωμάτωση» όλου του υπολογιστικού φορτίου σε μία μονάδα (processing module), έτσι ώστε να μπορεί να αποτελεί αυτόνομο υπολογιστικό τμήμα μίας συσκευής που χρησιμοποιεί αναγνώριση ή ταυτοποίηση ήχων. Οι μετρήσεις μέσω μικροφώνου αποδείχνεται ότι είναι σε μεγάλη συνέπεια (πάνω από 90%) με τις θεωρητικά προβλεπόμενες. Για παράδειγμα, όπως διαφαίνεται και από την σχετική εικόνα για τους MFCC στην αρμονική συχνότητα fo=2000Hz, η σχετική απόκλιση των συντελεστών μέσω MIC σε σχέση με τους θεωρητικά προβλεπόμενους είναι της τάξης μεγέθους του 10% (πχ για mel bin 4 είναι: 0,150-0,135=0,015 και σχετική απόκλιση 0,015/0,150=10%).

Στις ληφθείσες μετρήσεις και στους μετέπειτα υπολογισμούς τόσο του φάσματος κατά Fourier αλλά και των συντελεστών MFCC, σημαντικό ρόλο παίζει η συχνότητα δειγματοληψίας για την μετατροπή αναλογικού σε ψηφιακό σήμα, που στην πειραματική διαδικασία είχε τιμή 16000Hz. Σε μεγαλύτερη τιμή, για παράδειγμα στα 44100Hz, η βελτίωση της σχετικής ακρίβειας των πειραματικά αποκτηθέντων τιμών για το φάσμα και τους MFCC σε σχέση με τις θεωρητικές προβλέψεις μπορεί να έφθανε να υπάρχουν λιγότερο και από 5% αποκλίσεις.

Έχοντας υπολογίσει τους συντελεστές MFCC για τις διάφορες «δοκιμές» με τις αναφερθείσες συχνότητες, μπορεί να παρατηρηθεί πλην άλλων μία «φυσιολογική» συμπεριφορά της μετρικής Manhattan σε σχέση με τις κοντινές και μακρινές συχνότητες που εντοπίζονται γύρω από την εξεταζόμενη συχνότητα. Με άλλα λόγια, η απόσταση Manhattan μεγαλώνει σχετικά αναλογικά όσο αυξάνεται η διαφορά δύο αρμονικών συχνοτήτων, ενώ μικραίνει σχετικά αναλογικά όσο μειώνεται η διαφορά των δύο συχνοτήτων. Αν και οι αυξήσεις/μειώσεις είναι αναλογικές αλλά όχι γραμμικές, δείχνουν ότι στο χώρο των συχνοτήτων οι συντελεστές MFCC λειτουργούν ως άξονες αναφοράς σε αρμονικές συχνότητες και έτσι ένας σύνθετος ήχος δεν είναι παρά μια «υπερεπιφάνεια» στο χώρο R<sup>K</sup> των K-συντελεστών MFCC. Τέτοιες «τεμνόμενες επιφάνειες» παρουσιάζουν κοινό σύνολο σημείων άρα και αρμονικών συχνοτήτων ως φασματικό περιεχόμενο.

## 5.2 Συμπεράσματα - Προτάσεις

Με βάση τα περιγραφέντα στην παρούσα μελέτη, αναδεικνύεται ο καθοριστικός ρόλος της μετατροπής αναλογικού σε ψηφιακό σήμα με την τεχνική της δειγματοληψίας, ο ρόλος των μετασχηματισμών Fourier στην απόκτηση του «φασματικού» περιεχομένου ενός σήματος, αλλά και η μεγάλη χρησιμότητα των συντελεστών MFCC σε δημιουργία «ψηφιακού» αποτυπώματος των ήχων με μορφή που είναι ανεξάρτητη από τη συσκευή ή την πηγή που παράγει τον ήχο αλλά αφορά τα αυτούσια χαρακτηριστικά του ίδιου του ηχητικού σήματος.

