

**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΕΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΚΡΗΤΗΣ  
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΜΟΥΣΙΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ  
ΚΑΙ ΑΚΟΥΣΤΙΚΗΣ**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**ΣΥΝΘΕΣΗ ΗΧΟΥ ΜΕΣΩ ΦΑΣΜΑΣΜΑΤΙΚΩΝ ΕΠΕΞΕΡΓΑΣΙΩΝ ΣΕ  
ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ: ΠΑΓΩΜΑ ΦΑΣΜΑΤΟΣ & ΧΡΟΝΙΚΗ/ΣΥΧΝΟΤΙΚΗ  
ΑΝΑΔΟΜΗΣΗ ΦΑΣΜΑΤΟΣ**

ΔΗΜΟΣ ΦΥΝΤΑΝΙΔΗΣ (Α.Μ. 981)

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ  
ΝΙΚΟΛΑΣ ΒΑΛΣΑΜΑΚΗΣ

ΡΕΘΥΜΝΟ 2015



## Περίληψη.

Η εργασία έχει σαν στόχο να προσφέρει νέες δυνατότητες στη σύνθεση και επεξεργασία του ήχου. Η προσέγγιση στη σύνθεση γίνεται αποκλειστικά σε φασματικό επίπεδο και η λογική της λειτουργίας που εφαρμόζεται είναι ανάλυση – επεξεργασία – ανασύνθεση ήχου σε πραγματικό χρόνο.

Έχουν κατασκευαστεί δύο εφαρμογές με αυτή τη μεθοδολογία όπου και στις δυο γίνεται χρήση του αλγορίθμου γρήγορου μετασχηματισμού Φουριέ (FFT) συνδυαστικά με αλγόριθμους επεξεργασίας γραφικών δεδομένων.

Το περιβάλλον που αναπτύχθηκαν οι εφαρμογές είναι το πρόγραμμα Max/MSP Jitter 6.0 που υποστηρίζονται όλες οι απαραίτητες υποδομές για την μελέτη και υλοποίηση των αλγορίθμων και των μέσων διεπαφής για την ανάλυση, την επεξεργασία και την ανασύνθεση σε πραγματικό χρόνο.

Στην πρώτη εφαρμογή με τίτλο Φασματικό Πάγωμα (spectral freeze), επιλέγονται στιγμιαίες φασματικές πληροφορίες από ένα ηχητικό σήμα ως πάγωμα του φάσματος του στο χρόνο. Στόχος είναι ο μουσικός/χρήστης να μπορεί να συλλέγει κατά την επιθυμία του χροιάς από ήχους που θα εισάγει στον αλγόριθμο και έπειτα να τις διαχειρίζεται ελεύθερα δημιουργώντας ένα ηχητικό αποτέλεσμα.

Στην δεύτερη εφαρμογή με τίτλο Χρονική/Συχνοτική Αναδόμηση Φάσματος (spectral offset) επεξεργάζονται τα φασματικά δεδομένα απευθείας ως προς την χρονική και συχνοτική διάσταση τους. Ο μουσικός/χρήστης θα μπορεί να πραγματοποιεί μεταβολές επάνω στο πλαίσιο ενός φασματογραφήματος με μια ελευθερία κινήσεων. Οι μεταβολές και οι κινήσεις θα τροποποιούν το περιεχόμενο του φασματογραφήματος και ταυτόχρονα το σύστημα συνθέτει το ηχητικό αποτέλεσμα σε πραγματικό χρόνο.

**Λέξεις κλειδιά: λογισμικό σύνθεσης ήχου, φασματική επεξεργασία ήχου, ανάλυση/ανασύνθεση ήχου, μουσικός προγραμματισμός , Max/MSP.**

## **Abstract**

This study aims to offer new capabilities in sound synthesis and sound processing. Approaching the issue entirely on spectral level, the systematic logic that is implemented by the applications is analysis-process-resynthesis.

There have been developed two applications that functioning with the methods and properties of the Fast Furrier Transform (FFT) accompanied with graphical data processing algorithms.

The applications were developed in the programming environment of Max/MSP and Jitter 6.0. The software provides all the necessary infrastructures to study and develop the algorithms and the interfaces that execute the analysis, the processing and the resynthesis in real time.

On the first application, titled Spectral Freezing, the user can select instant spectral frames as a time sustained spectrum. Providing the option to the musician/user to collect timbres from sounds of his choice and conduct them to a final sound result.

On the second application titled Spectral Offset, the user can edit the spectral data directly on their time and frequency dimension with live output. The musician/user will be able to modify directly the frame of a spectrogram, this modifications will recreate the contents of the spectrum while the system synthesizes in real time the sound output.

**Keywords: sound synthesis, spectral process, sound analysis/resynthesis, music programming, Max/MSP.**

## Εισαγωγή.

Η εργασία έχει σαν στόχο να προσφέρει νέες δυνατότητες στη σύνθεση και επεξεργασία του ήχου. Η προσέγγιση στη σύνθεση γίνεται αποκλειστικά σε φασματικό επίπεδο και η λογική της λειτουργίας που εφαρμόζεται είναι ανάλυση – επεξεργασία – ανασύνθεση ήχου σε πραγματικό χρόνο.

Έχουν κατασκευαστεί δύο εφαρμογές με αυτή τη μεθοδολογία όπου και στις δυο γίνεται χρήση του αλγορίθμου γρήγορου μετασχηματισμού Φουριέ (FFT) συνδυαστικά με αλγόριθμους επεξεργασίας γραφικών δεδομένων.

Το περιβάλλον που αναπτύχθηκαν οι εφαρμογές είναι το πρόγραμμα Max/MSP Jitter 6.0 που υποστηρίζονται όλες οι απαραίτητες υποδομές για την μελέτη και υλοποίηση των αλγορίθμων και των μέσων διεπαφής για την ανάλυση, την επεξεργασία και την ανασύνθεση σε πραγματικό χρόνο.

Στην πρώτη εφαρμογή με τίτλο Φασματικό Πάγωμα (spectral freeze), επιλέγονται στιγμιαίες φασματικές πληροφορίες από ένα ηχητικό σήμα ως πάγωμα του φάσματος του στο χρόνο. Στόχος είναι ο μουσικός/χρήστης να μπορεί να συλλέγει κατά την επιθυμία του χροιάς από ήχους που θα εισάγει στον αλγόριθμο και έπειτα να τις διαχειρίζεται ελεύθερα δημιουργώντας ένα ηχητικό αποτέλεσμα.

Στην δεύτερη εφαρμογή με τίτλο Χρονική/Συχνοτική Αναδόμηση Φάσματος (spectral offset) επεξεργάζονται τα φασματικά δεδομένα απευθείας ως προς την χρονική και συχνοτική διάστασή τους. Ο μουσικός/χρήστης θα μπορεί να πραγματοποιεί μεταβολές επάνω στο πλαίσιο ενός φασματογραφήματος με μια ελευθερία κινήσεων. Οι μεταβολές και οι κινήσεις θα τροποποιούν το περιεχόμενο του φασματογραφήματος και ταυτόχρονα το σύστημα θα επιστρέφει το ηχητικό αποτέλεσμα σε πραγματικό χρόνο.

Το κείμενο της εργασίας έχει σαν στόχο να εξηγήσει και να αναλύσει τις έννοιες και τις μεθόδους γύρω από την σύνθεση ήχου μέσω φασματικών επεξεργασιών και να παρουσιάσει το θεωρητικό και τεχνικό υπόβαθρο των εφαρμογών που δημιουργήθηκαν. Η προσέγγιση στο ζήτημα γίνεται μέσα από την θεωρητική ανάλυση, την τεχνική ανάλυση και την υλοποίηση των μοντέλων και των συστημάτων που εξετάζονται.

Στο κεφάλαιο 1 αρχικά γίνεται αναφορά στις θεμέλιες τεχνικές της φασματικής ανάλυσης. Επισημαίνονται οι βασικότερες μέθοδοι και το πως αυτές υλοποιήθηκαν κατά το παρελθόν σε εφαρμοσμένα συστήματα. Στη συνέχεια της ενότητας εξετάζεται και εξηγείται η αρχή λειτουργίας των εργαλείων και των υποσυστημάτων που υλοποιούνται στην εργασία. Παραδείγματα και αναφορές και από παλιότερες υλοποιήσεις σε εφαρμογές κατά το παρελθόν συμπεριλαμβάνονται ως αναφορές για την καλύτερη κατανόηση των τεχνικών.

Στο κεφάλαιο 2 παρουσιάζονται οι πρωτότυπες εφαρμογές (φασματικό πάγωμα & χρονική/συχνοτική αναδόμηση φάσματος) που δημιουργήθηκαν στα πλαίσια αυτής της εργασίας. Εξετάζονται αναλυτικά ο τρόπος λειτουργίας για τα εργαλεία, τα αλγοριθμικά συστήματα καθώς και η κάθε παράμετρος κάθε αντικειμένου που χρησιμοποιείται στις εφαρμογές.

Στο κεφάλαιο 3 αναλύεται το αλγοριθμικό περιεχόμενο των εφαρμογών. Παρουσιάζεται σε κάθε επίπεδο του λειτουργικού τμήματος των εφαρμογών μια παρουσίαση που συνοδεύεται από επεξήγηση και απεικόνιση για την καλύτερη κατανόηση. Τα στάδια περιγραφής των περιεχομένων των εφαρμογών αναπτύσσονται από το επίπεδο της διεπαφής προς το επίπεδο του αλγοριθμικού υπόβαθρου και της αρχής λειτουργίας του.

Στο κεφάλαιο 4 παρέχονται υποδείξεις, παραδείγματα και τρόποι χρήσης των συστημάτων για κάθε εφαρμογή. Τα παραδείγματα συνοδεύονται με δείγματα ήχου που καταγράφηκαν μέσω των εφαρμογών και απεικονίσεις του περιβάλλοντος διεπαφής.

Τέλος, στο κεφάλαιο 5 παρέχονται συμπεράσματα από την ανάπτυξη και δοκιμή των εφαρμογών και των ηχητικών αποτελεσμάτων τους. Αναφέρονται σε τι βαθμό επιτευχθήκαν οι στόχοι που έχει η εργασία αρχικά και τι νέα δεδομένα εισήχθησαν σε αυτούς με το πέρασμα από την θεωρία στην πράξη.

