



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

UNIVERSITY OF PIRAEUS

ΑΝΑΛΥΣΗ ΑΛΓΟΡΙΘΜΟΥ ΑΝΙΧΝΕΥΣΗΣ ATTACK TRANSIENT
ΚΑΙ Η ΕΦΑΡΜΟΓΗ ΤΟΥ ΣΕ ΛΟΓΙΣΜΙΚΟ
ΜΟΝΤΕΛΟΠΟΙΗΜΕΝΗΣ ΣΥΝΘΕΣΗΣ ΚΥΜΑΤΟΜΟΡΦΩΝ.

**ATTACK TRANSIENT DETECTION ALGORITHM ANALYSIS AND ITS IMPLEMENTATION
USING A MODELED WAVEFORM SYNTHESIS SOFTWARE.**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ «ΠΛΗΡΟΦΟΡΙΚΗ»

ΠΑΝΑΓΙΩΤΗΣ ΓΕΩΡΓΙΟΥ / ΜΠΠΛ 13009

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: ΑΓΓΕΛΟΣ ΠΙΚΡΑΚΗΣ | ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ

ΙΟΥΝΙΟΣ 2017

ΤΡΙΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:

(ΥΠΟΓΡΑΦΗ)

(ΥΠΟΓΡΑΦΗ)

(ΥΠΟΓΡΑΦΗ)

ΑΓΓΕΛΟΣ ΠΙΚΡΑΚΗΣ | ΠΑΝΑΓΙΩΤΗΣ ΚΟΤΖΑΝΙΚΟΛΑΟΥ | ΜΙΧΑΗΛ ΨΑΡΑΚΗΣ
ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ

Πρόλογος και Ευχαριστίες.

Η εκπόνηση της μεταπτυχιακής διατριβής ήταν για τον γράφοντα, η ταυτόχρονα η πιο ευχάριστη και η πιο δύσκολη περίοδος των σπουδών. Η αιτία αυτής της αντίφασης, είναι η θεματολογία της διατριβής. Αυτή, αν και ανήκει στο πεδίο ενδιαφέροντος του γράφοντος, την επεξεργασία σημάτων, είναι μια ξεχωριστή επιστήμη από μόνη της και απαιτεί παχυλή γνώση για να μπορέσει να εφαρμοστεί.

Ως εκ τούτου, για την κατανόηση του αλγορίθμου που υλοποιήθηκε και της υλοποίησης του λογισμικού, χρειάστηκε να επενδυθεί πρώτα αρκετός χρόνος στην κατανόηση εννοιών οι οποίες ήταν άγνωστες και απαιτήθηκε μελέτη αρκετών συγγραμμάτων με θεματολογία την επεξεργασία σημάτων και την ψυχοακουστική.

Τα ωφέλη που αποκομήθηκαν είναι πολλαπλά, τόσο στην γνώση για την ερμηνεία προβλημάτων αλγοριθμικής φύσης σε προγραμματιστικούς όρους, όσο και στην υλοποίηση μιας αρχιτεκτονικής ενός λογισμικού επεξεργασίας σήματος.

Η παρούσα εργασία, αποτελεί το πρώτο μεγάλο βήμα -και ενδεχομένως το «εισιτήριο» για μια επαγγελματική σταδιοδρομία- στην επεξεργασία σημάτων ήχου, ένα πεδίο στο οποίο ο γράφων τρέφει ιδιαίτερη αγάπη.

Ιδιαίτερες ευχαριστίες θα πρέπει να δοθούν στον επιβλέποντα καθηγητή Κ. Άγγελο Πικράκη, πρωτίστως για την υπομονή που έδειξε στην επίβλεψη και στην λεπτομερή επεξήγηση, ακόμη και βασικών θεωριών όταν υπήρξε ανάγκη. Επίσης, ευχαριστίες θα πρέπει να δοθούν στους καθηγητές Κ.Γιώργο Τσιχριντζή και Κ. Χρήστο Δουληγέρη, για τις σημαντικές γνώσεις που αποκτήθηκαν μέσα από το μάθημα «Πολυμεσικά Σήματα και Συστήματα», όπως και στον Κ. Γιάννη Μουρτζόπουλο, καθηγητή του Πανεπιστημίου Πατρών, για τις κατευθυντήριες γραμμές που παρείχε.

*Αφιερωμένο στον Δημιουργό,
για το υπέροχο δώρο της Σύνθεσης.*

Περιεχόμενα

| | |
|---|----|
| Πρόλογος και Ευχαριστίες | 3 |
| Περίληψη | 6 |
| Κεφάλαιο 1: Εισαγωγικά..... | 7 |
| 1.1: Η έννοια της σύνθεσης ήχου. | 7 |
| 1.2: Τα είδη των ταλαντωτών στους συνθετητές. | 9 |
| 1.3: Η χρήση αποθηκευμένων δειγμάτων στους συνθετητές και οι ιδιότητές τους ως ταλαντωτές. | 10 |
| 1.4: Η χρήση δειγμάτων κατά την μέθοδο παραγωγής ήχου “Sample and Synthesis”. | 12 |
| Κεφάλαιο 2: Ο ρόλος του Attack Transient στην αντίληψη των ήχων..... | 15 |
| 2.1: Ο Ορισμός του transient. | 15 |
| 2.2: Ανάλυση του attack transient ως συμβάν. | 15 |
| 2.3: Μέθοδοι εντοπισμού και απομόνωσης του attack transient σε ένα σήμα. | 18 |
| Κεφάλαιο 3: Περιγραφή του αλγορίθμου ανάλυσης και εξαγωγής του Attack Transient. | 20 |
| 3.1: Περιορισμοί λειτουργίας και προεπεξεργασία σήματος εισόδου. | 20 |
| 3.2: Η επαναληπτική διαδικασία εξαγωγής του Attack transient. | 21 |
| 3.3: Η Υλοποίηση του αλγορίθμου στο Matlab. | 24 |
| Κεφάλαιο 4: Το λογισμικό MLAS (Modeled Linear Arithmetic Synthesis)..... | 27 |
| 4.1: Η φιλοσοφία πίσω από το λογισμικό. | 27 |
| 4.2 Περιγραφή των επιμέρους μερών του λογισμικού. | 29 |
| 4.3 Περιγραφή της λειτουργίας του λογισμικού. | 31 |
| 4.4 Εμπορικές προοπτικές χρήσης του αλγορίθμου και του λογισμικού..... | 37 |
| Βιβλιογραφία | 38 |

Περίληψη

Η σημαντικότητα του Attack Transient στην ταυτότητα ενός μουσικού οργάνου και στον καθορισμό του ηχοχρώματός του, έχει γίνει αντικείμενο ενασχόλησης αρκετών μελετητών. Μία από τις εφαρμογές της ανωτέρω θεωρίας έγινε στη δημιουργία της «Sample and Synthesis», μιας μεθόδου σύνθεσης κυματομορφών που έπαιξε σημαντικό ρόλο στην εμπορική επιτυχία πολλών συνθετητών.

Στην παρούσα μελέτη γίνεται εκτενής περιγραφή του Attack Transient ως γεγονός. Επιπλέον αναλύεται και υλοποιείται ένας αλγόριθμος ανίχνευσης Attack Transient σε Matlab. Η μελέτη επίσης περιλαμβάνει την περιγραφή της « S&S » και την χρήση του υλοποιημένου αλγόριθμου στο λογισμικό μοντελοποιημένης σύνθεσης κυματομορφών « Modeled Linear Arithmetic Synthesis ».

Abstract

Attack transient's importance in defining the identity of a musical instrument and its timbre, has been a subject of interest for many studiers. One of the implementations of the above theory can be found in "Sample and Synthesis", a method of waveform synthesis that played an important role in the commercial success of many synthesizers.

In the following study, there is an extensive description of Attack Transient as an event. Moreover, there is analysis and implementation of an attack transient detection algorithm in Matlab. The study also includes a description of "S&S" and the use of the implemented algorithm in the modeled waveform synthesis software, "Modeled Linear Arithmetic Synthesis".

Κεφάλαιο 1: Εισαγωγικά.

1.1: Η έννοια της σύνθεσης ήχου.

Η σύνθεση νέων ήχων είναι ένα αρχέγονο «πρόβλημα» το οποίο απασχολεί τον άνθρωπο από τις απαρχές της ύπαρξής του. Πρώιμοι συνθετητές ήχων θα μπορούσαν να χαρακτηριστούν τα πρώτα πνευστά που από καλάμια, τα κούτσουρα που παρήγαγαν ήχο κρούοντάς τα, ακόμη και το σφύριγμα που παράγει ο άνθρωπος, δημιουργώντας ένα πνευστό με τα χείλη του. Πρακτικά, εν αντιθέσει με την σύνδεση της λέξης συνθετητής (synthesizer) με τα ηλεκτρονικά μέσα σύνθεσης ήχου, ως συνθετητή θα μπορούσαμε να χαρακτηρίσουμε οποιοδήποτε μουσικό όργανο, από τη στιγμή που παράγει ήχο με την βοήθεια συνένωσης επιμέρους στοιχείων (1). Επί παραδείγματι στην κιθάρα, υπάρχουν οι χορδές, το ηχείο η ταστιέρα και πολλά άλλα μέλη τα οποία συμβάλλουν στην παραγωγή και μορφοποίηση του ήχου με τη συγκεκριμένη χροιά που καταλήγει στα αυτιά μας.

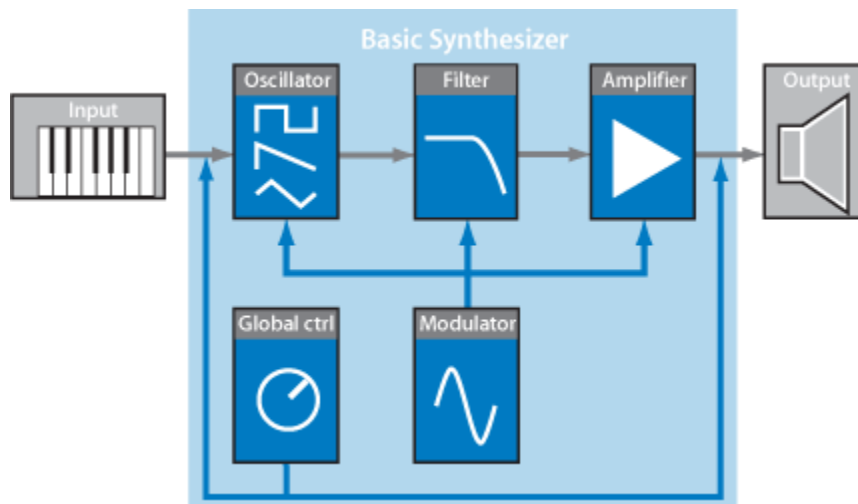
Σύνθεση στην ευρύτερη της έννοια, θα μπορούσε να οριστεί ως η διαδικασία της συνένωσης επιμέρους τμημάτων με σκοπό την συναρμογή ενός νέου, ολοκληρωμένου κατασκευάσματος. Η διαδικασία αυτή, διαχωρίζεται από αυτήν της απλής συνένωσης, καθότι η σύνθεση εμπεριέχει την διαδικασία της δημιουργικής εμπλοκής του ανθρώπου στην επιλογή των επιμέρους τμημάτων.(2)

Συγκεκριμένα, κατά την σύνθεση ενός ήχου, χρησιμοποιούνται μια πλειάδα τεχνικών γνώσεων, από τις επιστήμες της φυσικής των μαθηματικών, ακόμη και της βιολογίας είτε αυτός παράγεται με ηλεκτρονικά μέσα είτε με φυσικά.(3) Στο παράδειγμα του ήχου που παράγει μια κιθάρα, το ηχόχρωμα της εξαρτάται από το ξύλο που χρησιμοποιείται κατά την κατασκευή, το υλικό κατασκευής των χορδών (νάιλον, μεταλλικές) , το μέγεθος του ηχείου της κιθάρας, τον τρόπο παιχνιδιού του κιθαρίστα κ.α.