Έτσι, η ψηφιακή μετατροπή δειγμάτων ήχου που συλλέγονται από ένα απλό μικρόφωνο σε διασύνδεση με μικροελεγκτή, αποδεικνύεται αποτελεσματικός τρόπος παραγωγής ψηφιακών δειγμάτων, αφού συνδυάζει την ADC δυνατότητα του μC, και το interface του μC με εξωτερικές συσκευές μέσω των ψηφιακών του pins. Στο επόμενο βήμα, της επεξεργασίας του ψηφιακού σήματος, η επεξεργαστική ισχύς του μC, σε υψηλές συχνότητες ρολογιού (32MHz), καθώς και η δυνατότητα αποθήκευσης δεδομένων μόνιμα στην EEPROM, δίνουν σε τέτοια υπολογιστικά συστήματα όπως οι μC το ρόλο της αυτόνομης λειτουργίας μιάς υπολογιστικής μονάδας σε ένα συγκεκριμμένο «καθήκον», με τα πλεονεκτήματα του μικρού όγκου των συσκευών αυτών, την μεταφερσιμόητα και εγκατάσταση τέτοιων συσκευών αλλά και στις δυνατότητες διασύνδεσης τέτοιων συσκευών με άλλα περιφερειακά modules.

Σε σχέση με τους συντελεστές MFCC που περιγράφονται για τις αρμονικές συχνότητες, η διαδικασία μπορεί άμεσα να επεκταθεί και να περιλαμβάνει «ταυτοποίηση» και φυσικών σύνθετων ήχων και ανθρώπινης ομιλίας. Παρόμοιες κατευθύνσεις περιγράφονται να έχουν γίνει στη διεθνή βιβλιογραφία, αλλά δεν υπάρχουν μέχρι τώρα σχετικά στοιχεία για υλοποιήσεις με την Ελληνική Γλώσσα, όπως πχ για ελληνικά φωνήεντα, διφθόγγους κ.α.

Οι τεχνικές για τους συντελεστές MFCC μπορούν να «συγκεράσουν» και τους μετασχηματισμούς (wavelet transforms) που «συγχρονίζουν» τα δείγματα στο χρόνο για την εξαγωγή των αντίστοιχων «ενεργειακών» τιμών για τους συντελεστές, έτσι ώστε να βελτιωθεί η σχετική ακρίβεια υπολογισμών σε τιμές πάνω από 95% σε σχέση με τις θεωρητικές προβλέψεις.

## 5.3 ΕΠΙΛΟΓΟΣ

Στην παρούσα μελέτη υλοποιήθηκε καταγραφή και αναγνώριση ήχων (αρμονικές συχνότητες) με τη χρήση πειραματικού εξοπλισμού με μικροελεγκτή μC, μικροφώνου, γεννήτριας ακουστικών σημάτων και παλμογράφο, καθώς και σειριακή διασύνδεση δεδομένων με PC –Windows. Ο σκοπός της παρούσας μελέτης είναι αφενός μεν να αναδειχθεί ο ρόλος τέτοιων υπολογιστικών μονάδων σε αυτόνομη διεκπεραίωση υπολογιστικού φορτίου, αλλά και αφετέρου να αναδειχθεί η μεγάλη χρησιμότητα της έννοιας του «φάσματος» συχνοτήτων των σημάτων, μέσα από τις διαδικασίες αναλογικής σε ψηφιακή μορφή αλλά και χρήσης τεχνικών όπως ο μετασχηματισμός Fourier και οι μεθοδολογίες των συντελεστών MFCC.

Με τη χρήση δειγμάτων από αρμονικούς ήχους, σε συνδυασμό με τα στοιχεία που αναφέρθηκαν, περιγράφηκαν οι χρονικές και φασματικές κατανομές των σημάτων αυτών, καθώς και περαιτέρω επεξεργασία τους με τη μεθοδολογία των συντελεστών MFCC. Η τεχνική της «αναγνώρισης» και «ταυτοποίησης» των ηχητικών σημάτων, μέσω των συντελεστών MFCC και των δυνατοτήτων αποθήκευσης στοιχείων του μC, ανοίγει ένα μονοπάτι για περαιτέρω έρευνα και πειραματικές υλοποιήσεις σε σχέση με ήχους της Ελληνικής Γλώσσας, αλλά και με επέκταση των στοιχείων που περιγράφηκαν σε άλλες εφαρμογές, όπως αναγνώριση ήχων συναγερμών, ασθενοφόρων, μηχανημάτων σε λειτουργία κλπ. Οι δυνατότητες των μικροελεγκτών να επεξεργάζονται και να αποθηκεύουν τα ψηφιακά δεδομένα των ήχων όπως αυτά παράγονται από τις βαθμίδες ADC, καθώς και η ευκολία διασύνδεσης συσκευών καταγραφής ήχων (πχ μικρόφωνα) στους μικροελεγκτές, κάνουν τις συσκευές αυτές ισχυρά εργαλεία σε πολλές πρακτικές εφαρμογές.