## Περιεχόμενα

<b>1. Θεωρητική ανάλυση.....</b>	<b>12</b>
1.1 Βασικές Τεχνικές.....	12
1.2 Φασματογράφημα, PV και Max/MSP τεχνικές.....	17
1.2.1 Φασματογράφημα.....	17
1.2.2 Phase Vocoder.....	18
1.2.3 Max/MSP και Jitter.....	19
1.3 Επεξεργασίες Φασματογραφημάτων, οι εφαρμογές της εργασίας.....	21
1.3.1 Επεξεργασίες Φασματογραφημάτων.....	21
1.3.2 Οι εφαρμογές.....	22
1.3.2.1 Φασματικό Πάγωμα.....	23
1.3.2.2 Χρονική/Συχντοική Αναδόμηση Φάσματος.....	23
1.3.2.3 Ένα απλό παράδειγμα.....	24
<b>2. Θεωρητική ανάλυση εφαρμογών.....</b>	<b>26</b>
2.1 Φασματικό Πάγωμα(spectral freezing).....	26
2.2 Χρονική/Συχντοική Αναδόμηση Φάσματος.....	28
2.2.1 Εγγραφή Αναπαραγωγή Αποθήκευση.....	29
2.2.2 Επεξεργασία Αναπαραγωγή.....	31
<b>3. Τεχνική ανάλυση εφαρμογών.....</b>	<b>33</b>
3.1 Φασματικό Πάγωμα.....	33
3.1.1 Η διεπαφή και οι παράμετροι του Φασματικού παγώματος.....	33
3.1.2 Ανάλυση των αλγορίθμων του Φασματικού Παγώματος.....	35
3.2 Χρονική/Συχντοική Αναδόμησης Φάσματος.....	38
3.2.1 Η διεπαφή και οι παράμετροι της Χρονικής/Συχντοικής Αναδόμησης Φάσματος.....	39
3.2.2 Ανάλυση των αλγορίθμων της εφαρμογής Χρονική/Συχντοική Αναδόμηση Φάσματος.....	43
3.2.2.1 Ο αλγόριθμος καταγραφής.....	44

3.2.2.2 Ο αλγόριθμος επεξεργασίας.....	46
<b>4. Τρόπος λειτουργίας των εφαρμογών και ηχητικά παραδείγματα.....</b>	<b>60</b>
4.1 Η εφαρμογή του Φασματικού Παγώματος.....	60
4.2 Η εφαρμογή της Χρονικής/Συχνотικής Αναδόμησης Φάσματος.....	63
<b>5. Συμπεράσματα.....</b>	<b>69</b>
<b>Βιβλιογραφία/Αναφορές.....</b>	<b>71</b>



## Περιεχόμενες εικόνες

Εικόνα 1: Η δημιουργία παραθύρου ανάλυσης.....	13
Εικόνα 2: Η ανάλυση του σήματος με την μέθοδο STFT.....	13
Εικόνα 3: Παραθυρικές συναρτήσεις.....	14
Εικόνα 4: Αριστερά το παράθυρο χωρίς επικάλυψη, δεξιά το παράθυρο με επικάλυψη.....	14
Εικόνα 5: Επικάλυψη μεταξύ διαδοχικών παραθύρων.....	15
Εικόνα 6: Η ανασύνθεση με επικάλυψη παραθύρων.....	16
Εικόνα 7: Το φάσμα(πάνω) και το φασματογράφημα(κάτω).....	17
Εικόνα 8: Η διαδοχή της φάσης κατά την ανασύνθεση.....	19
Εικόνα 9: Η μορφή αποθήκευσης δεδομένων στο jit.matrix.....	20
Εικόνα 10: Η φάση χωρίς αναδίπλωση(αριστερά) και η φάση με αναδίπλωση.....	21
Εικόνα11: Η διεπαφή της εφαρμογής pitchshift_on_jitter.maxpat.....	24
Εικόνα12: Ο αλγόριθμος επεξεργασίας και ανάλυσης της εφαρμογής pitchshift_on_jitter.maxpat.....	25
Εικόνα13: Το φάσμα χωρίς επεξεργασία(πάνω), το φάσμα μετά την επεξεργασία(κάτω).....	25
Εικόνα14: Η διεπαφή της εφαρμογής του Φασματικού Παγώματος.....	33
Εικόνα15: Ο κεντρικός αλγόριθμος της εφαρμογής.....	35
Εικόνα16: Ο αλγόριθμος FFT.....	35
Εικόνα17: Εντός του p xspectrum.....	36
Εικόνα18: Οι μεταβλητές των φασματικών περιοχών.....	37
Εικόνα19: Η διεπαφή της εφαρμογής.....	39
Εικόνα20: Παράμετροι ελέγχου αρχικοποίησης και επεξεργασίας.....	39
Εικόνα21: Παράμετροι ελέγχου μετατόπισης και ανάδρασης.....	40
Εικόνα22: Οι συναρτήσεις μεταβολής των bins.....	41
Εικόνα23: Εντολές συναρτήσεων.....	41

Εικόνα24: Εντολές επεξεργασίας του συνόλου του φασματογραφήματος.....	41
Εικόνα25: Η έξοδος του ήχου της εφαρμογής.....	43
Εικόνα26: Ο κεντρικός αλγόριθμος καταγραφής.....	44
Εικόνα27: Το subpatch jftrecplay.maxpat.....	45
Εικόνα28: Ο κεντρικός αλγόριθμος επεξεργασίας.....	46
Εικόνα29: ο υποαλγόριθμος jitterengine.....	47
Εικόνα30: Ο υποαλγόριθμος preoffset.....	48
Εικόνα31: Η εγγραφή της συνάρτησης από το buffer~ στο jit.matrix.....	48
Εικόνα32: Προσομοίωση του plane των συχνοτήτων γραφικά.....	49
Εικόνα33: Προσομοίωση του plane των συχνοτήτων αριθμητικά.....	49
Εικόνα34: Προσομοίωση συνάρτησης για τις συχνότητες.....	50
Εικόνα35: Προσομοίωση της εφαρμογής της συνάρτησης συχνοτήτων στα cells.....	50
Εικόνα36: Προσομοίωση συνάρτησης των frames.....	51
Εικόνα37: Προσομοίωση του plane και με τις δυο συναρτήσεις.....	51
Εικόνα38: Προσομοίωση του plane όπως θα έχει ανακατασκευαστεί.....	51
Εικόνα39: Το subpatch jit-rota.....	52
Εικόνα40: Το subpatch routevalues.....	53
Εικόνα41: Προσομοίωση του plane όπως εισάγεται στο jit.rota.....	54
Εικόνα42: Προσομοίωση παραμόρφωσης του άξονα των frames.....	54
Εικόνα43: Το subpatch jit-feedback.....	55
Εικόνα44: Το subpatch interpolf.....	56
Εικόνα45: Το subpatch indexcontrol, συνάρτηση ανάγνωσης των frames.....	58
Εικόνα46: Ο αλγόριθμος της ανασύνθεσης jitfft.maxpat.....	59
Εικόνα47: Συνάρτηση πλάτους των συχνοτικών bins και οι μνήμες αποθήκευσης.....	60
Εικόνα48: Η αποθήκευση σε διαφορετικές μνήμες.....	61

Εικόνα49: Όλες οι μνήμες περιέχουν φασματικά δεδομένα.....	62
Εικόνα50: Ο μηχανισμός μίξης των frames.....	62
Εικόνα51: Συναρτήσεις μετατόπισης.....	63
Εικόνα52: 1. Το αυθεντικό φασματογράφημα 2. Η εφαρμογή συνάρτησης στον κατακόρυφο άξονα 3. Η εφαρμογή συνάρτησης και στους δυο άξονες.....	64
Εικόνα 53: Στάδια επεξεργασίας στον άξονα των συχνοτήτων.....	65
Εικόνα 54: Στάδια επεξεργασίας στον άξονα των frames.....	65
Εικόνα 55: Στάδια επεξεργασίας και στους δύο άξονες.....	66
Εικόνα 56: Στάδια πολλαπλών επεξεργασιών.....	67
Εικόνα57: Εφαρμογή πολλαπλών επεξεργασιών στο φασματογράφημα.....	67
Εικόνα58: Η εφαρμογή της αποτύπωσης.....	68

## 1. Θεωρητική ανάλυση.

### 1.1 Βασικές τεχνικές.

Η κύρια τεχνική στην οποία βασίζεται η ανάπτυξη όλων των αλγορίθμων των εφαρμογών είναι η Fast Fourier Transform (FFT). Η FFT βασίζεται στη θεωρία της Ανάλυσης Φουριέ η οποία αναφέρει ότι σύνθετα περιοδικά σήματα μπορούν να ερμηνευτούν ως το άθροισμα πολλαπλών ταυτόχρονων απλών σημάτων. Η γνώση της Ανάλυσης Φουριέ υπάρχει από το 19<sup>ο</sup> αιώνα και στο παρελθόν έχουν κατασκευαστεί μοντέλα και συστήματα που υλοποιούν την Ανάλυση. Η FFT ουσιαστικά είναι η μεταφορά του μαθηματικού συστήματος του Μετασχηματισμού Φουριέ (FT) στο διακριτό σύστημα που υλοποιείται σε ηλεκτρονικούς υπολογιστές. (Roads 1996)

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt$$

Σχέση 1.1: Μετασχηματισμός Φουριέ(FT)