Κατά τη σύνθεση ενός ήχου με ηλεκτρονικά μέσα, οι παράμετροι πληθαίνουν, καθώς το σύστημα των εμπλεκόμενων μερών στην δημιουργία ενός ήχου είναι σαφώς πιο σύνθετο. Το τελικό ηχόχρωμα εξαρτάται από το λογισμικό το οποίο χρησιμοποιείται για την παραμετροποίηση, τις πηγές παραγωγής ήχου που εφαρμόζονται (είδη ταλαντωτών), τις μεθόδους μορφοποίησης του ήχου και το υλικό το οποίο χρησιμοποιείται για την κατασκευή του (D/A converters, analog –digital amplification κ.α.)(4).

Η παρούσα μελέτη εστιάζει στους ταλαντωτές ηλεκτρονικών συνθετητών και έχει ως στόχο την μοντελοποίηση της μεθόδου παραγωγής ήχου sample and synthesis (S&S), που θα περιγραφεί παρακάτω. Δεν θα γίνει εκτενής περιγραφή των υπόλοιπων μερών του συνθετητή ως σύστημα, αλλά θα διερευνηθούν σε βάθος οι ποιότητες που χαρακτηρίζουν τους ταλαντωτές γενικότερα και ειδικότερα θα αναλυθεί το λογισμικό “Modeled Linear Arithmetic Synthesis” το οποίο και δημιουργήθηκε στα πλαίσια αυτής της εργασίας.

Στην παρακάτω εικόνα περιλαμβάνονται τα βασικά μέρη ενός ηλεκτρονικού συνθετητή και ο τρόπος συσχέτισης μεταξύ τους:



Η βασική δομή ενός συνθετητή.

Στην εικόνα περιγράφεται επίσης η διαδρομή που ακολουθεί ένα σήμα από την πηγή του (ταλαντωτής) έως και την έξοδό του από το σύστημα. Ως είσοδο δίνεται

μία πληροφορία (συνήθως midi στα σύγχρονα συστήματα) από ένα πληκτροφόρο που μεταξύ άλλων ορίζει την συχνότητα ταλάντωσης του ταλαντωτή. Εδώ απεικονίζεται η αφαιρετική μέθοδος σύνθεσης (subtractive) κατά την οποία κυματομορφές πλούσιες σε συχνοτικό περιεχόμενο, μορφοποιούνται με τη χρήση φίλτρων και περιβαλλουσών. Ωστόσο οποιαδήποτε μέθοδος σύνθεσης έχει ως αρχή της ροής του σήματος έναν ή περισσότερους ταλαντωτές. Η διαφοροποίηση βρίσκεται στις διαθέσιμες μεθόδους μορφοποίησης των ταλαντωτών.

1.2: Τα είδη των ταλαντωτών στους συνθετητές.

Ηλεκτρονικός ταλαντωτής, ονομάζεται το ηλεκτρονικό κύκλωμα, το οποίο παράγει ένα περιοδικό σήμα που είναι συνήθως ημιτονοειδές ή τετραγωνικό(8). Οι ταλαντωτές μετατρέπουν το συνεχές ρεύμα από μία παροχή ρεύματος σε εναλλασσόμενο.

Στους πρώτους συνθετητές, οι ταλαντωτές ήταν Voltage Controlled Oscillators. Η συχνότητα τους, μεταβάλλονταν ανάλογα με την τάση του ρεύματος. Αυτό πρακτικά σήμαινε ότι κάθε νότα αντιστοιχούσε σε διαφορετική τάση ηλεκτρικού ρεύματος. VCO συναντάμε σε μοντέρνους συνθετητές σπανιότερα, λόγω του υψηλού κόστους κατασκευής. Μετέπειτα τη δεκαετία του '80, λόγω της ασταθούς συχνότητας ταλάντωσης των συγκεκριμένων ταλαντωτών, προστέθηκε και ένα ψηφιακό κύκλωμα το οποίο φρόντιζε να κρατάει την συχνότητα σταθερή (Digital Controlled Oscillators).

Λόγοι όπως η μείωση του κόστους αλλά και οι διαθέσιμες κυματομορφές που μπορεί να παράγει ένα ηλεκτρονικό κύκλωμα, οδήγησαν στην αντικατάσταση των ταλαντωτών. Έτσι σε έναν ψηφιακό συνθετητή, ταλαντωτής μπορεί να είναι κάτι από τα ακόλουθα(9):

- Μία κυματομορφή που παράγεται σε πραγματικό χρόνο από έναν υλοποιημένο αλγόριθμο με την χρήση DSP ή CPU.
- Wavetable Synthesis. Δείγματα μιας περιόδου και συχνοτικά σχετιζόμενα μεταξύ τους αποθηκεύονται σε έναν πίνακα. Επιλέγεται μια κυματομορφή

από τον πίνακα για αναπαραγωγή, ή διατρέχεται ο πίνακας πραγματοποιώντας σειριακή αναπαραγωγή των αποθηκευμένων κυματομορφών.

- Αποθηκευμένα δείγματα ήχου σε μνήμη ram ή rom.
- Μία κυματομορφή που παράγεται ψηφιακά σε πραγματικό χρόνο με Frequency Modulation ή Phase Modulation.

Η αυξημένη επεξεργαστική ισχύς των ηλεκτρονικών υπολογιστών, έδωσε την δυνατότητα υλοποίησης αλγορίθμων χωρίς τη χρήση DSP. Έτσι ένα μεγάλο μέρος της αγοράς των συνθετητών είναι εικονικοί συνθετητές. Αυτοί είναι πρακτικά, λογισμικά που προσομοιώνουν όλες τις λειτουργίες ενός συνθετητή και μέσω γραφικής διεπαφής μπορεί κάποιος να μορφοποιεί τους ήχους, όπως αντίστοιχα θα έκανε με τα κουμπιά και τα ποτενσιόμετρα ενός πραγματικού συνθετητή. Στους εικονικούς συνθετητές συναντάμε πολλές φορές όλα τα παραπάνω είδη ταλαντωτών και τη δυνατότητα αλληλεπίδρασης μεταξύ τους.

1.3: Η χρήση αποθηκευμένων δειγμάτων στους συνθετητές και οι ιδιότητές τους ως ταλαντωτές.

Η δειγματοληψία ή sampling, είναι μία μέθοδος ψηφιακής κωδικοποίησης ενός αναλογικού σήματος. Η περισσότερο διαδεδομένη μέθοδος ονομάζεται Pulse Code Modulation ή PCM. Κατά τη διαδικασία της μεθόδου, λαμβάνονται ανά τακτά χρονικά διαστήματα δείγματα των εντάσεων ενός αναλογικού σήματος και ανάλογα με την ανάλυση της ψηφιοποίησης που δίνεται, κβαντίζονται τα δείγματα στις πλησιέστερες δυνατές τιμές (5).

Η ποιότητα του ήχου των δειγμάτων που λαμβάνονται εξαρτάται από δύο παράγοντες: Την συχνότητα που λαμβάνονται τα δείγματα και ο αριθμός των διαθέσιμων τιμών στις οποίες κβαντίζονται τα δείγματα (bit depth). Επί παραδείγματι, η ποιότητα ήχου ενός CD αντιστοιχεί σε 44.100 δείγματα το δευτερόλεπτο και 16 bit ανάλυση ή 65.536 διαθέσιμες τιμές κβάντισης.

Η μέθοδος PCM, η οποία εφευρέθηκε τη δεκαετία του '50 (6), χρησιμοποιείται στις διάφορες παραλλαγές της μέχρι και σήμερα για την ψηφιοποίηση ήχων. Μία από τις παραλλαγές τις, η LPCM χρησιμοποιείται για την κωδικοποίηση των αρχείων «wave» (.wav). Τα αρχεία τύπου wave αποτελούν και το κυριότερο είδος αρχείου που αναγνωρίζεται από τους συνθετητές που χρησιμοποιούν δείγματα.

Μετά τη διαδικασία της δειγματοληψίας, ακολουθεί η αποθήκευση του δείγματος. Αυτή μπορεί να γίνει απευθείας στην μνήμη του δειγματολήπτη (sampler), ή σε κάποια μονάδα αφαιρούμενου αποθηκευτικού χώρου (flash, sd, floppy κ.α.). Ύστερα ακολουθεί η διαδικασία του key mapping, όπου αντιστοιχίζεται βάση της συχνότητας του ήχου που έγινε δειγματοληψία το δείγμα σε συγκεκριμένο πλήκτρο του κλαβιέ. Στη συνέχεια, αναλόγως της έκτασης των δυνατοτήτων του συνθετητή, είτε ακολουθείται η ίδια διαδικασία με δείγματα διαφορετικής συχνότητας, είτε το δείγμα που επιλέχθηκε, αναπαράγεται με τη χρήση κάποιου αλγορίθμου σε διαφορετική συχνότητα, αναλόγως του πλήκτρου του κλαβιέ.

Αξίζει να σημειωθεί ότι, στους πρώτους συνθετητές που χρησιμοποιούνταν δείγματα, η αναπαραγωγή σε διαφορετική συχνότητα, σήμαινε απλώς ο πολλαπλασιασμός της αρχικής συχνότητας με έναν συντελεστή. Αυτό επέφερε μία σειρά προβλημάτων. Το πρώτο και πιο σημαντικό, ήταν ο χρόνος που διαρκούσε το τροποποιημένο δείγμα. Μεγαλύτερη συχνότητα σήμαινε μικρότερος χρόνος διάρκειας αναπαραγωγής και το αντίστροφο. Επιπροσθέτως, με την αυξομείωση της συχνότητας του δείγματος, μεταβάλλονταν και συχνότητες οι οποίες παράγονται κατά την συνήχηση μερών ενός μουσικού οργάνου. Αυτές είναι γνωστές και ως Formants και έχουν το χαρακτηριστικό ότι είναι σχετικά αμετάβλητες, όσο η θεμελιώδης συχνότητα του μουσικού οργάνου μεταβάλλεται. Αυτό προκαλεί μια αισθητή αλλοίωση στο τελικό ηχόχρωμα και θεωρείται ως ένας από τους βασικούς λόγους έλλειψης φυσικότητας ενός ήχου.