Με την παρούσα μελέτη, η βασική φιλοδοξία είναι να δώσει ένα «επιστημονικό ερέθισμα» στον αναγνώστη, στην χρήση των μC ως φορητές και αυτόνομες μονάδες σε εφαρμογές για καταγραφές ή επεξεργασία ήχων και ηχητικών σημάτων, αλλά και

στην κατεύθυνση της εφαρμογής των τεχνικών των μετασχηματισμών Fourier και της μεθοδολογίας των συντελεστών MFCC σε ηχητικές αναλύσεις για την Ελληνική Γλώσσα, που πέρα από τη διεθνώς αναγνωρισμένη μαθηματική της δομή έχει και το ηχητικό περιεχόμενο και σύνθεση σε αρμονία με την ανθρώπινη ομιλία που εκφράζει την ανθρώπινη σκέψη ως ένδειξη της πολυπλοκότερης λειτουργίας που έχει ο άνθρωπος σε ένα Κόσμο που μεταβάλλεται και συνεχώς εξελίσσεται.

## ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΠΑΡΑΠΟΜΠΕΣ

Brigham, E., (1988) 'Fast Fourier Transform and Its Applications, PPearson, 1st edition, 1988.

Rao, K.R., Kim D.N. & Hwang, J.J., (2010). Fast Fourier Transform - Algorithms and Application, Springer, 2010.

Kaiser, A., (2002) *Digital Signal Processing using the Fast Fourier Transform*, GRIN Verlag, 1st edition, 2002.

Owen, M., (2012), *Practical Signal Processing*, : Cambridge University Press, Reprint edition, 2012.

Ramirez, R., (1984), *The FFT, Fundamentals and Concepts*, : Prentice Hall, First Edition, 1984.

Mix, D., (2015) *Student's Guide to Discrete Fourier and z Transforms. Sampling, Multirate Processing, and the FFT*, Publisher: CreateSpace Independent Publishing Platform; Technical LAP Series, Volume 4, Student Edition, 22 March 2015.

Mulgrew, B. et al., (2002). *Digital Signal Processing: Concepts and Applications*, Published by: Palgrave, 2nd Edition, 9 September 2002.

Champeney, D.C., (1973). *Fourier Transforms and Their Physical Applications*, Published by: Academic Press, 1 June 1973.

Leaver, C., (2010). *Introduction to Atmel AVR Microcontroller Development*, Published by: Sylvania Books.

Grace, T., (2015). *Programming and Interfacing ATMEL's AVRs*, Published by: Cengage Learning PTR, 1<sup>st</sup> edition.

Barr, M. & Massa, A., (2006). *Programming Embedded Systems: With C and GNU Development Tools*, O'Reilly Media: Second edition.

Tarafder, R., M., (2017), *Performance Analysis and Improvement of Internal Architecture of Atmel Microcontroller*, Published by: LAP LAMBERT Academic Publishing.

Barnett, R, Cox, S. & O'Cull, L., (2006), *Embedded C Programming and the Atmel AVR*, Published by: Cengage Learning, 2 edition.

Gadre, D., (2009). *Programming and Customizing the AVR Microcontroller*, Published by:MacGraw-Hill, 2 edition.

Huang, H., (2013). *The Atmel AVR Microcontroller: MEGA and XMEGA in Assembly and C*, Published by: Cengage Learning US.

Oliver, B.M., Pierce, J.R., & Shannon, C.E., (1948). 'The Philosophy of PCM'. *Proceedings of the IRE.* V. **36** (11): 1324–1331.

Yang, D. & Wu, S., (2014). *Fundamentals of Liquid Crystal Devices*, Published by: Wiley, 2nd edition, Series in Display Technology.

Cristaldi, D., Pennisi, S. & Pulvirenti, F. (2009). *Liquid Crystal Display Drivers: Techniques and Circuits*, Published by: Springer, 2009 edition.

Davis, S. & Mermelstein, P., (1980). 'Comparison of parametric representations or monosyllabic word recognition in continuously spoken sentences' *IEEE Transactions on ASSP28*: 357–366.

Stallings, W., (1984). 'Digital Signaling Techniques', *IEEE Communications Magazine*, Vol. 22, No. 12, December 1984.

Cooley, J., Tukey, W. & John, W., (1965). 'An algorithm for the machine calculation of complex Fourier series', *Mathematics of Computation*, 19 (90): 297–301.

Cooley, J., Lewis, P., & Welch, P., (1969). 'The finite Fourier transform', *IEEE Transactions on Audio and Electroacoustics*, **17** (2): 77–85.

Proakis, J., Manolakis, G. & Dimitri, G., (1996). *Digital Signal Processing: Principles, Algorithms and Applications*, (3 ed.), Upper Saddle River, New Jersey:Prentice-Hall, 3<sup>rd</sup> Edition,

Duhamel, P., Piron, B. & Etcheto, J, (1988). 'On computing the inverse DFT', *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 36 (2), pp. 285–286.