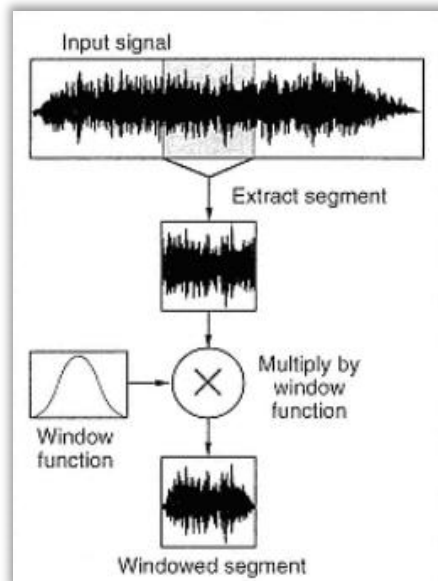
Σύμφωνα με τον Roads ο Μετασχηματισμός Φουριέ (FT) είναι μια μαθηματική συνάρτηση που αντιστοιχεί κάθε κυματομορφή συνεχόμενου χρόνου σε ένα άθροισμα από στοιχειώδη ημιτονοειδείς κυματομορφές η καθεμία με συγκεκριμένο πλάτος και φάση. (Roads 1996, σ. 550, σ. 1094). Για σύνθετα σήματα η διαδικασία της ανάλυσης γίνεται εξαιρετικά πολύπλοκη. Έτσι επινοήθηκε ένα μοντέλο ανάλυσης διακριτού χρόνου (Discrete Fourier Transform) που θα κατακερματίζει ένα συνεχόμενο σήμα σε διαδοχικά χρονικά παράθυρα ανάλυσης και στη συνέχεια θα εφαρμόζεται η ανάλυση σε κάθε παράθυρο. Ακολουθώντας αυτή τη λογική, εξελίχτηκε η μέθοδος Short-Time Fourier Transform (STFT), μια μέθοδος FFT που πραγματοποιεί ικανοποιητική ανάλυση σε σύνθετα σήματα με μεγαλύτερη ευκολία και ταχύτητα. (Roads 1996)

$$X[n, k] = \sum_{m=-\infty}^{\infty} \{x[m]h[n - m]\}e^{-j(2\pi/N)km}$$

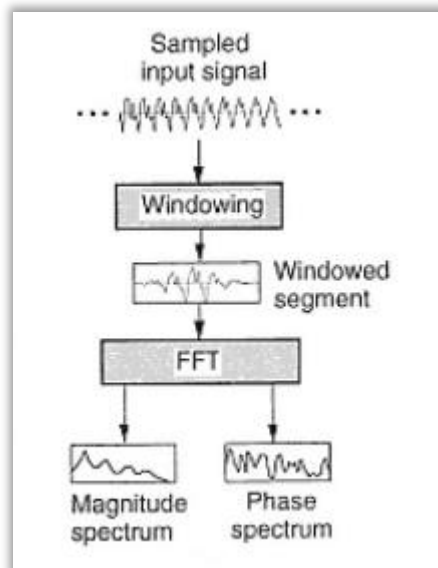
Σχέση 1.2: Το STFT

Το STFT ξεκινά να επεξεργάζεται το σήμα εισόδου  $x[m]$  χωρίζοντάς το σε ίσα μερίσματα πολύ σύντομου χρόνου, χρησιμοποιώντας μια συνάρτηση παραθύρου  $h[m]$  (Roads 1996, σ. 550). Κάθε παραθυρική συνάρτηση  $h[m]$  ουσιαστικά είναι μια περιβάλλουσα πλάτους, η οποία εφαρμόζεται σε διαδοχικές χρονικές στιγμές στο σήμα. Το αποτέλεσμα από την εφαρμογή της παραθυρικής συνάρτησης είναι το παράθυρο ανάλυσης η αλλιώς ένα FFT frame όπου εκτελείται ο FFT αλγόριθμος ανάλυσης. Αυτά που επιστρέφει η ανάλυση σε ένα FFT frame είναι τα πλάτη των συχνοτήτων και τα πλάτη των φάσεων.

Η μαθηματική ερμηνεία αυτών των αποτελεσμάτων είναι ένας μιγαδικός αριθμός, του οποίου το πραγματικό μέρος είναι η συνάρτηση του πλάτους της συχνότητας και το φανταστικό, αυτό της φάσης.

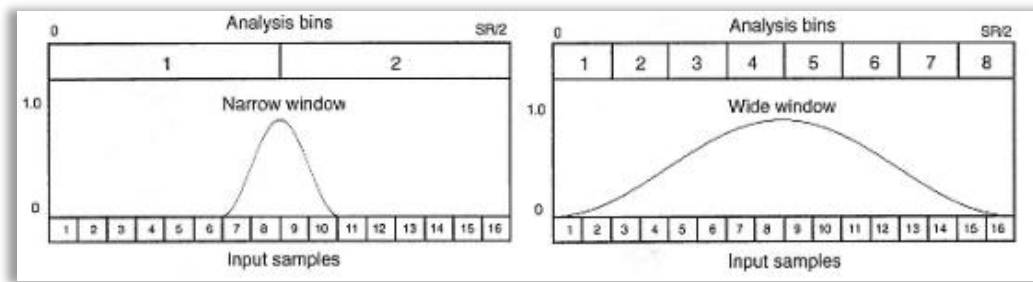


Εικόνα 1: Η δημιουργία παραθύρου ανάλυσης (Roads 1996, σ. 550).



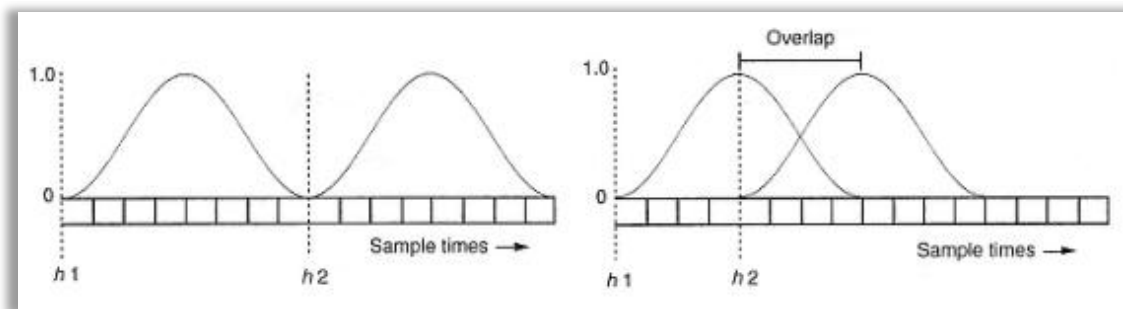
Εικόνα 2: Η ανάλυση του σήματος με την μέθοδο STFT (Roads 1996, σ. 551).

Υπάρχουν ορισμένοι τύποι παραθυρικών συναρτήσεων, των οποίων ο ρόλος είναι η καλύτερη εφαρμογή της καταγραφής του φάσματος, όπως για παράδειγμα σε σήματα που έχουν αρμονικό χαρακτήρα, χρησιμοποιούνται παράθυρα με στενότερο εύρος περιβάλλουσας, ενώ σε σήματα σύνθετα π.χ. ο θόρυβος χρειάζονται μεγαλύτερο εύρος (Roads 1996, σ. 1103).



Εικόνα 3: Παραθυρικές συναρτήσεις (Roads 1996, σ. 560)

Ο τρόπος εφαρμογής της παραθυρικής συνάρτησης είναι μέσω του πολλαπλασιασμού της περιβάλλουσας με το σήμα. Σαν αποτέλεσμα αυτού μπορούν να δημιουργούνται ανεπιθύμητα φαινόμενα τύπου διαμόρφωσης πλάτους (AM modulation), επηρεάζοντας έτσι το φασματικό περιεχόμενο. Η μέθοδος που αποτρέπει να συμβαίνει αυτό λέγεται window overlapping και λειτουργεί εφαρμόζοντας σύμπτυξη μεταξύ των διαδοχικών παραθύρων, εξαλείφοντας έτσι τις μεγάλες διακυμάνσεις.

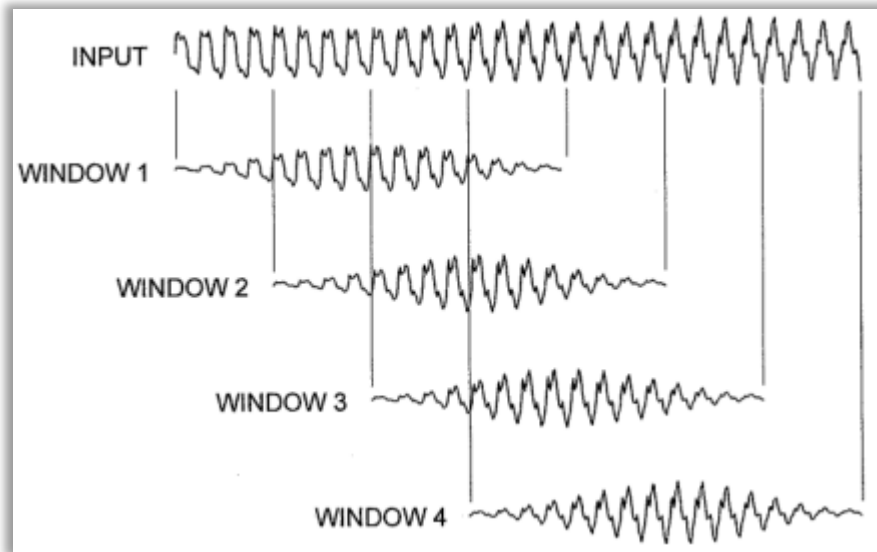


Εικόνα 4: Αριστερά το παράθυρο χωρίς επικάλυψη, δεξιά το παράθυρο με επικάλυψη (Roads 1996, σ. 569).

Το επικαλυπτόμενο εύρος που δημιουργείται από τη σύμπτυξη δυο διπλανών παραθύρων, εξαρτάται από το μέγεθος του παραθύρου και τον βαθμό επικάλυψης (overlap factor). Το εύρος αυτό, ονομάζεται hop size  $H$  (μέγεθος μεταπήδησης) και παρακάτω θα δούμε πώς ορίζεται. (Roads 1996)

$$X[l, k] = \sum_{m=0}^{M-1} h[m]x[m + (lH)]e^{-j(2\pi/N)km}$$

Σχέση 1.3: Το STFT με επικάλυψη παραθύρου.



Εικόνα 5: Επικάλυψη μεταξύ διαδοχικών παραθύρων (Dogde 1997, σ. 250).

Στο διακριτό σύστημα υπάρχουν όρια και οι τιμές των μεγεθών είναι υπολογίσιμες. Έτσι ένα σήμα εισόδου έχει συγκεκριμένα χαρακτηριστικά όπως και οι παράμετροι της ανάλυσης.

Το παράθυρο ανάλυσης έχει ένα συγκεκριμένο μέγεθος  $N$ , ένα συγκεκριμένο hop size  $H$ , σήμα εισόδου  $x[n]$  και μια συγκεκριμένη συχνότητα δειγματοληψίας  $f_s$ . Κατά συνέπεια η ανάλυση έχει ένα συγκεκριμένο βάθος ακριβείας. Όντας στο δυαδικό σύστημα, ένας υπολογιστικής ορίζει τα μεγέθη με δυνάμεις του 2. Το σήμα αποτελείται από διαδοχικά δείγματα (samples). Το παράθυρο ανάλυσης ως μικρότερο τμήμα του σήματος, αποτελείται και αυτό από δείγματα, πολύ λιγότερα όμως από το συνολικό σήμα. Για καλύτερη απόκριση του συστήματος και για τις ανάγκες της ανάλυσης, το μέγεθος του παραθύρου και του hop size σε δείγματα θα ορίζεται και αυτό ως δύναμη του 2. Παράδειγμα για κάθε ένα παράθυρο ανάλυσης μπορεί να είναι  $2^{10}$  δηλαδή 1024 δείγματα και το overlap factor  $2^1$ . Έτσι το hop size ορίζεται  $(\text{window size})/(\text{overlap factor}) = 512$  δείγματα.