Τα ανωτέρω προβλήματα λύθηκαν με ποικίλους τρόπους. Αρχικά εφαρμόζοντας time stretching αλγορίθμους, διορθώνοντας έτσι τον χρόνο διάρκειας της αναπαραγωγής. Με τη χρήση αναπαραγωγής σε βρόγχο, εξασφαλίστηκε ότι μία νότα μπορεί να έχει παραπάνω διάρκεια όπου χρειάζεται, ασχέτως με τον χρόνο

διάρκειας του δείγματος. Αργότερα, εταιρίες όπως η Roland, παρουσίασαν συνθετητές που με την χρήση formant shifting, μπορούσε κάποιος να κάνει χρήση ενός και μόνο δείγματος σε μεγάλη συχνοτική έκταση χωρίς αλλοίωση στο ηχόχρωμα.

Με αύξηση της διαθέσιμης μνήμης στους συνθετητές, δόθηκε και η δυνατότητα αποθήκευσης και αναπαραγωγής δειγμάτων που αντιστοιχούσαν σε κάθε νότα του κλαβιέ. Δόθηκε επίσης ευχέρεια αποθήκευσης μεγάλων σε διάρκεια δειγμάτων έτσι ώστε η ανάγκη αναπαραγωγής σε βρόγχο να ελαχιστοποιηθεί κατά την χρήση του.

Δύο από τους πρώτους συνθετητές, που χρησιμοποιήθηκαν εκτενώς στην παραγωγή εμπορικής μουσικής τη δεκαετία του '80 ήταν οι CMI Fairlight και EMU Emulator. Ξεκίνησαν να παράγονται στις αρχές της δεκαετίας του 80 και στις αρχικές τους εκδόσεις, αναγνώριζαν δείγματα 8bit ανάλυσης και είχαν δυνατότητα αναπαραγωγής το μεν Fairlight στα 10 khz (7), το δε Emulator στα 27 khz. Λόγω της περιορισμένης μνήμης ram (208 KB και 128 KB αντίστοιχα), τα δείγματα διαρκούσαν 1-2 δευτερόλεπτα. Η ικανότητα αναπαραγωγής των δειγμάτων σε βρόγχο, έπαιξε μεγάλο ρόλο στην δυνατότητα χρήσης τους σε συχνοτικό εύρος μεγαλύτερο από αυτό της συχνότητας του ήχου στην οποία έγινε η δειγματοληψία. Οι περιορισμοί που είχαν δεν αποτέλεσαν ωστόσο εμπόδιο στην πολύ επιτυχημένη εμπορική τους πορεία.

1.4: Η χρήση δειγμάτων κατά την μέθοδο παραγωγής ήχου “Sample and Synthesis”.

Οι ήχοι έχουν διάφορα στάδια ανάπτυξης στον χρόνο. Αυτά κατά προσέγγιση, μπορούν να θεωρηθούν τα 4 ακόλουθα: Το στάδιο έναρξης του ήχου γνωστό και ως attack, το στάδιο decay, που μεσολαβεί από την έναρξη μέχρι το κομμάτι του ήχου που παραμένει σχετικά αμετάβλητο στο χρόνο (Steady State ή Sustain) και το τέλος του ήχου (release).

Ο Βαθμός συμμετοχής του κάθε μέρους στον συνολικό χρόνο που διαρκεί ο ήχος, ανήκει στα συμπεριφορικά γνωρίσματα του κάθε μουσικού οργάνου. Για παράδειγμα τα έγχορδα όργανα έχουν συνήθως μεγάλο attack σε σύγκριση με τα κρουστά, στα οποία είναι πολύ μικρό.

Οι συνθετητές που έκαναν χρήση δειγμάτων, είχαν ένα μεγάλο μειονέκτημα. Την υψηλή τιμή. Και αυτό γιατί η μνήμη ram, που είναι απαραίτητη για το φόρτωμα δειγμάτων στον συνθετητή ήταν πολύ ακριβή.

Η πρώτη που έδωσε την ευκαιρία στο ευρύ κοινό να έχουν πρόσβαση σε ήχους μεγαλύτερου ρεαλισμού με μικρό κόστος, ήταν η Roland. Βασιζόμενη στη θεωρία της ψυχοακουστικής που θα αναλυθεί σε επόμενο κεφάλαιο, ότι η περισσότερη πληροφόρηση για την ταυτότητα ενός οργάνου βρίσκεται στο attack transient μέρος ενός ήχου, αποθήκευσε σε μνήμη rom attack transients μουσικών οργάνων. Το υπολειπόμενο μέρος ενός ήχου μπορούσε να προσομοιωθεί με η χρήση ταλαντωτών που παρήγαγαν αλγοριθμικά τις βασικές για έναν συνθετητή κυματομορφές (square, sawtooth) και με την χρήση αποθηκευμένων steady state δειγμάτων. Η μορφοποίηση των αλγοριθμικά παραγόμενων κυματομορφών γίνονταν μέσω αφαιρετικής σύνθεσης, δηλαδή απλούστευση του συχνοτικού περιεχομένου τους μέσω φίλτρων.

Ο Roland D-50 θεωρείται ο πρωτοπόρος S&S συνθετητής. Η μνήμη rom είχε μέγεθος 128 KB και είχε αποθηκευμένα δείγματα ανάλυσης 8bit και συχνοτικού εύρους 32 KHz. Τα δείγματα είχαν διάρκεια ενός δευτερολέπτου και δεν μπορούσαν να χρησιμοποιηθούν φίλτρα για την μορφοποίησή τους. Εκτός από attack transients, ανάμεσα στα δείγματα υπήρχαν και ήχοι που μπορούσαν να παιχτούν σε βρόγχο. Τα δείγματα μπορούσαν να αλληλεπιδράσουν με τις αλγοριθμικά παραγόμενες κυματομορφές μέσω ring modulation. Η μέθοδος σύνθεσης ονομάστηκε Linear Arithmetic Synthesis. Στην πορεία και άλλες εταιρίες παρουσίασαν παραλλαγές στην μέθοδο, όπως η Korg με την Advanced Integrated Synthesis στο M1 και η Ensoniq με την Cross Wave Synthesis.

Παραλλαγές της μεθόδου χρησιμοποιήθηκαν για πάνω από δέκα χρόνια σε συνθετητές, έως ότου η μνήμη ήταν πια φτηνή και μπορούσαν να αποθηκευτούν ολόκληρα δείγματα με μικρό κόστος στη μνήμη rom. Ωστόσο, πολλοί ήχοι που

δημιουργήθηκαν σε αυτό το διάστημα θεωρούνται μοναδικοί, και ήταν αποτέλεσμα της δημιουργικής συνένωσης attack transients και κυματομορφών από τελείως διαφορετικές πηγές προέλευσης ήχων. Ο ήχος Fantasia του D50, του οποίου παραλλαγές βρίσκονται σε όλους τους σύγχρονους συνθετητές, αποτελείται από ένα attack transient από δείγμα κύμβαλου, έναν βρόγχο από δείγμα glass harp και δύο πριονωτές κυματομορφές. Μέσω της μορφοποίησης των ανωτέρω τμημάτων, ο χρήστης δύσκολα μπορεί να καθορίσει τους επιμέρους ήχους χωρίς να χρησιμοποιήσει τη διεπαφή του συνθετητή και να δει τι αντιστοιχεί σε κάθε ταλαντωτή.

Μία μεγάλη μερίδα σύγχρονων συνθετητών χρησιμοποιεί δείγματα ως βάση για τον ήχο που παράγουν, αποθηκευμένα σε μνήμη rom. Οι ήχοι πιάνου, για παράδειγμα εξακολουθούν να είναι στην πλειοψηφία τους αποθηκευμένα δείγματα, πολύ μεγάλης διάρκειας. Πολλές φορές σε μία νότα αντιστοιχούν παραπάνω από ένα δείγματα, ηχογραφημένα σε διαφορετικές δυναμικές. Έτσι σε μεγαλύτερης έντασης πάτημα ενός πλήκτρου του κλαβιέ, αντιστοιχεί άλλο δείγμα από αυτό που θα αναπαράγονταν σε μικρής έντασης πάτημα. Η μέθοδος αυτή, αν και χρησιμοποιεί δείγματα ως ταλαντωτές, ξεφεύγει από την έννοια του S&S και θα μπορούσε να προσεγγιστεί ως sample playback.

Κεφάλαιο 2: Ο ρόλος του Attack Transient στην αντίληψη των ήχων.

2.1: Ο Ορισμός του transient.

Ως transient ορίζεται το κομμάτι εκείνο ενός σήματος ήχου που αναπτύσσεται γρήγορα και απρόβλεπτα μέσα σε ένα μικρό χρονικό διάστημα (11). Τα transients χωρίζονται στις εξής υποκατηγορίες:

- Attack transients. Εντοπίζονται κυρίως στην έναρξη μιας νότας και μπορεί να είναι απότομες μεταβολές της έντασης, της φάσης και του συχνοτικού περιεχόμενου.
- Fast decay transients που αφορούν απότομες πτωτικές μεταβολές της έντασης.
- Fast transitions (π.χ. πορταμέντο, βιμπράτο)
- Θόρυβος (π.χ. ο ήχος της βροχής)

Τα ανωτέρω συμβάντα μπορεί να συμβούν κατά τη διάρκεια ενός σήματος και παράλληλα στην περίπτωση που ο ήχος είναι πολυφωνικός. Αυτό σημαίνει ότι έχουμε πολλές διαφορετικές πηγές ήχων εγγεγραμμένες ταυτόχρονα. Τα ανωτέρω συμβάντα θα μπορούσαν να πραγματοποιούνται ταυτόχρονα όταν για παράδειγμα το σήμα μας είναι μια ηχογραφημένη συναυλία που πραγματοποιήθηκε σε εξωτερικό χώρο.

2.2: Ανάλυση του attack transient ως συμβάν.

Από όλα τα παραπάνω είδη των transients, αυτό του Attack θεωρείται ότι είναι το σημαντικότερο όσο αναφορά την πληροφορία που περιέχεται στο

συγκεκριμένο τμήμα της νότας και αφορά την ταυτότητα του μουσικού οργάνου. (12), (13) Αυτό θεωρητικά σημαίνει ότι το ποιο μουσικό όργανο ακούγεται, είναι σε ένα μεγάλο βαθμό εγγεγραμμένο στο attack transient τμήμα του ήχου. Μία εξήγηση σε αυτό θα μπορούσε να είναι ότι η αντιληπτική ικανότητα του ανθρώπου να αναγνωρίζει την προέλευση των ήχων διαμορφώθηκε με τέτοιο τρόπο έτσι ώστε να κάνει την κατηγοριοποίηση βάση του μέρους του ήχου που θα διακρίνονταν περισσότερο ανάμεσα σε όλους τους ήχους της φύσης. Για αυτό και μέσα σε ένα συνονθύλευμα ήχων, μεγαλύτερη πιθανότητα έχουμε να διακρίνουμε τις απότομες μεταβολές που συμβαίνουν στο attack transient μέρος ενός ήχου, παρά στο περιοδικό του μέρος όπου το ηχόχρωμα αναπόφευκτα θα αλλοιωθεί από τους υπόλοιπους ήχους του περιβάλλοντος.