Lu,Y. & Wu, Z. (2013). 'Maximum likelihood subband polynomial regression for robust speech recognition', *Applied Acoustics*, vol.74, pp. 640-646.

Rabiner, L.R. & Juang, B., (1993). *Fundamentals of Speech Recognition*, Prentice Hall.

Bridle, J.S. and Brown, M.D., (1974). 'An Experimental Automatic Word-Recognition System", *JSRU Report No. 1003, Joint Speech Research Unit*, Ruislip, England.

Morgan, N., Bourlard, H. & Hermansky, H., (2004). 'Automatic Speech Recognition: An Auditory Perspective', In Steven Greenberg & William A. Ainsworth (eds.). *Speech Processing in the Auditory System*. Springer. p. 315.

Young S.J., Woodland, P. & Byrnet, W.J., (1993). *HTK: Hidden Markov Model Toolkit*, V 1.5, Cambridge University Engineering Department Speech Group and Entropic Research Laboratories Inc.

Jain, A. & Harris, J., (2004). *Speaker identification using MFCC and HMM based techniques*, University of Florida, 25 April 2004.

Ganchev, T., Fakotakis, N. and Kokkinakis, G., (2005). 'Comparative evaluation of various MFCC implementations on the speaker verification task' in *10th International Conference on Speech and Computer* (SPECOM 2005), Vol. 1, pp. 191–194.

Xiong, X., (2009). 'Robust speech features and acoustic models for speech recognition', PhD. Thesis, 194 p., Nanyang Technological University, Singapore.

Anusuya, M. & Katti, S., (2011). 'Front end analysis of speech recognition: a review', *International Journal of Speech Technology*, vol. 14, pp. 99-145, 2011.

Axelson, J., (2007). Serial Port Complete: COM Ports, USB Virtual COM Ports, and Ports for Embedded Systems, Published by: Lakeview Research; 2 edition, December 1, 2007.

Campbell, J., (1993). '*C Programmer's Guide to Serial Communications*', Published by: Sams, Subsequent edition, 1 September 1993.

Papazoglou, P., (2018). *An Educational Guide to the AVR Microcontroller Programming*, Published by: CreateSpace Independent Publishing Platform, 1 edition, March 16, 2018.

Mazidi, M. et al., (2010). AVR Microcontroller and Embedded Systems: Using Assembly and C, Published by: Pearson, 1st edition, January 31, 2010.

Willi, E., (2014). '*AVR Programming: Learning to Write Software for Hardware*', Published by: Maker Media, Inc, 1st edition, 17 February 2014.

# Πηγές στο Internet

(Μικροελεγκτές)

[1]https://www.microchip.com/design-centers/8-bit/avr-mcus

[2]https://www.microchip.com/design-centers/microcontrollers

(MCU του μικροελεγκτή)

[3]https://www.microchip.com/wwwproducts/en/ATxmega256A3U

(LCD οθόνες)

[4]<u>https://www.youtube.com/watch?v=DoqorBJ1-38</u>

(RS232)

[5]https://www.youtube.com/watch?v=XVEnxipCIJ0

(UART)

[6] <u>https://www.youtube.com/watch?v=aW61hlliiS0</u>

## ΠΑΡΑΡΤΗΜΑ Ι

## **Α )** κωλικάς fft

```
//FFT
void _fft(cplx buf[], cplx out[], int n, int step)
{
      if (step < n) {
            fft(out, buf, n, step * 2);
            fft(out + step, buf + step, n, step * 2);
            for (int i = 0; i < n; i += 2 * step) {</pre>
            cplx t = (cos(PI * i / n) -I*sin(PI*i/n))* out[i + step];
                  buf[i / 2] = out[i] + t;
                  buf[(i + n)/2] = out[i] - t;
            }
      }
}
void fft(cplx buf[], int n)
{
      cplx out[n];
      for (int i = 0; i < n; i++) out[i] = buf[i];</pre>
      fft(buf, out, n, 1);
}
int find syxnotha(cplx xtimes[])
{
```

```
double maxisxys=0.0;
int uu=0;
int marka=-1;
double zza=0.0;
double zzii=0.0;
double tempa=0.0;
fft(xtimes, NFFT);
for (uu=0; uu<NFFT; uu++)</pre>
{
      zza=creal(xtimes[uu]);
      zzii=cimag(xtimes[uu]);
      tempa=zza*zza+zzii*zzii;
      tosxu[uu]=tempa/(double)NFFT;
}
maxisxys=tosxu[0];
marka=0;
for (uu=1; uu< (NFFT/2); uu++)</pre>
{
      if (tosxu[uu]>maxisxys)
      {
            maxisxys=tosxu[uu];
            marka=uu;
      }
      else
      {
            ;
      }
```

```
}
return marka;
}
```