Η ανάλυση χωρίζει το παράθυρο σε δύο ίσα μερίσματα, ένα για τα πλάτη των συχνοτήτων και ένα για αυτό των φάσεων. Έτσι κάθε μέρισμα έχει μέγεθος  $N/2$ . Τα δείγματα κάθε μερίσματος πλέον, αφορούν διαφορετικά μεγέθη. Έτσι για το μέρισμα με τα πλάτη των συχνοτήτων, τα δείγματα ονομάζονται frequency bins και για το μέρισμα με τα πλάτη των φάσεων, phase bins. Με  $k$  ορίζεται ο δείκτης των συχνοτικών bins (frequency bin index) παίρνοντας ακέραιες τιμές.

Η ακρίβεια της ανάλυσης εξαρτάται από όλα αυτά τα μεγέθη. Έστω ότι για συχνότητα δειγματοληψίας στο σήμα έχουμε  $f_s = 44,1$  kHz και για μέγεθος ανάλυσης  $N = 1024$  δείγματα. Αυτό συνεπάγεται ότι η μεγαλύτερη συχνότητα που θα μπορούσε να έχει το σήμα είναι τα 22 kHz ενώ το πλήθος των frequency bins είναι 512. Το συχνοτικό εύρος που αντιστοιχεί στο πρώτο bin ( $k=1$ ) είναι  $(f_s/2)/\text{bin size}$  δηλαδή  $22050/512 = 43,07$  Hz. Έτσι ορίζεται η θεμέλιος συχνότητα της ανάλυσης και όλες οι συχνότητες που θα εμφανιστούν στο φάσμα που έχει υπολογιστεί με τον αλγόριθμο FFT, θα είναι ακέραια πολλαπλάσια της θεμελίου έχοντας αρμονική σχέση μεταξύ τους. (Roads 1996)

$$fk = (k/N) * fs$$

Σχέση 1.4: Η θεμέλιος των bins του FFT.

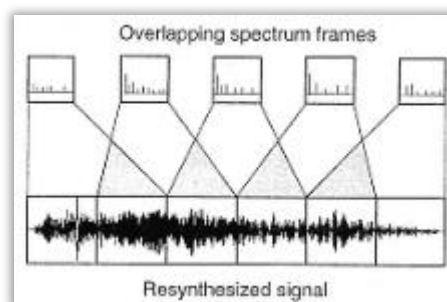
Η συσχέτιση της συχνοτικής ανάλυσης με το χρονικό εύρος του παραθύρου είναι κρίσιμη. Για μεγαλύτερο βάθος ανάλυσης, χρειάζεται μεγαλύτερο παράθυρο. Αντίθετα ένα μικρό παράθυρο θα οδηγήσει σε φτωχότερες αναλύσεις.

Μετά από μια επιτυχημένη ανάλυση του FFT αλγορίθμου σε ένα σήμα, τα δεδομένα πλάτη συχνοτήτων και φάσης μπορούν να χρησιμοποιηθούν για ανασύνθεση και στη συνέχεια να επιστρέψουν ως μια έκδοση του αρχικού σήματος. (Roads 1996)

$$X[n] = 1/N \sum_{k=0}^{N-1} X(k)e^{-j(2\pi/N)km}$$

Σχέση 1.5: Η Ανασύνθεση σε διακριτό χρόνο.

Η διαδικασία της ανασύνθεσης FFT είναι η αναστροφή της διαδικασίας της ανάλυσης. Κάθε γινόμενο ενός παραθύρου ανάλυσης (συχνότητα/φάση) χρησιμοποιείται για να κατασκευασθεί ένα νέο παράθυρο σήματος ίδιου εύρους με αυτό της ανάλυσης.



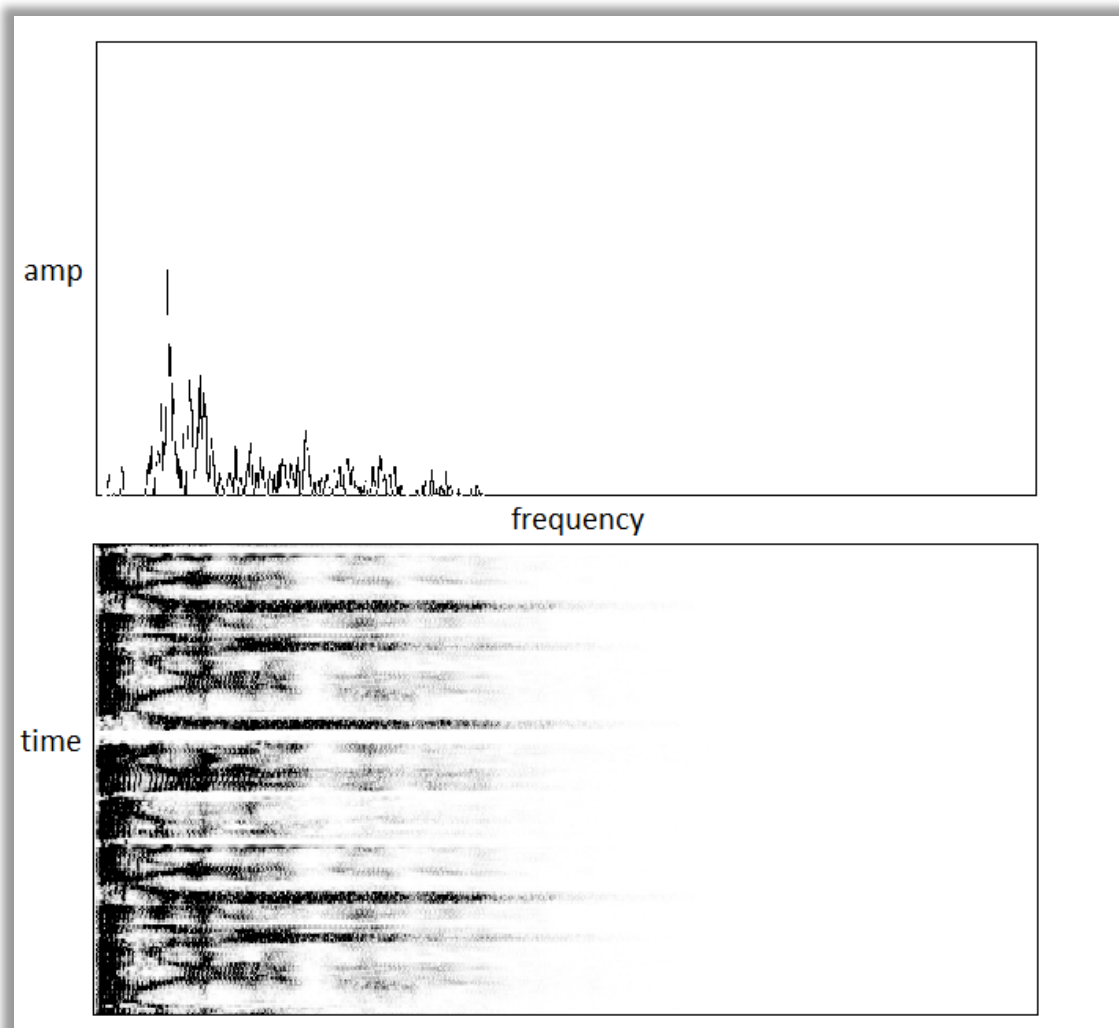
Εικόνα 6: Η ανασύνθεση με επικάλυψη παραθύρων (Roads 1996, σ. 554).

Όπως στην ανάλυση έτσι και στην ανασύνθεση υπάρχει σύμπτυξη μεταξύ των παραθύρων, με τη διαφορά ότι η άθροιση των παραθύρων παραμένει όπως είναι δημιουργώντας μια συνέχεια στο ανασυντιθέμενο σήμα. (Roads 1996, σ. 553).



## 1.2 Φασματογράφημα, PV και Max/MSP τεχνικές.

### 1.2.1 Φασματογράφημα.



Εικόνα 7: Το φάσμα(πάνω) και το φασματογράφημα(κάτω).

Η αναπαράσταση του περιεχομένου ενός παραθύρου ανάλυσης για τα πλάτη των συχνοτήτων σε ένα γράφημα, δείχνει την διακύμανση της ενέργειας ανά συχνότητα. Αυτό είναι γνωστό ως φάσμα (spectrum). Ένα φασματογράφημα (spectrogram) δείχνει σε ένα γράφημα το φάσμα όπως αυτό εξελίσσεται στο χρόνο. Έτσι παρέχεται η δυνατότητα να παρατηρηθούν περισσότερες ιδιότητες του σήματος που αναλύεται, όπως για παράδειγμα νότες, φωνοσυχνότητες, χροιές μουσικών οργάνων και γενικότερα μεταβολές ως προς το χρόνο (Roads 1996). Το φασματογράφημα ως μέσο διεπαφής είναι πρακτικό ώστε να εκτελούνται διάφορες επεξεργασίες.

Η διαφοράς του συνολικού φασματογραφήματος με το αποτέλεσμα ανάλυσης του STFT αλγορίθμου είναι στον τρόπο απεικόνισης του. Όλες οι παράμετροι παραμένουν ίδιες. Έτσι μπορούμε να σκεφτούμε το φασματογράφημα ως ένα πίνακα, του οποίου οι σειρές είναι το πλήθος των συχνοτήτων και οι στήλες τα μερίσματα του χρόνου, ενώ κάθε κελί του πίνακα κράτα την τιμή για ένα bin που υπολογίζει ο αλγόριθμος. Τα χαρακτηριστικά που παίζουν σημαντικό ρόλο στο φασματογράφημα είναι ο τρόπος αποτύπωσης του εύρους της συχνότητας, του πλάτους αυτών και αν τα συγκεκριμένα χαρακτηριστικά θα είναι γραμμικά ή λογαριθμικά. Το μέγεθος μεταπήδησης (hop size) ορίζει την απόσταση χρονικά από μία στήλη στην επόμενη, ενώ το μέγεθος του παραθύρου ανάλυσης ορίζει το πλήθος των κελιών (Roads 1996, σ. 564).