Όπως αναλύθηκε στο προηγούμενο κεφάλαιο, υπάρχουν 4 στάδια ανάπτυξης μιας νότας στον χρόνο. Από αυτά το decay και sustain κομμάτια, έχουν μελετηθεί ενδελεχώς λόγω της σχετικής περιοδικότητας που εμφανίζουν στον χρόνο. (14) Η ενότητα των 2 αυτών διαδοχικών κομματιών είναι γνωστή και ως steady-state και μελετάται ως ένα ενιαίο κομμάτι. Σε κάποιες μελέτες, στο steady-state συμπεριλαμβάνεται και το release κομμάτι της νότας.(15) Για της ανάγκες της παρούσας εργασίας και του λογισμικού που δημιουργήθηκε στα πλαίσια αυτής, ως steady state κομμάτι, θα θεωρείται το υπόλοιπο μέρος του ήχου, εκτός του attack transient.

Το attack transient λόγω της έλλειψης περιοδικότητας του τμήματος αλλά και των ασαφών ορίων του συμβάντος, κυρίως στο πού τελειώνει, αποτελεί ένα σχετικά ανεξερεύνητο πεδίο. Ως ένας άτυπος κανόνας, τα attack transients διαρκούν περίπου 50 ms και μπορεί να φτάσουν και τα 100 ms. Μέσα σε αυτό το χρονικό διάστημα παρουσιάζονται απότομες μεταβολές τόσο στην ένταση όσο και στην φάση διαφόρων συχνοτήτων. Τα transients μπορούν να κατηγοριοποιηθούν ως αδύναμα ή ισχυρά, βασιζόμενη στην ισχύ της περιβάλλουσας. Επίσης μπορούν να χαρακτηρισθούν ως γρήγορα ή αργά ανάλογα με τον ρυθμό που μεταβάλλεται η ένταση της περιβάλλουσας. Γρήγορα transient θεωρούνται αυτά που έχουν περιβάλλουσες που είναι λεπτόκυρτες ενώ οι περιβάλλουσες που είναι πλατύκυρτες αντιστοιχούν σε αργά transients.(16)

Αξίζει να γίνει αναφορά σε μερικές περιπτώσεις στις οποίες το attack transient ερευνήθηκε ως συμβάν:

- Keeler το 1972 (17), σε μία μελέτη όπου είχε στόχο την εύρεση στοιχείων που διαφοροποιούν τα πνευστά, χρησιμοποίησε τη διάρκεια του attack transient, ως ένα από τα συστατικά στοιχεία διαφοροποίησης των οργάνων.
- Ο John William Gordon στη μελέτη που δημοσίευσε το 1984 “Perception of Attack Transients In Musical Tones”, ερευνά μέσα από μία σειρά πειραμάτων τη σημαντικότητα του αντιληπτού χρόνου του attack transient ως συμβάν. Διαπιστώνει ότι, η διάρκεια του Attack επηρεάζει την χρονική στιγμή που γίνεται αντιληπτό από τον ακροατή. Σημαντικό ρόλο παίζει κατά συνέπεια και στον συντονισμό διαφορετικών οργάνων σε μία ορχήστρα, καθώς και στη αίσθηση του ρυθμού. Μουσικά όργανα που παράγουν ήχους με attack transients μικρότερης διάρκειας αποδίδουν καλύτερα τον ρυθμό σε ένα μουσικό κομμάτι, ενώ όργανα με μεγαλύτερης διάρκειας attack, δίνουν την αίσθηση στον ακροατή ότι «καθυστερούν».
- Σε μία άλλη μελέτη του ο Ralph C. Thayer Jr. Το 1974 (18) προσπάθησε να αποδείξει ότι δεν είναι μόνο το συχνотικό περιεχόμενο αυτό που καθορίζει την ταυτότητα του οργάνου και να ενισχύσει την άποψη που υποστηρίζει την σχέση ταυτότητας οργάνου- attack transient. Αν και η μελέτη διεξήχθη με πολλούς περιορισμούς (3 συχνότητες από κάθε όργανα και μόλις 4 πνευστά μουσικά όργανα) και δε θα μπορούσε να οδηγήσει σε γενικεύσεις, εντούτοις τα αποτελέσματα παρουσιάζουν ιδιαίτερο ενδιαφέρον. Στη μελέτη διαχωρίστηκαν τα attack transients και τα steady states και συγκολλήθηκαν διαφορετικά μέρη από διαφορετικά όργανα. Οι ακροατές που άκουσαν τα τροποποιημένα δείγματα ήχου, επηρεάστηκαν εμφανώς στην αναγνώριση των steady states των δειγμάτων που άκουγαν.
- Οι E. L. Saldanha και John F. Corso (19) αναπαρήγαγαν σε ακροατές 3 διαφορετικές συχνότητες από 10 μουσικά όργανα. Αποκόπτοντας το attack transient από τους ήχους, η αναγνώριση των μουσικών οργάνων από τους ακροατές μειώθηκε από το 47% στο 32%. Ανέφεραν επίσης την σημαντικότητα της διάρκειας του attack transient στον συγχρονισμό μιας

ορχήστρας (όσο μικρότερη η διάρκεια, τόσο δυσκολότερος ο συντονισμός) και την αλλαγή του γενικού ηχοχρώματος ενός κομματιού, όταν αλλάξει ο τρόπος παιξίματος ενός οργάνου. Στον στακάτο τρόπο παιξίματος όπου την διάρκεια της νότας καταλαμβάνει το attack, τα attack transients καθορίζουν το συχνοτικό περιεχόμενο, ενώ στο λεγκάτο που το attack συνήθως είναι απών, τα steady states καθορίζουν το συχνοτικό περιεχόμενο.

- Οι Miller και Carterette (20) πειραματίστηκαν με συνθετικούς ήχους και με αυτόν τον τρόπο τροποποίησαν παραμέτρους, όπως το συχνοτικό περιεχόμενό τους, το σχήμα της περιβάλλουσας του attack και την χρονική στιγμή εμφάνισης συγκεκριμένων αρμονικών. Η μελέτη επιβεβαιώνει ότι το attack transient αποτελεί σημαντική παράμετρο στον καθορισμό του ηχοχρώματος ενός ήχου.

2.3: Μέθοδοι εντοπισμού και απομόνωσης του attack transient σε ένα σήμα.

Αν και δεν είναι σαφώς ορισμένα τα όρια αρχής και τέλους ενός attack transient από κάποιον κανόνα, εντούτοις υπάρχει μία πληθώρα μεθόδων διαχωρισμού ενός attack transient από το steady state. Αυτές οι μέθοδοι θα μπορούσαν να κατηγοριοποιηθούν σε τρεις κλάσεις, βάση της δομής των εξόδων που δίνουν (21):

Η πρώτη κλάση μεθόδων βασίζεται στη γραμμική πρόβλεψη. Αποδομείται ο ήχος σε ένα σήμα διέγερσης και ένα συνηχούμενο φίλτρο. Συγκεκριμένα, οι συναρτήσεις των Yule-Walker επιτρέπουν την εκτίμηση της βέλτιστης τάξης του φίλτρου και ελαχιστοποιούν το σφάλμα της πρόβλεψης. Όταν γίνει η εκτίμηση του φίλτρου, το σήμα διέγερσης είναι το αποτέλεσμα του φιλτραρίσματος του σήματος με το αντίστροφο φίλτρο. Εάν έχει γίνει σωστή εκτίμηση του φίλτρου, τότε η περισσότερη από την ενέργεια του σήματος διέγερσής εντοπίζεται στην περιοχή όπου βρίσκεται το attack transient του αρχικού σήματος.

Τα βέλτιστα αποτελέσματα αυτής της μεθόδου δίνονται συνήθως όταν τα σήματα είναι αποτέλεσμα μιας μεταβλητής στο χρόνο διέγερσης, και είναι φιλτραρισμένα και ενισχυμένα από έναν ενισχυτή. Εντούτοις, υπάρχουν σημαντικοί περιορισμοί. Ο πρώτος αφορά την τάξη του φίλτρου που θα πρέπει να έχει εκτιμηθεί από πριν, γεγονός αρκετά δύσκολο να πραγματοποιηθεί. Επιπλέον, αυτή η μέθοδος είναι αρκετά αναποτελεσματική στην ανίχνευση attack transient σε διαδοχικές νότες.

Στη δεύτερη κλάση γίνεται έμμεσος εντοπισμός του attack transient με την απομόνωση του Steady State κομματιού, το οποίο και λόγω της περιοδικότητας χαρακτηρίζεται ως το τονικό μέρος του ήχου. Εάν η απομόνωση είναι επιτυχής, στο εναπομείναν κομμάτι που όπως και στην προηγούμενη κλάση, θα παρατηρείται μια απότομη «έκρηξη» ενέργειας και θα ορίζεται ως το attack transient. Ένας τρόπος επίτευξης είναι με την τμηματοποίηση του Short Time Fourier Transform. Πραγματοποιείται αρχικά ανάλυση του STFT. Βάση της ανάλυσης του STFT, γίνεται αναζήτηση σε κάθε frequency bin που μας παρέχεται, αναζητώντας την παρουσία ενός attack transient. Η αναγνώρισή του προκύπτει από την πρόβλεψη της φάσης σε κάθε frequency bin k , όπου στα steady states αναπτύσσεται γραμμικά στον χρόνο. Η πρόβλεψη της φάσης ϕ_n δίνεται από τον παρακάτω τύπο:

$$\phi_n^{pred} = 2\phi_{n-1} - \phi_{n-2}$$

Ύστερα η ϕ_n^{pred} συγκρίνεται με την ϕ_n και αν η διαφορά τους είναι μικρότερη από μια παράμετρο ϵ τότε διαπιστώνεται παρουσία steady state. Εναλλακτικά διαπιστώνεται παρουσία attack transient.

Στην Τρίτη κλάση στόχος είναι η αποδόμηση του τμήματος σε τρία συστατικά στοιχεία: Στο περιοδικό (το οποίο εκφράζεται και ως άθροισμα ημιτόνων) στο transient και σε θόρυβο. Η διαδικασία αυτή γίνεται είτε τμηματικά είτε ταυτόχρονα, με την ταυτόχρονη εξαγωγή να έχει αναπόφευκτα μεγαλύτερο κόστος υπολογιστικής ισχύος. Η κλάση μεθόδων ονομάζεται και STN (Sines+Transient+Noise).

Ο υλοποιημένος αλγόριθμος εμπίπτει έμμεσα στην δεύτερη κατηγορία, με την διαφοροποίηση ότι ανιχνεύεται η «έκρηξη» ενέργειας του attack transient μέσω του spectral magnitude και απομονώνεται από το περιοδικό κομμάτι (steady state).

Λεπτομερέστερη ανάλυση του αλγορίθμου θα πραγματοποιηθεί στο επόμενο κεφάλαιο.

Κεφάλαιο 3: Περιγραφή του αλγορίθμου ανάλυσης και εξαγωγής του Attack Transient.

3.1: Περιορισμοί λειτουργίας και προεπεξεργασία σήματος εισόδου.