#### ΚΩΔΙΚΑΣ Υπολογισμού συντελεστών MFCC

```
//MFCC
void find mel mfcc()
{
      double mspaca=0.0;
      int fm minus=0;
      int fm m=0;
      int fm plus=0;
      FREQS[0] = 0.0;
      FREQS[MTotal - 1] = FSAMPL/2.0;
      mels[0] = 2595.0 * log10(1.0 + (FREQS[0] / 700.0));
      mels[MTotal - 1] = 2595.0 * log10(1.0 + (FREQS[MTotal - 1] /
700.0));
      mspaca=(mels[MTotal-1]-mels[0])/(double)(MTotal);
      for(yy=1;yy<(MTotal-1);yy++)</pre>
      {
            mels[yy]=mels[0]+mspaca*(double)yy;
      }
      for (yy = 1; yy < (MTotal-1); yy++)
      {
            FREQS[yy] = 700.0 * (pow(10.0, (mels[yy] / 2595.0)) -
1.0);
      }
      for (yy = 0; yy < MTotal; yy++)
      {
```

```
bins[yy] = (int) (floor(FREQS[yy] * ((double)NFFT+1.0) /
FSAMPL));
      }
      for (yy = 1; yy < MTotal-1; yy++)
      {
            energy[yy-1]=0.0;
            fbank=0.0;
            fm minus=bins[yy-1];
            fm_m=bins[yy];
            fm_plus=bins[yy+1];
            for(kk=0;kk<(NFFT/2);kk++)</pre>
            {
            if ((kk>=fm minus) && (kk<fm m))
            {
                  fbank =(double)(kk-fm minus)/(double)(fm m-
fm minus);
            }
            else if (kk==fm_m)
            {
                  fbank=1.0;
            }
            else if ((kk>fm m) && (kk<=fm plus))
            {
                  fbank=(double)(fm plus-kk)/(double)(fm plus-fm m);
            }
            else
            {
                  fbank=0.0;
            }
                  energy[yy-1] += (isxys[kk] * fbank);
```

```
}
             encepstr[yy-1] = log10(energy[yy-1]);
      }
for(yy=0;yy<MTotal;yy++)</pre>
{
      kyros[yy]=encepstr[yy]+I*0.0;
}
fft(kyros,MTotal);
for(yy=0;yy<MTotal;yy++)</pre>
{
      energy[yy]=sqrt(creal(kyros[yy])*creal(kyros[yy])+cimag((kyros[
yy])*cimag(kyros[yy])));
}
findDCT(encepstr,MTotal);
}
void findDCT(double totransf[],int nto)
{
      int uu=0;
      for (yy=0; yy< (nto-2); yy++)</pre>
       {
             energy[yy]=0.0;
             for (uu=0; uu< (nto-2); uu++)</pre>
             {
             energy[yy] +=
totransf[uu]*cos(PI*(double)yy*((double)(uu)-0.50)/(double)(nto-2));
             }
      }
}
```

## **Β**) ΚΩΔΙΚΑΣ Ρ**c**-μC με UART

```
/* χρήση του USARTCO και αρχικοποίηση buffers. */
USART InterruptDriver Initialize(&USART data, &USART,
USART DREINTLVL LO gc);
         /* USARTCO, 8 Data bits, No Parity, 1 Stop bit. */
         USART Format Set(USART data.usart, USART CHSIZE 8BIT gc,
                USART PMODE DISABLED gc, false);
         /* Enable RXC interrupt. */
         USART RxdInterruptLevel Set(USART data.usart,
USART_RXCINTLVL_LO_gc);
         /* Ανάθεση γιά Baudrate to 9600 bps:
         * Baudrate select = (1/(16*(((I/O clock
frequency)/Baudrate)-1)
         *
                           = 207
         */
        USART_Baudrate_Set(&USART, 207 , 0);
         /* Enable каι τα RX and TX. */
         USART_Rx_Enable(USART_data.usart);
         USART Tx Enable(USART data.usart);
```

ISR(USARTC0\_RXC\_vect)

```
{
    USART_RXComplete(&USART_data);
}
ISR(USARTC0_DRE_vect)
{
    USART_DataRegEmpty(&USART_data);
}
```

///