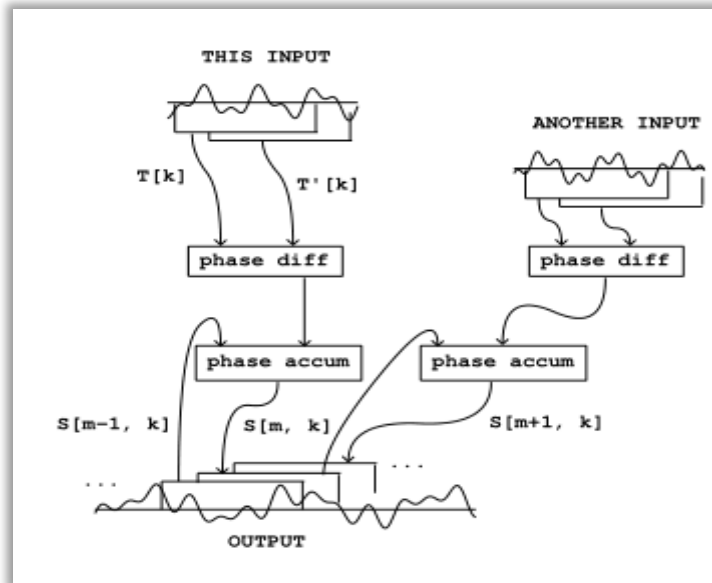
### 1.2.2 Phase Vocoder.

Μια τεχνική που πρέπει να αναφερθεί επίσης προτού ολοκληρωθεί η θεωρία είναι το Phase Vocoder (PV). Στο παρελθόν αναπτύχθηκε με σκοπό την συμπίεση και κωδικοποίηση σημάτων φωνής, με μια πιο συμβατική μέθοδο ανάλυσης και ανασύνθεσης. Το αποτέλεσμα όμως οδηγούσε σε πολύ μεγαλύτερο όγκο δεδομένων από τα αρχικά. Έτσι για κάποια χρόνια θεωρούνταν αδιέξοδη ως μέθοδος ανάλυσης, μέχρι την ανακάλυψη ενός phase vocoder που βασίζονταν στον αλγόριθμο FFT και την εφαρμογή του στην σύνθεση ήχου σε υπολογιστές (Moorer 1978). (Dodge, Jerse 1997).

$$X[k] = \sum_{m=-\infty}^{\infty} \{x[m]e^{-j(2\pi/N)km}\}h[n-m]$$

Σχέση 1.6: Το STFT με PV.

Βλέποντας το PV ως το επόμενο στάδιο στην STFT ανάλυση και ως ένα δημοφιλές εργαλείο ανάλυσης, τα οφέλη είναι σημαντικά (Drljic 2011). Το προστιθέμενο όφελος είναι η δυνατότητα να υπολογίζει την διαφορά συχνότητας από την κεντρική συχνότητα στο bin (Dodge, Jerse 1997).



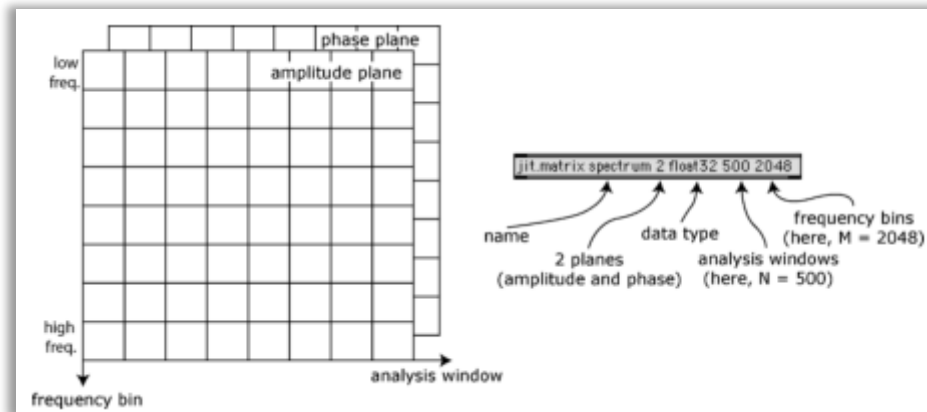
Εικόνα 8: Η διαδοχή της φάσης κατά την ανασύνθεση (Puckette 2006, σ. 282).

Καθώς το PV έχει να κάνει με τη φάση κατά το στάδιο της ανάλυσης, η διαφορά φάσης μεταξύ των διαδοχικών παραθύρων ανάλυσης σε ένα συγκεκριμένο bin, καθορίζει την ακριβή συχνότητα της ενέργειας σε αυτό το bin (Dudas, Lippe 2006). Έτσι για παράδειγμα μια ανάλυση με θεμέλιο τα 50 Hz όταν πάει να επεξεργαστεί ένα σήμα ημιτόνου στα 440 Hz, θα υπολογίσει ως κεντρική συχνότητα τα 450 Hz στο  $9^\circ$  bin, ενώ με τη διαφορά συχνοτήτων/φάσεων στο  $9^\circ$  bin, το PV βρίσκει ότι είναι 10 Hz χαμηλότερα (Dodge, Jerse 1997).

### 1.2.3 Max/MSP και Jitter.

Το προγραμματιστικό περιβάλλον Max/MSP και Jitter, είναι μια γλώσσα οπτικού και αντικειμενοστραφούς προγραμματισμού, όπου παρέχεται η δυνατότητα να υλοποιούνται με σχετικά απλό τρόπο οι πολύπλοκες τεχνικές PV και STFT. Επίσης μπορούν να υλοποιούνται διεργασίες ταυτόχρονα στο πεδίο του χρόνου και σε αυτό των συχνοτήτων, σε πραγματικό χρόνο εκτέλεσης.

Οι αλγόριθμοι PV και STFT είναι ενθυλακωμένοι στα Max/MSP αντικείμενα: `fft~`, `ifft~`, `prfft~`, `fftin~`, `fftout~`, `fftinfo~`, με δυνατότητα ρύθμισης όλων των παραμέτρων τους. Τα δεδομένα της ανάλυσης PV και STFT μπορούν να καταγραφούν και να αποθηκευτούν στα Max/MSP και Jitter αντικείμενα: `buffer~` `jit.matrix`, `jit.poke~`, `jit.catch~`, που αποτελούν δισδιάστατους πίνακες καταγραφής φάσματος για αποθήκευση συχνοτήτων και φάσεων σε κάθε κανάλι και τρισδιάστατους πίνακες για αποθήκευση επιπρόσθετα της διάστασης του χρόνου ως μορφή φασματογραφήματος.



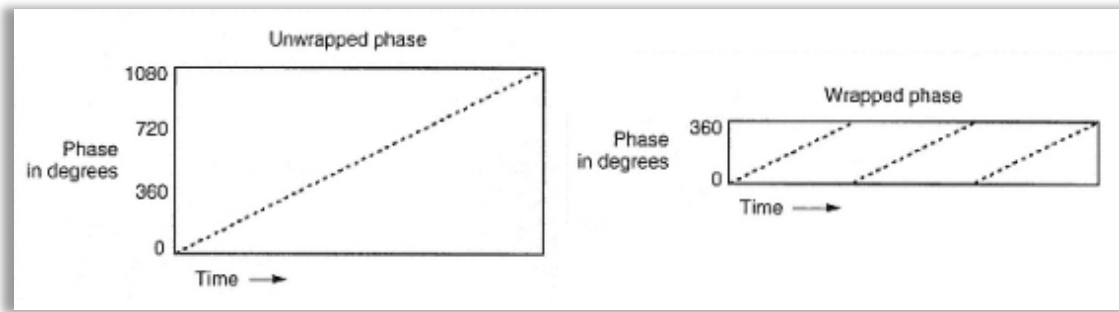
Εικόνα 9: Η μορφή αποθήκευσης δεδομένων στο *jit.matrix* (Charles 2008, σ. 90).

Τα φασματικά δεδομένα προβάλλονται στην έξοδο του συστήματος μέσω ενός δισδιάστατου πλαισίου Jitter διεργασίας. Το περιεχόμενο μπορεί να είναι δισδιάστατος πίνακας φάσματος, ή τρισδιάστατος πίνακας φασματογραφήματος. Το Jitter σήμα έχει τέσσερα κανάλια σήματος εικόνας ARGB. Για την χρήση φασματικών δεδομένων χρησιμοποιείται μόνο το πρώτο, που αποτελεί το πλάτος (alpha plane). Έτσι στην έξοδο το αποτέλεσμα είναι μια ασπρόμαυρη εικόνα.

Οι διαστάσεις του πλαισίου ορίζονται στα αντικείμενα με τέτοιο τρόπο ώστε να ταιριάζουν στα χαρακτηριστικά της ανάλυσης FFT. Έτσι το μέγεθος του δισδιάστατου πίνακα για ένα απλό φάσμα είναι όσο και το μέγεθος ενός FFT frame και το εύρος τιμών. Όταν πρόκειται για φασματογράφημα, η κάθε σειρά του πίνακα έχει το μέγεθος ενός FFT frame και η κάθε στήλη είναι το FFT frame πολλαπλασιασμένη επί το hop size (Droljc 2011, σ. 71) (Charles 2008). Έτσι ορίζεται και το χρονικό εύρος που καταγράφεται και απεικονίζεται στο φασματογράφημα.

Η αποτύπωση του φάσματος υπακούοντας στη φύση του FFT είναι σε γραμμική μορφή. Με επεξεργασία όμως του περιεχομένου μπορεί να αποτυπώνεται και σε λογαριθμική μορφή, με κόστος την απώλεια ανάλυσης στις χαμηλότερες συχνότητες. (βλ. Droljc 2011, σ. 72).

Τα δεδομένα της ανάλυσης που χρησιμοποιούνται και προβάλλονται ως διεπαφή είναι τα πλάτη των συχνοτήτων. Τα δεδομένα της φάσης τροποποιούνται και χρησιμοποιούνται για την ανακατασκευή του σήματος σύμφωνα με τις ιδιότητες του PV. Τα αντικείμενα `framedelta~` και `phasewrap~` υπολογίζουν την διαφορά φάσης μεταξύ των διαδοχικών FFT frames και την περιορίζουν σε εύρος  $360^\circ$  μεταξύ των τιμών  $-\pi$ ,  $\pi$ . Κατά την ανασύνθεση η φάση ανακατασκευάζεται με το αντικείμενο `frameaccum~`. (βλ. Droljc 2011, σ. 66).



Εικόνα 10: Η φάση χωρίς αναδίπλωση(αριστερά) και η φάση με αναδίπλωση(δεξιά)( Roads 1996, σ. 1098).