Για την εκπόνηση της παρούσας μελέτης υλοποιήθηκε ο αλγόριθμος που περιγράφεται στη δημοσίευση «A TRANSIENT DETECTION ALGORITHM FOR AUDIO USING ITERATIVE ANALYSIS OF STFT» (22). Η επιλογή του αλγορίθμου έγινε με τα κριτήρια ότι θα μπορεί να λειτουργεί χωρίς επιπλέον παραμετροποίηση από το χρήστη κατά την εφαρμογή του και θα έχει χαμηλές απαιτήσεις σε μνήμη και υπολογιστική ισχύ για τη λειτουργία του.

Ο αλγόριθμος έχει τους εξής σημαντικούς περιορισμούς: 1) Λειτουργεί μόνο σε μονοφωνικά σήματα και 2) είναι λειτουργικός σε attack transients που αναπτύσσονται με σχετικά μεγάλη ταχύτητα (δηλαδή είναι μικρής διάρκειας) και εμφανίζουν σημαντικές μεταβολές στις εντάσεις στο φασματογράφημα. Αυτό σημαίνει ότι θα είναι σε θέση να ανιχνεύσει attack transients ήχων κιθάρας, πιάνου, κύμβαλων κλπ., αλλά δεν θα είναι αποτελεσματικός στο να ανιχνεύσει το attack ενός gong ή ενός βιολιού που θα παράγει ήχο με τρόπο παιξίματος “arco” όπου το attack δεν είναι κρουστικής μορφής.

Πριν την έναρξη της εφαρμογής του αλγορίθμου, πραγματοποιούμε μία σειρά από βήματα επεξεργασίας του σήματος (έστω σήμα εισόδου x) στο οποίο θα εφαρμοστεί ο αλγόριθμος. Αυτά είναι τα εξής ακόλουθα:

- Α) Επαναδειματοληπούμε το δείγμα στα 16.000 Hz για να διορθώσουμε τυχόν μεταβλητό ρυθμό δειγματοληψίας,
- β) Κανονικοποιούμε το δείγμα μας με τον παρακάτω τρόπο:

$$x_{norm}[n] = \frac{x[n]}{\max(|x[n]|)}$$

γ) Πραγματοποιούμε short time Fourier transform χρησιμοποιώντας πλαίσια διάρκειας 44ms και επικάλυψη 33ms. Κάθε πλαίσιο πολλαπλασιάζεται με ένα παράθυρο της μορφής Blackmann- Harris μήκους $N=700$ δειγμάτων. Οι τιμές που δίνεται στον αλγόριθμο είναι 40ms, 30ms και 640 αντίστοιχα, εντούτοις μετά από πειραματισμούς επιλέχθηκαν οι ανωτέρω τιμές. Το μετασχηματισμένο δείγμα ήχου δίνεται από τον παρακάτω τύπο για I δείγματα και k frequency bins:

$$X(i, k) = \sum_{n=0}^{N-1} x[n] \cdot w[n - iR] \cdot e^{-j \cdot 2\pi \cdot n \cdot k / N}$$

όπου $w[n]$ είναι η συνάρτηση παραθύρωσης, N είναι ο αριθμός των δειγμάτων του παραθύρου και R είναι η μετακύλιση σε δείγματα.

Αυτά είναι και τα πρώτα βήματα ανάλυσης του σήματος. Η τελευταία συνάρτηση, αποτελεί ουσιαστικά και την είσοδο στον αλγόριθμο όπως θα δούμε παρακάτω. Οι επόμενες συναρτήσεις περιέχονται στην επαναληπτική διαδικασία που ακολουθείται για την εξαγωγή του attack transient.

3.2: Η επαναληπτική διαδικασία εξαγωγής του Attack transient.

Τα βήματα της επαναληπτικής διαδικασίας, τροποποιήθηκαν για την παρούσα μελέτη, έτσι ώστε να συμφωνούν με την προγραμματιστική υλοποίηση του αλγορίθμου.

Βήμα 1): Βάση του STFT που υπολογίστηκε στην προηγούμενη παράγραφο, υπολογίζονται και οι παρακάτω συναρτήσεις T_- και T_+ που εντοπίζουν κάθετες ακμές στο φασματογράφημα:

$$\begin{aligned} T_-(i, k) &= |X(i, k)| - |X(i-1, k)|, \\ T_+(i, k) &= |X(i, k)| - |X(i+1, k)|. \end{aligned}$$

Αυτές οι κάθετες ακμές σηματοδοτούν την έναρξη (onset) και το τέλος (offset) ενός attack transient. Οι συναρτήσεις T_- και T_+ πρέπει να έχουν υψηλές τιμές σε πλαίσια που αντιστοιχούν σε transients.

Βήμα 2): Υπολογίζουμε την $F(i, j)$ που συνδυάζει τις T_- και T_+ συναρτήσεις και πραγματοποιεί half wave rectification, μία διαδικασία που διασφαλίζει ότι θα εντοπιστούν μονάχα έναρξη από την T_- και τέλος από την T_+ αντίστοιχα. Η παράμετρος ν έχει ως στόχο να συνυπολογίσει το συχνοτικό “άπλωμα” του attack transient, απορρίπτοντας ενθόρυβες μεταβολές του φασματογραφήματος:

$$\begin{aligned} F(i, j) &= 0.5 \left\{ \sum_{k=j-\nu}^{j+\nu} \{1 + \text{sgn}(T_-(i, k))\} \cdot T_-(i, k) \right. \\ &\quad \left. + \{1 + \text{sgn}(T_+(i, k))\} \cdot T_+(i, k) \right\}, \end{aligned}$$

Όπου

$$\text{sgn}(\theta) = \begin{cases} 1 & \text{if } \theta \geq 0, \\ -1 & \text{if } \theta < 0. \end{cases}$$

Βήμα 3) Υπολογίζουμε το adaptive threshold $\lambda(i, j)$, ένα κατώφλι που μεταβάλλεται βάση του χρόνου για κάθε time-frequency bin στη συνάρτηση $F(i, j)$:

$$\lambda(i, j) = \beta \times \frac{\sum_{l=i-\tau}^{i+\tau} F(l, j)}{2\tau + 1},$$

Όπου β είναι μια παράμετρος που ελέγχει την ένταση του transient που πρέπει να εξαχθεί από το σήμα. Εάν σε κάθε frequency bin της συνάρτησης F η τιμή της είναι μεγαλύτερη από το κατώφλι, τότε στη συνάρτηση $\Gamma(i,j)$ δίνεται η τιμή «1» :

$$\Gamma(i, j) = \begin{cases} 1 & \text{if } F(i, j) > \lambda(i, j), \\ 0 & \text{if } F(i, j) \leq \lambda(i, j). \end{cases}$$

Αθροίζοντας τα frequency bins της Γ ενός πλαισίου, βρίσκουμε το σύνολο των frequency bins που έχουν περισσότερη ενέργεια από τους γείτονές τους. Αυτό μπορεί να φανερώσει εάν υπάρχει ή όχι attack transient:

$$\Sigma_{\Gamma}(i) = \sum_{j=0}^{N-1} \Gamma(i, j).$$

Βήμα 4) Αν το παραπάνω άθροισμα είναι μεγαλύτερο από ένα δεύτερο κατώφλι που ονομάζεται λ_{Thr} , τότε το αντίστοιχο frame θεωρείται ως transient frame και ένα μικρό κομμάτι δ από τη συνάρτηση του μετασχηματισμού Fourier που υπολογίσαμε αποκόπτεται από το frame και προστίθεται στη συνάρτηση P για να χτιστεί το transient:

$$P(i, j) = \begin{cases} P(i, j) & \text{if } \Sigma_{\Gamma} < \lambda_{Thr}, \\ P(i, j) + \delta \cdot X(i, j) & \text{if } \Sigma_{\Gamma} \geq \lambda_{Thr}, \end{cases}$$

Όπου το j μεταβάλλεται από 0 έως $N-1$.

Στην περίπτωση που ανιχνεύονται transients, η συνάρτηση $X(i,j)$ τροποποιείται ακολούθως:

$$X(i, j) = \begin{cases} X(i, j) & \text{if } \Sigma_{\Gamma} < \lambda_{Thr} , \\ (1 - \delta) \cdot X(i, j) & \text{if } \Sigma_{\Gamma} \geq \lambda_{Thr} , \end{cases}$$

Όπου το j μεταβάλλεται από 0 έως $N-1$.

Όλα τα παραπάνω βήματα είναι μέρος ενός βρόχου, και συγκεκριμένα με βέλτιστο αριθμό επαναλήψεων $N_of_Iterations = 20$

Για τις παραμέτρους που χρησιμοποιήθηκαν στις παραπάνω συναρτήσεις ως βέλτιστες τιμές για την λειτουργία του αλγορίθμου θεωρούνται οι παρακάτω:

| ΠΑΡΑΜΕΤΡΟΣ | ΑΡΙΘΜΗΤΙΚΗ ΤΙΜΗ |
|------------------|-----------------|
| ΜΕΓΕΘΟΣ ΠΛΑΙΣΙΟΥ | 700 (ΔΕΙΓΜΑΤΑ) |
| ν | 3 |
| δ | 0.1 |
| τ | 3 |
| β | 2 |
| M | 20 |
| $Fthr$ | 106 |

Πίνακας τιμών των παραμέτρων.

3.3: Η Υλοποίηση του αλγορίθμου στο Matlab.

Η υλοποίηση του αλγορίθμου βασίστηκε στα παραπάνω βήματα που περιεγράφηκαν. Στο κεφάλαιο αυτό θα γίνει ανάλυση σε μερικά σημεία που παρουσιάζουν ενδιαφέρον:

- Κατά το short time Fourier transform, έγινε διαχωρισμός του πραγματικού μέρους του μετασχηματισμού από τη φάση, όπου και αποθηκεύτηκε σε ξεχωριστή μήτρα. Ως είσοδος στην επαναληπτική αλγοριθμική διαδικασία είναι το magnitude spectrum, δηλαδή το R. Έπειτα αναστράφηκε ο R για να ακολουθηθεί η μέθοδος του αλγορίθμου:

```
% perform STFT
while indx + wlen <= zlen
    % windowing
    xw = z(indx+1:indx+wlen).*w;

    % FFT
    X = fft(xw, nfft);

    % update the stft matrix
    stftTemp(:,col) = X(1:rownTemp);
    R(:,col) = abs(X(1:rownTemp));
    Phase(:,col)= angle(X(1:rownTemp));
    %whos phase;
    % update the indexes
    indx = indx + h;
    col = col + 1;
End
stft=R'; %inverse the matrix, in the way the algorithm uses it
%plot(stftTemp');
```

- Αφότου ανακατασκευάστηκε το σήμα με ISTFT, αφαιρέθηκε η τυχόν υπάρχουσα σιγή στο σήμα έτσι ώστε να μείνει αυτούσιο το attack transient :

```
%remove silence
frame_len = 0.005*Fs; % 0.005 per frame
N = length(x);
num_frames = floor(N/frame_len);
new_sig = zeros(N,1);
count = 0;
```