Η υλοποίηση και αποθήκευση μιας εφαρμογής μπορεί να γίνει σε αρχεία .pat και .maxpat τα οποία περιέχουν τον πηγαίο κώδικα με δυνατότητα τροποποίησης και περεταίρω ανάπτυξης, αλλά μπορεί να γίνει και σε αρχεία binary για την χρήση ως αυτόνομης εφαρμογής (stand alone), χωρίς όμως την δυνατότητα περεταίρω ανάπτυξης.

### 1.3 Επεξεργασίες Φασματογραφημάτων, οι εφαρμογές της εργασίας.

Κατά το παρελθόν αναπτύχθηκαν εφαρμογές που εκτελούν γραφικές επεξεργασίες σε φασματογραφήματα με αλγόριθμους βασιζόμενους στις FFT, PV και STFT τεχνικές. Όμως ο όγκος των δεδομένων σε σχέση με τις δυνατότητες των υπολογιστών της εποχής για το ευρύ κοινό, καθιστούσαν αδύνατο ώστε οι επεξεργασίες να γίνονται σε πραγματικό χρόνο.(Charles 2008).

#### 1.3.1 Επεξεργασίες Φασματογραφημάτων.

Κάποιες από τις πρώτες εφαρμογές με επεξεργασίες φασματογραφημάτων σε μη πραγματικό χρόνο που αξίζει να αναφερθούν, είναι : το SpecDraw που αναπτύχθηκε από τον Gerhard Eckel το 1989 στο IRCAM σε σύστημα NeXT και παρείχε την δυνατότητα σχεδίασης πολλαπλών φίλτρων πάνω στο φασματογράφημα, το AudioSculpt και το SuperPV που αναπτύχθηκαν αργότερα το 1996 στο IRCAM και βασίστηκαν στο SpecDraw προσθέτοντας περισσότερες δυνατότητες στη ρύθμιση των παραμέτρων του FFT(Charles 2008). Επίσης η εφαρμογή MetaSynth αναπτύχθηκε το 2005 και παρέχει την δυνατότητα επεξεργασίας του φασματογραφήματος με εργαλεία γραφικής επεξεργασίας. (Charles 2008). Ακόμα η εφαρμογή SPEAR το 2006, με δυνατότητα σχεδίασης εξ' ολοκλήρου ενός φασματογραφήματος.

Καθώς όμως η τεχνολογία εξελίσσεται, παρέχεται περισσότερη υπολογιστική ισχύς για την ανάπτυξη και εκτέλεση εφαρμογών. Έτσι εφαρμογές με διαδραστικό περιβάλλον και λειτουργίες σε πραγματικό χρόνο, μπορούν να κατασκευαστούν και να χρησιμοποιηθούν ευκολότερα. Όπως αναφέρθηκε παραπάνω, το περιβάλλον Max/MSP Jitter λειτουργεί σε πραγματικό χρόνο και εκτελεί FFT ανάλυση/ανασύνθεση. Η πρώτη προσέγγιση έγινε από τον Luke Dubois, προγραμματιστή της Jitter, που κατασκεύασε το 2005 μια εφαρμογή που αποθηκεύει FFT δεδομένα σε ένα Jitter πίνακα στη συνέχεια τα τροποποιεί και τέλος τα αναπαράγει στην έξοδο ως ήχο μέσω FFT ανασύνθεσης. (Charles 2008).

Αργότερα ο Jean Francois Charles το 2008 ανέπτυξε εφαρμογές Max/MSP/Jitter καθιερώνοντας μια σειρά από μεθόδους στον τρόπο ανάπτυξης και σχεδίασης εφαρμογών που διευρύνουν κατά πολύ της δυνατότητες επεξεργασίας σε πραγματικό χρόνο. Οι μέθοδοι που ανέπτυξε ο Charles στις οποίες σε μεγάλο βαθμό βασίστηκε η παρούσα εργασία, βρίσκονται στις εφαρμογές για την καταγραφή και αναπαραγωγή *2-record-play-frame-by-frame.pat* και *3-record-play-speed-control.pat*, για την ομαλότερη αναπαραγωγή των frames κατά την ανασύνθεση *6-interpolate-2frame.pat*. Επίσης στις εφαρμογές διεπαφής με τον πλαίσιο του φασματογραφήματος *4-explore-the-sonogram.pat* και *5-graphical-transformations.pat*.(Charles 2008 σ. 87-102).

Το 2011 ο Tadej Droljc στην μελέτη του STFT Analysis Driven Sonographic Sound Processing in Real-Time using Max/MSP and Jitter, ανέπτυξε εφαρμογές που εξέλιξαν την εργασία του Charles, εφαρμόζοντας γνωστές τεχνικές ηχητικών επεξεργασιών, παράλληλα με τεχνικές επεξεργασίας εικόνας και σχεδίασης γραφικών επάνω στο φάσμα και στο φασματογράφημα. Τεχνικές σχεδίασης και ανάπτυξης του Droljc βοήθησαν στην κατανόηση και κατασκευή των ιδεών και των εφαρμογών της παρούσας εργασίας σε σημαντικό βαθμό επίσης.

### **1.3.2 Οι εφαρμογές.**

Οι εφαρμογές της εργασίας είναι το Φασματικό Πάγωμα (spectral freeze) και η Χρονική/Συχνοτική Αναδόμηση Φάσματος (spectral offset). Η ιδέα και στις δύο εφαρμογές είναι η σύνθεση και σχεδίαση ήχων σε πραγματικό χρόνο με διεπαφή την εφαρμογή, τον ηλεκτρονικό υπολογιστή και το ηχητικό σύστημα.

### **1.3.2.1 Φασματικό Πάγωμα.**

Το Φασματικό Πάγωμα (spectral freeze) εκτελεί συνεχόμενη ανάλυση σε ένα σήμα εισόδου. Ο μουσικός/χρήστης έχει την δυνατότητα να κρατήσει μεμονωμένα παράθυρα ανάλυσης από διαφορετικές χρονικές στιγμές, σε ένα σύστημα καταγραφής κατ' επιλογή του, να επιλέξει σε τι ποσοστό και εάν θα προστίθεται κάθε παράθυρο ανασύνθεσης και να επιλέξει πόσα παράθυρα θα συνδυάζονται. Τέλος θα έχει την δυνατότητα μέσω μίας διεπαφής να πραγματοποιεί με απλή κίνηση του δείκτη ένα ηχητικό αποτέλεσμα.

Σύμφωνα με τον Roads, η τεχνική Spectral Interpolation Synthesis (SIS) επινοήθηκε το 1990 από τους Serra, Rubine, Dannenberg και πρόκειται για μια πειραματική τεχνική που παράγει ήχους διαφορετικής χρονικής διάρκειας μέσω επικάλυψης αναλυμένων φασμάτων. Πρόκειται για προσθετική σύνθεση που υλοποιείται στο πεδίο των συχνοτήτων με χρήση ηχογραφημένων και αναλυμένων δειγμάτων ήχου. (Roads 1996, σ. 151). Ο Roads επίσης αναφέρει την ανάγκη για κωδικοποίηση των φασματικών δεδομένων και την χρήση συναρτήσεων που θα ελέγχουν την μεταβίβαση μεταξύ των φασμάτων. Αναφέρει επίσης την δυσκολία στην μεταχείριση των απότομων δυναμικών μεταβολών στα δείγματα ήχων.

Η ιδέα της εφαρμογής, έχει κοινά χαρακτηριστικά με την τεχνική SIS. Ο στόχος είναι ουσιαστικά ο ίδιος, δηλαδή η προσθετική σύνθεση με χρήση φασματικών δεδομένων. Είναι όμως μια εφαρμογή πραγματικού χρόνου και γι' αυτό πολλές παράμετροι είναι απαραίτητο να ρυθμίζονται με σχετικά γρήγορο τρόπο και να ανάγονται σε απλοποιημένες διεπαφές.

Με τις ευκολίες που παρέχουν οι τεχνικές καταγραφής φασματικών δεδομένων στο περιβάλλον Max/MSP Jitter δηλαδή με τη χρήση αντικειμένων τύπου multislider/nodes, η εφαρμογή καταφέρνει να υλοποιήσει μια μορφή της τεχνικής SIS και το ηχητικό αποτέλεσμα να έχει ένα ενδιαφέρον.

### **1.3.2.2 Χρονική/Συχνοτική Αναδόμηση Φάσματος.**

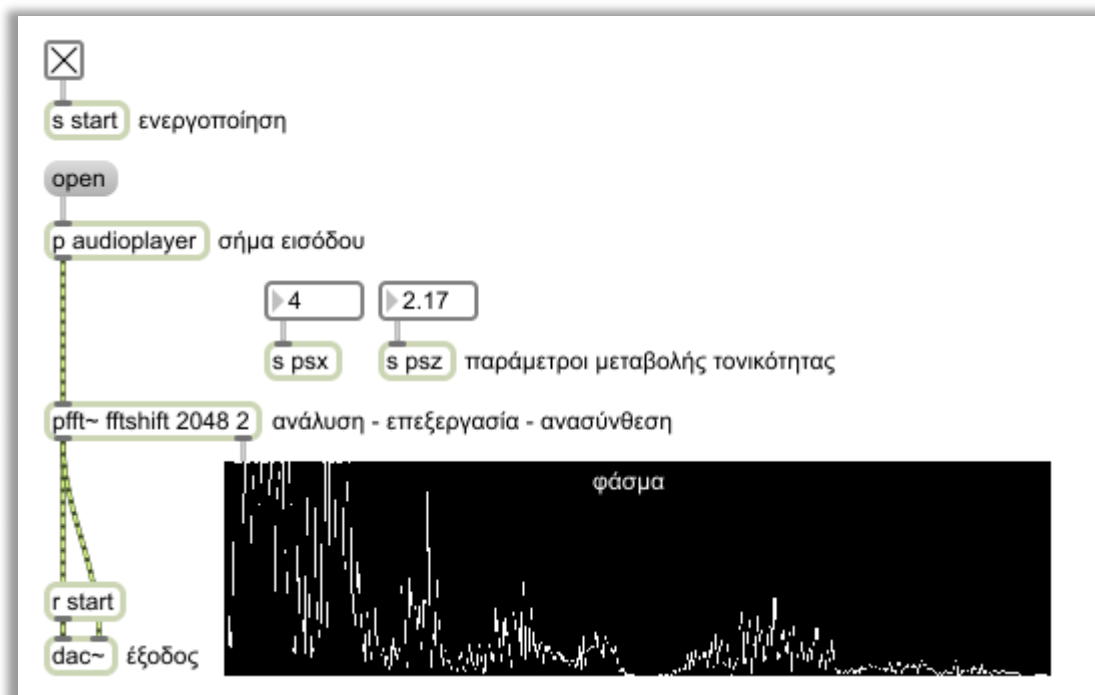
Η Χρονική/Συχνοτική Αναδόμηση Φάσματος (spectral offset) παρέχει στον μουσικό/χρήστη την δυνατότητα να καταγράψει ένα φασματογράφημα από ένα δείγμα ήχου και στη συνέχεια να φορτώσει το φασματογράφημα προς επεξεργασία. Η επεξεργασία δημιουργεί ανεξάρτητες μετατοπίσεις σε κάθε bin/frame του φασματογραφήματος, ενώ παράλληλα αναπαράγει το περιεχόμενο του προς το αντικείμενο της ανασύνθεσης. Η θέση ανάγνωσης του φασματογραφήματος καθώς και η διεργασία της μεταβολής των bins/frames είναι πάντα στην επιλογή του μουσικού/χρήστη.