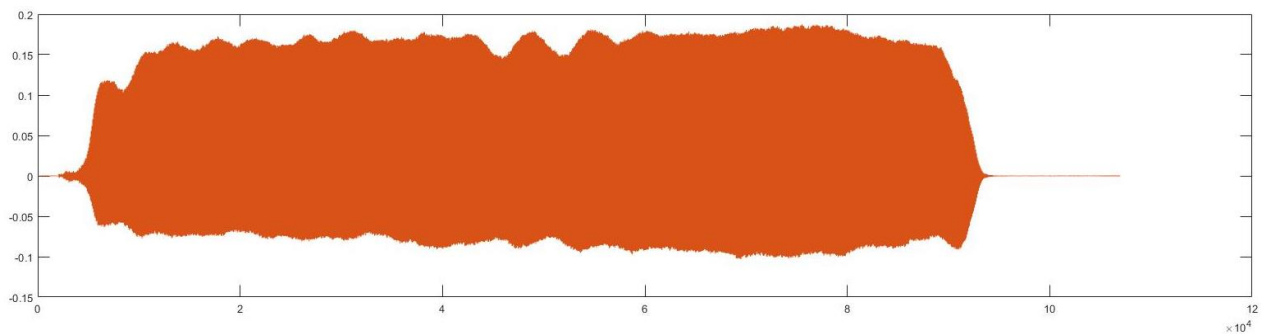
```

for k=1:num_frames
    frame = x((k-1)*frame_len+1 : frame_len*k);
    max_val = max(frame);
    %only append signal at amplitude >0.005
    if(max_val > 0.005)
        count = count+1;
        new_sig((count-1)*frame_len+1 : frame_len*count) = frame;
    end
end

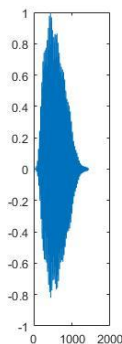
% remove trailing zero in signal
Attack =new_sig(1:find(new_sig, 1, 'last'));

```

- Οι παρακάτω κυματομορφές απεικονίζουν αντίστοιχα μια νότα φλάουτου που αποτελεί την είσοδο στον αλγόριθμο για επεξεργασία και την έξοδό του, δηλαδή το αποκομμένο attack transient.



Η γραφική απεικόνιση της κυματομορφής μιας νότας φλάουτου.



Η γραφική απεικόνιση του attack transient της νότας φλάουτου που είναι και η έξοδος του αλγορίθμου.

Κεφάλαιο 4: Το λογισμικό MLAS (Modeled Linear Arithmetic Synthesis)

4.1: Η φιλοσοφία πίσω από το λογισμικό.

Κατά την μακροχρόνια ενασχόληση του γράφοντος με τους συνθετητές, τόσο με hardware synths όσο και με software synths, διαπιστώθηκε ότι ενώ πολλές εταιρίες έχουν παράξει συνθετητές με βάση την μέθοδο Sample and Synthesis που αναλύσαμε στο πρώτο κεφάλαιο υπήρξε ένα κενό στην υλοποίησή τους.

Ενώ οι δυνατότητες δημιουργίας ήχων μέσω της χρήσης δειγμάτων έχουν αυξηθεί στο μοντέλο S&S δεν έχει βρεθεί ακόμη τρόπος μιας περισσότερο εξατομικευμένης ενοποίησης των δειγμάτων με τέτοιον τρόπο, έτσι ώστε ο χρήσης να έχει να έχει ένα τελικό, ομοιογενές αποτέλεσμα σε ένα μεγάλο συχνοτικό εύρος του κλαβιέ. Αυτό για να υλοποιηθεί υπάρχουν δύο επιλογές. Α) Η χειροκίνητη αποκοπή των attack transients και των steady states και αποθήκευση τους σε κάποιο αποθηκευτικό μέσο, ύστερα αποθήκευση των επιλεγμένων στη μνήμη ram του συνθετητή, επεξεργασία και συγκόλληση με την χρήση amplitude envelopes και άλλων εργαλείων επεξεργασίας τυχόν διαθέσιμα. Μία διαδικασία εξαιρετικά χρονοβόρα. Β) Η αυτοματοποιημένη αποκοπή και συγκόλληση των επιμέρους μερών, και την δυνατότητα επιλογής δειγμάτων από τον χρήστη.

Η έως σήμερα επιλογή ήταν επί της ουσίας η χρήση ήδη αποθηκευμένων attack και steady state στη μνήμη. Αυτά πολλές φορές ήταν δείγματα από διάφορες συχνότητες, η στις πιο πρώιμες εκδοχές τους ένα και μόνο attack ή steady state από κάθε όργανο. Η αναπαραγωγή τους σε χαμηλότερες ή υψηλότερες συχνότητες επέφερε και τα προβλήματα που αναφέρθηκαν. Πολύ συχνό σε αυτές τις περιπτώσεις ήταν και το φαινόμενο της αναδίπλωσης (aliasing).

Πολλοί ισχυρίζονται ότι το φαινόμενο της συχνοτικής μεταβολής των formants κατά την αλγοριθμική μεταβολή των συχνοτήτων των δειγμάτων,

προσέδιδε μια μουσικότητα και έναν χαρακτήρα σε κάθε μουσικό όργανο της εκάστοτε εταιρίας. Επί παραδείγματι, το πιάνο που περιλαμβάνονταν στο Korg M1 ήταν αποτέλεσμα S&S. Ο χαρακτήρας του είναι αδύνατον να αντιγραφεί με κανέναν άλλο τρόπο σύνθεσης, καθότι περιορισμοί όπως αυτός της χαμηλής συχνότητας δειγματοληψίας για τα σημερινά δεδομένα (32 kHz) και μικρά steady state δείγματα παιγμένα σε βρόγχο συμβάλλουν επίσης στη μοναδικότητά του. Ο συγκεκριμένος ήχος πιάνου, δε θεωρείται καθόλου ρεαλιστικός για τα σημερινά δεδομένα, συνέδεσε εντούτοις τον χαρακτηριστικό του ήχο με τη μουσική House, με αποτέλεσμα να χρησιμοποιείται ως ο κατεξοχήν ήχος πιάνου στις house παραγωγές έως τη στιγμή που γράφονται αυτές οι γραμμές.

Τα σημερινά δεδομένα έχουν αλλάξει. Ο περιορισμός της μνήμης δεν υπάρχει πια με αποτέλεσμα της δυνατότητας δειγματοσμού ολόκληρου του συχνοτικού εύρους (όλων των νοτών) που δύναται να αναπαράξει ένα όργανο. Επιπροσθέτως, τα ποιοτικά επίπεδα στη μουσική παραγωγή έχουν ανέβει. Αυτό επιβάλλει την προσαρμογή της μεθόδου S&S.

Ως βάση της μοντελοποίησης, χρησιμοποιήθηκαν οι αρχές της Linear Arithmetic Synthesis της Roland που είναι από τις πρώτες μορφές S&S που χρησιμοποιήθηκαν σε συνθετητή.

Στόχος του MLAS είναι η αυτοματοποιημένη συνένωση των attack transient και steady state δειγμάτων που δίνει ο χρήστης με στόχο την δημιουργία υβριδικών οργάνων. Συνεπώς εισάγονται οι εξής νέες δυνατότητες στην S&S:

A) Η αυτοματοποιημένη αποκοπή ενός Attack Transient με την χρήση του αλγορίθμου που αναπτύχθηκε παραπάνω.

B) Η αυτοματοποιημένη συγκόλλησή των δύο μερών χωρίς την ανάγκη επεξεργασίας του attack και steady state από τον χρήστη. Αυτό πραγματοποιείται κρατώντας αμετάβλητο το attack transient. Η επεξεργασία από το λογισμικό, γίνεται στην κυματομορφή που θα χρησιμοποιηθεί ως steady state. Σε αυτήν εφαρμόζουμε έναν amplitude envelope, γραμμικό, αυξανόμενης έντασης ξεκινώντας από την αρχή του σήματος. Βάση πειραματισμών, το μέγεθός του ορίστηκε στα $n=1000$ δείγματα. Το attack transient αθροίζεται στο μορφοποιημένο σήμα από την έναρξή του, με τέτοιον

τρόπο ώστε τα δύο σήματα να έχουν κοινή αρχή. Στην πράξη έχουμε αλληλοεπικάλυψη των δύο σημάτων και με την χρήση του *amplitude envelope* στο ένα από τα δύο σήματα επιτυγχάνεται τόσο η μεγιστοποίηση της συχνοτικής παρουσίας του *attack transient* στο τελικό σήμα, όσο και η ομαλή μετάβαση από το ένα σήμα στο άλλο.

Γ) Η εισαγωγή δειγμάτων από τον χρήστη. Αν και η εισαγωγή δειγμάτων και η μορφοποίησή τους είναι γνωστή από συνθετητές *sample based*, η εισαγωγή δειγμάτων σε *S&S* συνθετητές δεν ήταν δυνατή. Μόνο σε κάποιες περιπτώσεις, υπήρχε η δυνατότητα επέκτασης των κυματομορφών μέσω καρτών *rom*, χωρίς όμως τη δυνατότητα επιπλέον επεξεργασίας.

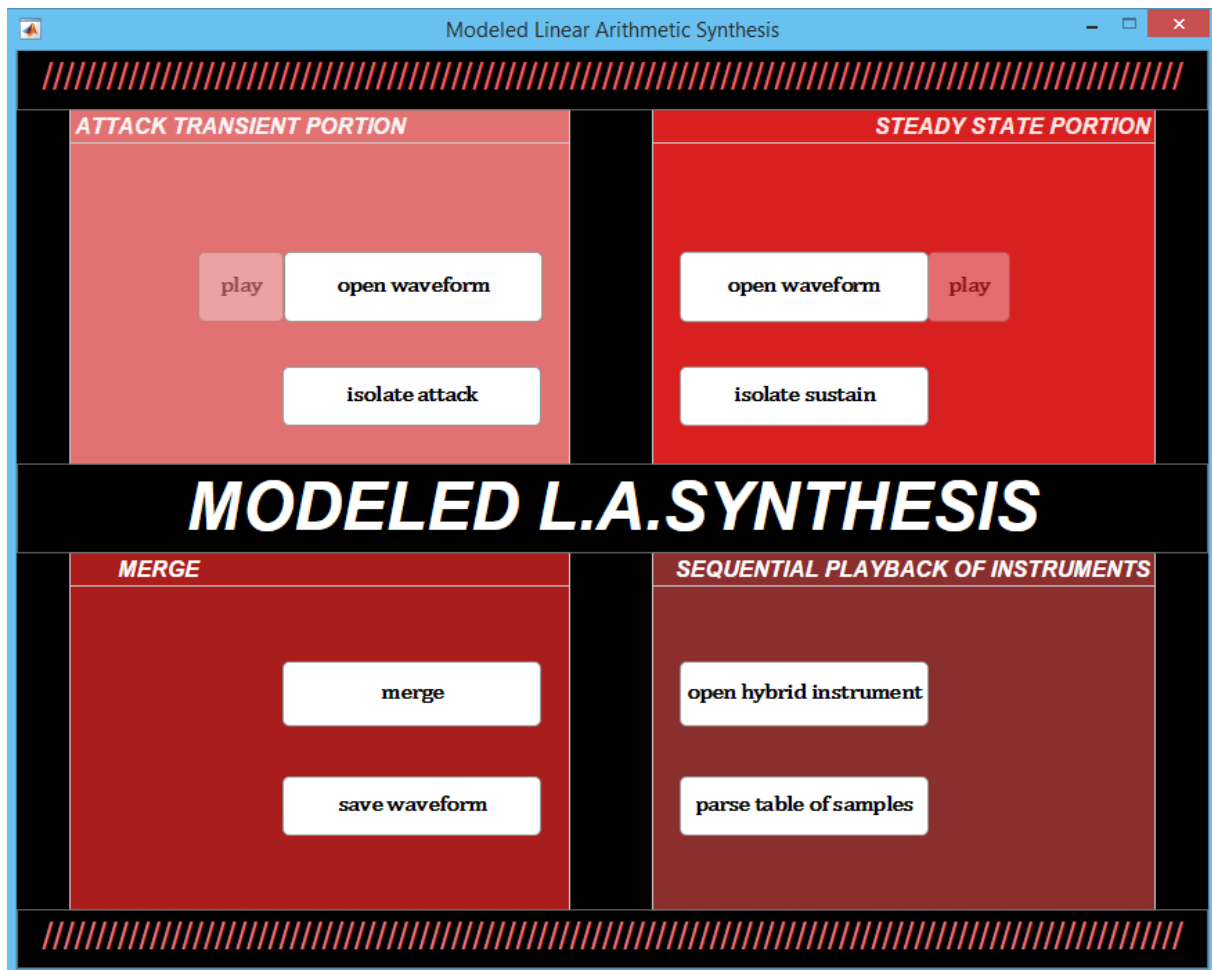
Να σημειωθεί ότι το λογισμικό περιορίζεται στη σύνθεση του νέου υβριδικού οργάνου το οποίο μπορεί να αποτελέσει τη βάση για επιπλέον μορφοποίηση από έναν ολοκληρωμένο συνθετητή.