Τα βασικότερα χαρακτηριστικά της επεξεργασίας είναι η μετατόπιση bins/frames ομαδικά ή ακέραια, η αποτύπωση σε παρελθοντικές θέσεις επίσης ομαδικά ή ακέραια και η αναδίπλωση/παραμόρφωση του φάσματος προς κάθε διάσταση του φασματογραφήματος. Έτσι η εφαρμογή υπάγεται σε μια κατηγορία τεχνικών που έχει να κάνει με την διαμόρφωση θέσης. Τεχνικές τέτοιας μορφής είναι το spectral delay (Droljc 2011)(Boyle 2004) με επιλογή θέσης ως προς το χρόνο και ένα τύπου pitch shift με επιλογή θέσης ως προς τις συχνότητες. Κατά συνέπια τα παραγόμενα ηχητικά αποτελέσματα της εφαρμογής θα έχουν κατά κάποιο τρόπο συγγενικές ιδιότητες με ηχητικά αποτελέσματα που θα προκύψουν με χρήση των παραπάνω αναφερθέντων τεχνικών.

Στα επόμενα κεφάλαια θα γίνει επεξήγηση των παραμέτρων, των τεχνικών αναλυτικά και του τρόπου λειτουργίας κάθε εφαρμογής. Ηχητικά παραδείγματα κατασκευασμένα μόνο μέσω των εφαρμογών και καταγεγραμμένα φάσματα, θα χρησιμοποιηθούν ως επίδειξη των ιδιοτήτων των εφαρμογών για καλύτερη κατανόηση.

### 1.3.2.3 Ένα απλό παράδειγμα.

Για την καλύτερη κατανόηση του τρόπου λογικής της λειτουργίας μίας εφαρμογής φασματικών επεξεργασιών και ως παράδειγμα πριν αναλυθούν οι παράμετροι των κυρίως εφαρμογών, αναπτύχθηκε η εφαρμογή pitchshift\_on\_jitter.maxpat, όπου υλοποιείται μεταβολή τονικότητας σε ένα σήμα εισόδου σε πραγματικό χρόνο.

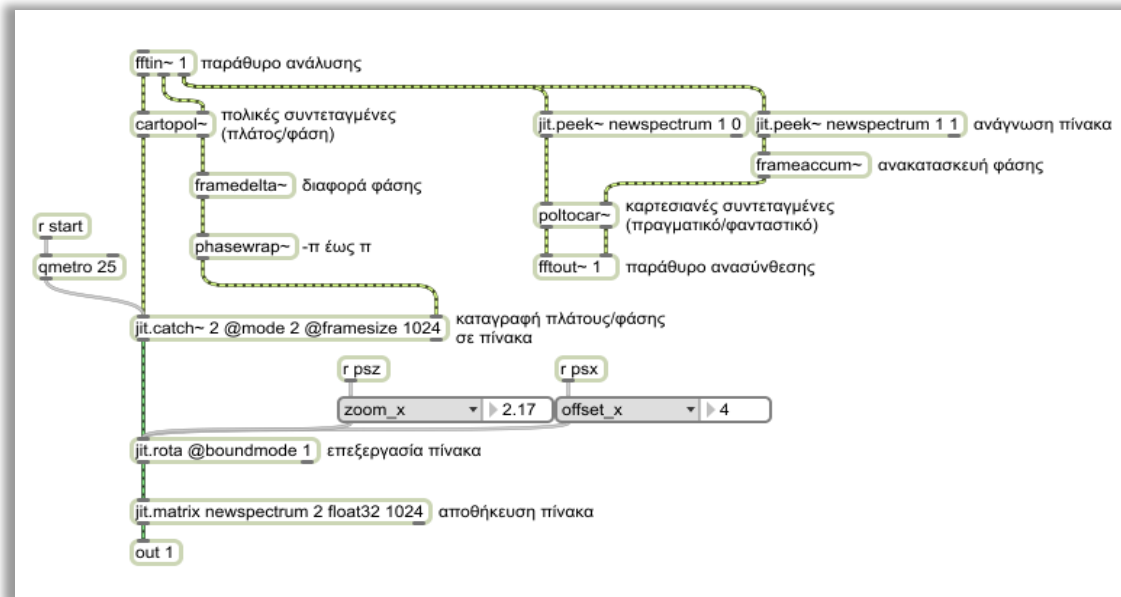


Εικόνα11: Η διεπαφή της εφαρμογής pitchshift\_on\_jitter.maxpat.



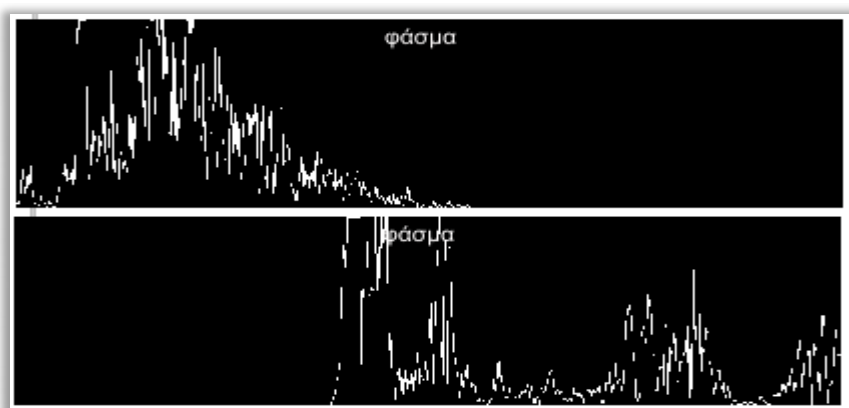
Ένα σήμα εισόδου αναλύεται σε συχνότητα και φάση μέσω των `rfft~` και `fftin~` αντικειμένων ανάλυσης και στη συνέχεια καταγράφεται σε δύο παράλληλους δισδιάστατους πίνακες ίδιου μεγέθους με το `jit.catch~`.

Το περιεχόμενο του κάθε πίνακα επεξεργάζεται το αντικείμενο `jit.rota`, πραγματοποιώντας ομαδοποιημένη μεταβολή των γραμμών του πίνακα δεξιά ή αριστερά και αλλαγή του μήκους τους, με αυξομείωση αυτών. Οι παράμετροι των μεταβολών είναι στην επιλογή του χρήστη και ελέγχονται από δύο αριθμούς.



Εικόνα12: Ο αλγόριθμος επεξεργασίας και ανάλυσης της εφαρμογής `pitchshift_on_jitter.maxpat`.

Στις γραμμές του πίνακα ο άξονας x αντιπροσωπεύει τις συχνότητες, κατά συνέπεια η μεταβολή του επηρεάζει το φάσμα άρα και το τελικό αποτέλεσμα.



Εικόνα13: Το φάσμα χωρίς επεξεργασία (πάνω), το φάσμα μετά την επεξεργασία (κάτω).

Το επεξεργασμένο φάσμα αποθηκεύεται και αναπαράγεται προς ανασύνθεση στα αντικείμενα `jit.matrix` και `jit.peak~`. Το αποτέλεσμα της ανασύνθεσης οδηγείται στην έξοδο του συστήματος για ακρόαση.

## 2. Θεωρητική ανάλυση εφαρμογών

### 2.1 Φασματικό Πάγωμα(spectral freeze)

Η εφαρμογή του φασματικού παγώματος(spectral freezing) λαμβάνει μέσω της ανάλυσης του FFT αλγόριθμου στιγμιαίες φασματικές πληροφορίες από ένα δείγμα ήχου και τις αποθηκεύει προσωρινά σε μια σειρά από πίνακες τιμών δύο διαστάσεων. Οι τιμές του κάθε πίνακα προωθούνται ως δεδομένα στην ανασύνθεση του FFT αλγόριθμου και το αποτέλεσμα συνδέεται στην έξοδο προς ακρόαση.

Ο σκοπός της εφαρμογής είναι να παρέχεται η δυνατότητα να διατηρείται ένα χαρακτηριστικό σημείο της χροιάς του δείγματος ήχου που αναπαράγεται σε πραγματικό χρόνο, να μπορεί να δίνει στον ήχο μια ελεγχόμενη συνέχεια (sustain) και να μπορεί να συνθέτει νέες χροιές με συνδυασμό των ήδη υπαρχόντων.

Καθώς το παραγόμενο αποτέλεσμα προέρχεται από τις πληροφορίες της χροιάς του δείγματος ήχου, οι οποίες είναι στιγμιαίες, δεν υπάρχει μεταφορά πληροφοριών δυναμικής και περιβάλλουσας έντασης, που έχει το δείγμα. Οι πληροφορίες της χροιάς είναι οι φασματικές πληροφορίες που επιστρέφει ο αλγόριθμος FFT κατά την ανάλυση και αποτελούνται από το πραγματικό και το φανταστικό μέρος. Με τροποποίηση αυτών των δεδομένων στο πολικό σύστημα τιμών, παίρνουμε τα πλάτη των συχνοτήτων και τα πλάτη της φάσης. Τα δεδομένα συχνότητας/φάσης είναι συμβατότερα για αποθήκευση και επεξεργασία στα γραφικά αντικείμενα (βλ. Charles 2008).

Καθώς ο FFT αλγόριθμος παραθυροποιεί το δείγμα για να το εισάγει στην ανάλυση, έχει ένα συγκεκριμένο μέγεθος παραθύρου ανάλυσης στο οποίο χωράει το πλάτος και η φάση σε ίσα τμήματα σύμφωνα και με τη θεωρία. Έτσι για την πληροφορία του πλάτους ορίζεται το μισό μέγεθος ανάλυσης και για τη φάση το άλλο μισό. Για μέγεθος ανάλυσης 2048 θα οριστούν 1024 μερίσματα για τις συχνότητες και 1024 για τις φάσεις, όπου το μικρότερο ακέραιο μέρισμα ορίζεται ως ένα bin για συχνότητα και φάση.