4.2 Περιγραφή των επιμέρους μερών του λογισμικού.

Το λογισμικό αποτελείται από την κεντρική διεπαφή η οποία χωρίζεται σε τέσσερα επιμέρους διακριτά μέρη με τους εξής τίτλους:

- 1) *ATTACK TRANSIENT PORTION*. Μέσω των εικονικών κομβίων, ο χρήστης επιλέγει το δείγμα μουσικού από το οποίο θα χρησιμοποιήσει το *attack transient* το οποίο και αποκόπτεται αυτόματα.
- 2) *STEADY STATE PORTION*. Μέσω των εικονικών κομβίων, ο χρήστης επιλέγει το δείγμα μουσικού από το οποίο θα χρησιμοποιήσει το *steady state* το οποίο και αποκόπτεται αυτόματα.
- 3) *MERGE*. Μέσω των εικονικών κομβίων, ο χρήστης συνθέτει τον υβριδικό ήχο από τα *attack transient* και *steady state* που παράχθηκαν από τα δύο ανωτέρω μέρη και τον αποθηκεύει σε φάκελο της επιλογής του.

- 4) *SEQUENTIAL PLAYBACK OF INSTRUMENTS*. Αφότου ολοκληρωθεί η συνένωση των νοτών και η αποθήκευσή τους, ο χρήστης φορτώνει το υβριδικό μουσικό όργανο της επιλογής του. Ύστερα μπορεί να ακούσει τις νότες του οργάνου να αναπαράγονται διαδοχικά από την χαμηλότερη προς την υψηλότερη με μικρές παύσεις ανάμεσά τους.



Η γραφική διεπαφή του λογισμικού.

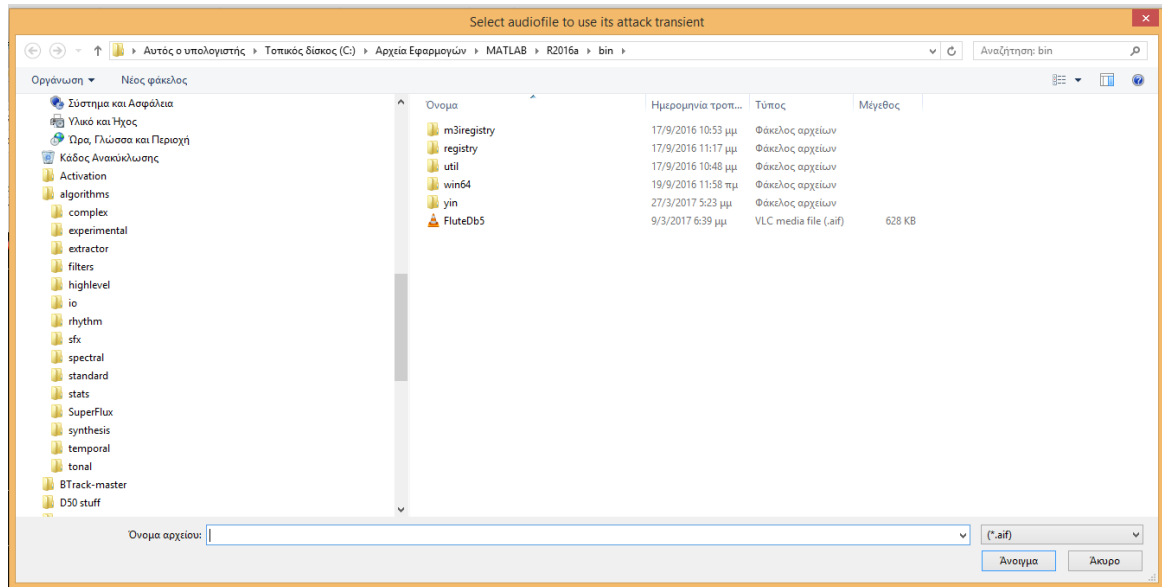
4.3 Περιγραφή της λειτουργίας του λογισμικού.

Το λογισμικό γράφτηκε χρησιμοποιώντας το Matlab. Για την δημιουργία της γραφικής διεπαφή χρησιμοποιήθηκε το εργαλείο App Designer. Πλην των συναρτήσεων που περιλαμβάνονται στο πακέτο του Matlab, υλοποιήθηκε ο αλγόριθμος εύρεσης του attack transient και χρησιμοποιήθηκε ο υλοποιημένος αλγόριθμος εύρεσης θεμελιώδους συχνότητας «Yin»(23), που βασίζεται στην αυτοσυσχέτιση. Το λογισμικό έχει τη δυνατότητα επεξεργασίας και αναπαραγωγής αρχείων ήχου της μορφής “.wav” και “.aif” που θεωρούνται οι πιο διαδεδομένες μορφές ασυμπίεστων αρχείων ήχου.

Για τον πειραματισμό και την δημιουργία υβριδικών μουσικών οργάνων χρησιμοποιήθηκαν δείγματα της βάσης δεδομένων δειγμάτων ήχου μουσικών οργάνων του University of Iowa. Τα δείγματα είναι ηχογραφημένα με πολύ υψηλές προδιαγραφές και διατίθενται ελεύθερα στην παρακάτω ηλεκτρονική διεύθυνση του πανεπιστημίου: <http://theremin.music.uiowa.edu/MIS.html>

Τα τμήματα «*ATTACK TRANSIENT PORTION*», «*STEADY STATE PORTION*», «*SEQUENTIAL PLAYBACK OF INSTRUMENTS*» λειτουργούν ανεξάρτητα μεταξύ τους. Το τμήμα «*MERGE*» από τη στιγμή που λειτουργεί έχοντας ως είσοδο τις εξόδους των τμημάτων «*ATTACK TRANSIENT PORTION*» και «*STEADY STATE PORTION*» μπορεί να εκτελεστεί μόνο μετά την ολοκλήρωση της εκτέλεσης των δύο παραπάνω τμημάτων.

Στο τμήμα «*ATTACK TRANSIENT PORTION*» ο χρήστης επιλέγει το δείγμα που θα χρησιμοποιηθεί για την αποκοπή του attack transient. Πατώντας το κουμπί “open waveform” ο χρήστης καλείται να επιλέξει το αρχείο που θα χρησιμοποιηθεί για το attack transient.



Το παράθυρο επιλογής αρχείου ήχου.

Μετά το άνοιγμα, ο χρήστης μπορεί να ακούσει το αρχείο που επιλέχθηκε πατώντας το κουμπί της ακρόασης. Ύστερα, έχει τη δυνατότητα να αποκόψει αυτόματα το attack transient. Σε αυτό το τμήμα γίνεται χρήση του αλγόριθμου για την εύρεση του attack transient που υλοποιήθηκε για τις ανάγκες αυτής της μελέτης. Ύστερα πραγματοποιείται αυτόματη αναπαραγωγή του attack transient. Ακολουθεί ο κώδικας που εκτελείται κατά το πάτημα του κουμπιού “isolate attack”:

```
global Attack;
global AttTrans;
if isempty(AttTrans)==1
    msgbox('Please load the attack portion.','error');
else
    Attack=TranDet;% run attack transient detection algorithm
    sound(Attack,16000);% listen to the result
```

Στο τμήμα «*STEADY STATE PORTION*» Ο χρήστης επιλέγει πατώντας το κουμπί open waveform, μέσω ενός παραθύρου το δείγμα που θα χρησιμοποιηθεί για την αποκοπή του steady state. Πατώντας το κουμπί της ακρόασης, μπορεί να ακούσει το αρχείο που επιλέχθηκε. Η αποκοπή του Steady State, δεν απαιτεί τη χρήση εξειδικευμένου αλγόριθμου καθότι είναι ένα σχετικά σταθερό συχνοτικά και σε

επίπεδο έντασης τμήμα μιας νότας. Σε αυτό μπορεί να περιλαμβάνεται και το τμήμα διακοπής της αναπαραγωγής της νότας από τον οργανοπαίκτη, αλλά αυτό δεν επηρεάζει τη διαδικασία συνένωσης. Επιπλέον, πραγματοποιείται η αυτοματοποιημένη αποκοπή της σιγής σε περίπτωση που υπάρχει στην αρχή και στο τέλος της κυματομορφής. Έτσι γνωρίζοντας από τη θεωρία ότι ο μέγιστος χρόνο διάρκειας ενός attack transient φτάνει τα 100 ms, επιλέγεται ένα πολλαπλάσια μεγαλύτερο κομμάτι (300 ms) από την αρχή της νότας, το οποίο και αποκόπτεται. Το υπολειπόμενο είναι το steady state που θα χρησιμοποιηθεί κατά τη συνένωση. Μετά την αποκοπή πραγματοποιείται η αυτόματη αναπαραγωγή του steady state. Ακολουθεί ο κώδικας που εκτελείται κατά το πάτημα του κουμπιού “isolate sustain”:

```
global SteadStat;
global Sustain;
if isempty(SteadStat)==1
    msgbox('Please load the steady state portion.','error');
else
    [x, Fs] = audioread(SteadStat); %read opened audio file
    x = (x(:,1)+x(:,2))/2; %stereo to mono
    %remove silence
    frame_len = 0.01*Fs; % 0.01 per frame
    N = length(x);
    num_frames = floor(N/frame_len);
    new_sig = zeros(N,1);
    count = 0;

    for k=1:num_frames
        frame = x((k-1)*frame_len+1 : frame_len*k);
        max_val = max(frame);
        %only append signal at amplitude >0.005
        if(max_val > 0.005)
            count = count+1;
            new_sig((count-1)*frame_len+1 : frame_len*count) = frame;
        end
    end

    % remove trailing zero in signal
    y =new_sig(1:find(new_sig, 1, 'last'));
```

```

re=resample(y,16000,Fs);%resampling
z = re/max(abs(re(:)));%normalisation
Sustain =z(4800:end); %select the steady state
sound(Sustain, 16000); %listen to the selected part of the
waveform
end

```

Στο τμήμα «MERGE» πραγματοποιείται η συνένωση. Προαπαιτούμενη συνθήκη, η ύπαρξη attack transient και steady state. Σε περίπτωση απουσίας ενός εκ των δύο ή και των δύο κομματιών που απαιτούνται για συνένωση, εμφανίζεται μήνυμα σφάλματος. Η μέθοδος συνένωσης προέρχεται από την αυθεντική μέθοδο συνένωσης που συναντάμε στην L.A. Synthesis. Αυτή περιλαμβάνει την εφαρμογή ενός amplitude envelope σε $n=1000$ δείγματα, γραμμικού το οποίο είναι τύπου «fade in». Στη συνέχεια υπάρχει αλληλοεπικάλυψη του steady state από το attack transient από το σημείο έναρξης του steady state έως και όσα δείγματα διαρκεί το attack transient. Η συνένωση γίνεται αγνοώντας την περιοδικότητα βασιζόμενοι ότι το attack transient είναι σχετικά μη περιοδικό. Επιπλέον να σημειωθεί ότι ο τρόπος της συνένωσης είναι σχετικά «πρωτόγονος» συγκριτικά με τις διαθέσιμες μεθόδους (π.χ. PSOLA), εντούτοις το M.L.A.S. όπως και η μέθοδος στην οποία βασίζεται, έχει σαν στόχο την μουσικότητα του αποτελέσματος της συνένωσης και όχι την ακρίβεια. Μετά τη συνένωση ακολουθείται αυτόματη αναπαραγωγή του υβριδικού μουσικού οργάνου. Ύστερα ο χρήστης επιλέγει να αποθηκεύσει το αρχείο ήχου στον φάκελο της επιλογής του. Το αρχείο ήχου που αποθηκεύεται μπορεί να είναι της μορφής “.wav”. Ακολουθεί ο κώδικας που εκτελείται κατά το πάτημα του κουμπιού “merge”:

```

n = 1000; % number of fade-in samples.
W = linspace(1,0,n)'; %'
global Attack;
global Sustain;
global Merged;
if isempty(Attack)==1 & isempty(Sustain)==0 % error messages
    msgbox('Please load the attack portion.','error');
end

```

```

if isempty(Sustain)==1 & isempty(Attack)==0
    msgbox('Please load the steady state portion.','error');
end
if isempty(Attack)==1 & isempty(Sustain)==1
    msgbox('Please load the attack and the steady state
portions.','error');
end
if isempty(Attack)==0 & isempty(Sustain)==0 %only if attack and
steady state acquired
Sustain(1:n) = Sustain(1:n).*(1-W);
Merged = zeros(size(Sustain,1), 1);% initialise Merged vector
Merged(1:size(Attack,1)) = Attack;% use the attack vector
Merged(1:end) = Merged(1:end) + Sustain;% add the sustain to the mix

sound(Merged,16000);
delete 'merged.wav';
audiowrite('merged.wav',Merged,16000);

```

Τέλος, στο τμήμα «*SEQUENTIAL PLAYBACK OF INSTRUMENTS*» δίνεται η δυνατότητα της ακρόασης του δημιουργημένου υβριδικού οργάνου. Ο χρήστης επιλέγει όλες τις νότες του υβριδικού οργάνου που δημιούργησε και μπορεί να τις ακροαστεί διαδοχικά από νότα με την χαμηλότερη θεμελιώδη αρμονική συχνότητα προς την νότα με την υψηλότερη.

Το συγκεκριμένο τμήμα λειτουργεί έχοντας επιλέξει παραπάνω από ένα αρχείο ήχου/νότας. Εναλλακτικά, εμφανίζεται μήνυμα σφάλματος. Αφότου γίνει επιλογή των νοτών, γίνεται ανίχνευση της θεμελιώδους συχνότητας μέσω του αλγορίθμου «YIN». Αυτή αποθηκεύεται σε μία μήτρα μαζί με την νότα. Στη συνέχεια, ταξινομούνται οι νότες βάση της τιμής της θεμελιώδους αρμονικής τους με αύξουσα σειρά. Αφότου ολοκληρωθεί η παραπάνω διαδικασία, ο χρήστης μπορεί να επιλέξει να ακροαστεί τις νότες με την σειριακή αναπαραγωγή των δειγμάτων. Ακολουθεί ο κώδικας που εκτελείται κατά το πάτημα του κουμπιού “open hybrid instrument”:

```

[filename, pathname, FilterIndex] =
uigetfile({'*.wav'; '*.aif'}, 'Select Multiple Notes of Hybrid
Instrument', 'MultiSelect', 'on');
z=size(filename,2);
if iscell(filename)==0
    msgbox('Please select more than one note to fill the table of
samples.', 'error'); %error message if less than two notes are
selected
else
    global ScaleOfNotes;
    global Fs;
    global prd;
    ScaleOfNotes=cell(z,2); %create empty cell matrix to
store notes
    for i=1:z
        filename{i};
        note= strcat(pathname, filename{i});
        [x,Fs]= audioread(note);
        if Fs~=16000
            Fs=resample(x,16000,Fs);
        end
        N = ndims(x); %if not mono
        if N > 2;
            x= (x(:,1)+x(:,2))/2;
        end
        r=yin2(x(5000:8000)); % run yin algorithm to find
fundamental frequency
        ScaleOfNotes{i,1}= x; %fill note in cell matrix
        ScaleOfNotes{i,2}= 1/r; % write next to each note
it's fundamental frequency
    end
    ScaleOfNotes=sortrows(ScaleOfNotes,2);
end

```

Με τη χρήση του M.L.A.S. δημιουργήθηκαν μία σειρά από παραδείγματα υβριδικών οργάνων τα οποία είναι διαθέσιμα μαζί με το λογισμικό.

4.4 Εμπορικές προοπτικές χρήσης του αλγορίθμου και του λογισμικού.

Η σημασία ανίχνευσης του attack transient δεν περιορίζεται μονάχα στην μουσικολογική της προσέγγιση που κατά κύριο λόγο εφαρμόζεται στην παρούσα μελέτη. Ο αλγόριθμος θα μπορούσε να χρησιμοποιηθεί στη εξαγωγή του κρουστικού περιεχομένου από την αρχή διαδοχικών κυματομορφών. Αυτό μπορεί να οδηγήσει στην διευκόλυνση αναγνώρισης μοτίβων, όπως για παράδειγμα το ρυθμικό κομμάτι των προσωδιακών χαρακτηριστικών της ομιλίας.

Το λογισμικό M.L.A.S. έχει αρκετά περιορισμένη λειτουργικότητα. Εξυπηρετεί ωστόσο τις ανάγκες της μελέτης στα πλαίσια της μεταπτυχιακής διατριβής. Εστιάζει περισσότερο σε συγκεκριμένο μέρος του συστήματος συνθετητής, τον ταλαντωτή. Επιπλέον, πρωταρχικός στόχος της υλοποίησης ήταν η πρακτική εφαρμογή του αλγορίθμου ανίχνευσης attack transient και της θεωρίας που συνοδεύει την σημαντικότητα του attack για το ηχόχρωμά ενός ήχου.

Η παραγόμενες νέτες από το λογισμικό, αποτελούν την βάση για περαιτέρω επεξεργασία από τα επόμενα μέρη της αλυσίδας του συστήματος ενός συνθετητή (φίλτρα, ενίσχυση κ.λπ.). Συνεπώς, το λογισμικό θα μπορούσε να λειτουργήσει ως τη βάση για τη λειτουργία ενός ολοκληρωμένου εικονικού συνθετητή S&S.

Μια επίσης προοπτική, που δεν υλοποιήθηκε λόγω των περιορισμένων δυνατοτήτων που προσφέρει το Matlab, είναι η εφαρμογή του πρωτοκόλλου MIDI (Musical Instrument Digital Interface). Εάν σε μελλοντική έκδοση του Matlab δοθεί ολοκληρωμένο MIDI implementation, σε μία ανάλογα τροποποιημένη έκδοση του λογισμικού, ο χρήστης θα μπορεί να έχει μία ολοκληρωμένη εμπειρία αναπαράγοντας τους ήχους των υβριδικών οργάνων, με την χρήση ενός πληκτροφόρου.

Βιβλιογραφία

- (1), (2),(3), (5), (8) Martin Russ “Sound Synthesis and Sampling”, Third Edition, Focal Press
- (4), (10) “Designing Software Synthesizer Plug-Ins in C++: For RackAFX, VST3, and Audio Units” Will Pirkle , CRC Press
- (6) "National Inventors Hall of Fame announces 2004 class of inventors". Science Blog. February 11, 2004.
- (7) "Fairlight – The Whole Story". Audio Media. January 1996.
- (9) Snelgrove, Martin (2011). "Oscillator". McGraw-Hill Encyclopedia of Science and Technology, 10th Ed., Science Access online service. McGraw-Hill.
- (11) H. Thornburg: "Detection and modeling of transient audio signals with prior information", PhD Thesis, Stanford University,2005.
- (12) “Analysis of transient musical sounds by auto-regressive modeling” Florian Keiler, Can Karadogan, Udo Zölzer, and Albrecht Schneider, Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03), London, UK, September 8-11, 2003, introduction
- (13),(14), (17) “The Use of the Attack Transient Envelope in Instrument Recognition” Benedict Tan & Deep Sen, School of Electrical Engineering & Telecommunications University of New South Wales, Sydney, Australia, introduction
- (15), (18) “The effect of the attack transient on aural recognition of instrumental timbres” Ralph C. Thayer Jr, SAGE Social Science Collections
- (16), (22) «A transient detection algorithm for audio using iterative analysis of stft»,12th International Society for Music Information Retrieval Conference (ISMIR 2011), Balaji Thoshkahna, Francois Xavier Nsabimana, K.R.Ramakrishnan,
- (19) “Timbre Cues and the Identification of Musical Instruments”, E. L. Saldanha Department of Psychology, Wells College, Aurora, New York John F. Corso,

Department of Psychology, State University of New York, Cortland, New York The
Journal of the Acoustical Society of America 36, 2021 (1964)

(20) “Music Cognition” W. Jay Dowling, Dane L. Harwood, Academic Press series 1
p.7

(21) A Review on Techniques for the Extraction of Transients in Musical Signals
Laurent Daudet, Laboratoire d’Acoustique Musicale, Universit´e Pierre et Marie Curie
(Paris 6), 11 rue de Loumel, 75015 Paris, France, introduction

(23) «YIN, a fundamental frequency estimator for speech and music» Alain de
Cheveigne´, Hideki Kawahara, 2002 Acoustical Society of America.