Όσο ο αλγόριθμος εκτελείται, επεξεργάζεται μια σειρά από παράθυρα. Εφόσον ο ήχος που αναλύεται δεν είναι περιοδικός, κάθε παράθυρο ανάλυσης θα αποδίδει τελείως διαφορετικά δεδομένα. Παρόλα αυτά, για ευκολία στην επιλογή πιο συγκεκριμένων χαρακτηριστικών στην χροιά του δείγματος η τιμή από κάθε συχνοτικό bin πολλαπλασιάζεται από 0 έως 1, διαμορφώνοντας έτσι εκ νέου τον καταμερισμό ενέργειας στο φάσμα.

Ο έλεγχος του πολλαπλασιασμού γίνεται με την σχεδίαση γραφικής συνάρτησης σε buffer και πολλαπλασιασμό κάθε sample του buffer με κάθε bin του FFT αλγορίθμου. Το μέγεθος του buffer ορίζεται σε millisecond. Για να καλύπτει όλο το φάσμα, δηλαδή όλα τα bins, πρέπει να είναι  $\text{bin size/sampling rate}$ , για παράδειγμα για 2048 μέγεθος παραθύρου ανάλυσης και 44100 sampling rate, το buffer είναι  $1024/44100 \cdot 10^{-3}$  δηλαδή 23,21ms. Η συνάρτηση που σχεδιάζεται στο buffer αποτελεί ένα ηχητικό φίλτρο συχνοτήτων ελεύθερου σχηματισμού.

Στο αμέσως επόμενο στάδιο το αντικείμενο του πακέτου της γραφικής επεξεργασίας Jitter jit.catch, καλείται να καταγράψει τα μερίσματα συχνοτήτων/φάσης που περιέχει ένα παράθυρο ανάλυσης, αφού έχει υποστεί και την αυξομείωση τιμών από το φίλτρο. Κάθε bin αποθηκεύεται προσωρινά στο αντικείμενο αποθήκευσης τιμών jit.matrix, που αρχικοποιείται να έχει δύο πίνακες τους οποίους ονομάζει planes, ένα για τις συχνότητες και ένα για τη φάση, μεγέθους ίσου με το μισό του παραθύρου ανάλυσης. Οπότε για μέγεθος ανάλυσης 2048 κάθε πίνακας/plane θα αποτελείται από 1024 cells. Ο τύπος δεδομένων είναι float32 για συμβατότητα με πληροφορίες σημάτων DSP. Κάθε πίνακας γεμίζει η αντικαθιστά τα cells με παρελθοντικές τιμές του, κάθε φορά που το αντικείμενο καταγραφής jit.catch ενεργοποιείται. Αυτή η δυνατότητα είναι στην επιλογή του χρήστη και όχι αυτόματα.

Οι δυο πίνακες έχουν «φωτογραφίσει» την στιγμιαία κβαντοποιημένη κατάσταση πλάτους/φάσης. Με την χρήση των αντικειμένων γραφημάτων jit.graph και jit.rwindow αποτυπώνεται το περιεχόμενο του πρώτου πίνακα/plane για το πλάτος συχνοτήτων. Το ορατό πλέον περιεχόμενο αποτελεί την συχνοτική απόκριση μίας συγκεκριμένης χρονικής στιγμής του αρχικού ήχου. Για το plane της φάσης, η γραφική αποτύπωση δεν προσφέρει κάτι καθώς μετά και την τροποποίηση τιμών, η φάση παίρνει τιμές  $-\pi$  και  $\pi$  και προσεγγίζει σήμα θορύβου. Η χρησιμότητα του πίνακα της φάσης είναι για την ανασύνθεση όπως έχει αναφερθεί και στη θεωρία.

Οι πίνακες που είναι ενσωματωμένοι στο αντικείμενο αποθήκευσης jit.matrix αποτελούν μια ομάδα. Πολλαπλά αντικείμενα τέτοιου τύπου τοποθετούνται παράλληλα με σκοπό να καταγράφονται πληροφορίες πλάτους συχνοτήτων και φάσης σε διαφορετικές στιγμές του δείγματος το καθένα και να υπάρχει η δυνατότητα να γίνεται συλλογή των προτιμότερων και συνδυαστική μίξη φασματικών δεδομένων.

Τα συνδυασμένα δεδομένα από την ποσοστιαία πρόσθεση των πινάκων αποθηκεύονται προσωρινά σε ένα τελικό αντικείμενο αποθήκευσης με τα αντίστοιχα χαρακτηριστικά.

Δυο αντικείμενα ανάγνωσης `jit.peek~` καλούνται να διαβάσουν τα `planes` του τελικού αντικειμένου αποθήκευσης `jit.matrix`. Το κάθε αντικείμενο ανάγνωσης διαβάζει όλο το `plane` από την πρώτη μέχρι την τελευταία θέση σύμφωνα με την ίδια ταχύτητα που ο αλγόριθμος FFT ανανεώνει κάθε παράθυρο ανάλυσης.

Τα δεδομένα που επιστρέφουν τα `planes` περιέχουν συχνότητα/φάση. Έτσι γίνεται ξανά τροποποίηση τιμών, αυτή τη φορά αντίστροφα από πολικό σε καρτεσιανό σύστημα και συνδέονται στην έξοδο του αλγόριθμου προς ανασύνθεση. Η ανασύνθεση διαβάζει το ίδιο συμβατά δεδομένα με αυτά που δίνει η ανάλυση, με την διαφορά ότι είναι συνεχώς επαναλαμβανόμενα και το πότε θα γίνει η επόμενη μετάβαση είναι στην επιλογή του χρήστη. Το αποτέλεσμα που θα είναι προς ακρόαση θα έχει μια σταθερή χροιά και ακαθόριστη χρονική διάρκεια.

## **2.2 Χρονική/Συχνотική Αναδόμηση Φάσματος (spectral offset).**

Η εφαρμογή της Χρονικής/Συχνотικής Αναδόμησης Φάσματος έχει σαν στόχο να προσφέρει την δυνατότητα επέμβασης απευθείας πάνω στο φασματογράφημα ενός καταγεγραμμένου και αναλυμένου ήχου. Οι πληροφορίες της φασματικής ανάλυσης ενός ήχου αποθηκεύονται σε `planes` δεδομένων που υποβάλλονται σε επεξεργασία και στέλνονται προς ανασύνθεση. Το αποτέλεσμα βγαίνει στην έξοδο. Ο χαρακτήρας της επεξεργασίας είναι η ανεξάρτητη μετατόπιση περιοχών των `planes` σε πραγματικό χρόνο.

Έχει οριστεί ως το κύριο μέσο ελέγχου για την επεξεργασία να είναι το ίδιο το φασματογράφημα. Μέσα στον αλγόριθμο το φασματογράφημα ερμηνεύεται ως ένας τρισδιάστατος πίνακας τιμών. Τα γραφικά αντικείμενα `jit.matrix` που καλούνται να αποθηκεύσουν δεδομένα, είναι αντίστοιχα με αυτά που χρησιμοποιήθηκαν στην εφαρμογή Φασματικού Παγώματος, με τη διαφορά ότι περιέχουν μια παραπάνω διάσταση. Με τις δύο διαστάσεις έχουμε την τιμή πλάτους/φάσης και την συχνότητα, ενώ η τρίτη διάσταση είναι ο χρόνος που ερμηνεύει τις πρώτες δυο διαστάσεις σε διαφορετικές καταστάσεις. Το αντικείμενο ουσιαστικά κρατά ακολουθίες από παράλληλα δισδιάστατα `planes` φάσματος τα οποία ονομάζει `frames`.

### 2.2.1 Εγγραφή Αναπαραγωγή Αποθήκευση.

Το εύρος τιμών για το πλάτος που αποθηκεύεται είναι επίσης τύπου float32 ίδιου με το DSP για να χωρά με ακρίβεια η τροποποιημένη τιμή. Το εύρος συχνοτήτων εξαρτάται από το μέγεθος της ανάλυσης. Μεγαλύτερη ανάλυση αποδίδει χαμηλότερη θεμέλιο συχνότητα στην FFT και καλύτερη ακρίβεια στο φάσμα, σύμφωνα και με την θεωρία. Το εύρος του χρόνου μπορεί να οριστεί στο αντικείμενο και είναι μια ακέραια τιμή όπου το κάθε μέρισμα ερμηνεύεται ως frame. Κάθε frame μπορεί να προσπελαστεί με διαφορετική ταχύτητα. Αυτό συνεπάγεται ότι για διαφορετικές ταχύτητες καταχωρούνται διαφορετικά μερίσματα χρόνου από το αρχικό δείγμα στο τρισδιάστατο plane, τα οποία στη συνέχεια αναπαράγονται από το τρισδιάστατο plane στην έξοδο του συστήματος.

Οι χρήση του αντικειμένου αποθήκευσης jit.matrix προορίζονταν για καταγραφή κινούμενης εικόνας και με τροποποιήσεις χρησιμοποιείται ως φασματογράφος. Για την αποτύπωση του περιεχομένου γίνεται ξανά χρήση ενός αντικειμένου γραφημάτων jit.pwindow ως μέσο απεικόνισης και η έξοδος του συστήματος είναι ένα ορθογώνιο πλαίσιο, με την διαφορά ότι υπάρχει διάδραση καθώς το ίδιο πλαίσιο είναι και το μέσο ελέγχου των λειτουργικών παραμέτρων της εφαρμογής.

Το μέσο απεικόνισης έχει τις ιδιότητες του φασματογραφήματος. Ο οριζόντιος άξονας του πλαισίου είναι ο χρόνος ενώ ο κάθετος άξονας το συχνοτικό εύρος σε γραμμική μορφή. Το χρώμα του κενού πλαισίου είναι μαύρο και αντιπροσωπεύει την έλλειψη ενέργειας στο φάσμα. Όταν το plane περιέχει πληροφορίες, στο πλαίσιο με άσπρο χρώμα αναγράφεται η ύπαρξη ενέργειας και με διαφορετική φωτεινότητα ορίζεται η διακύμανση της έντασης στο φάσμα.

Η διαδικασία της καταγραφής και καταχώρισης των φασματικών πληροφοριών είναι παρόμοια με αυτή που χρησιμοποιείται στην εφαρμογή του Φασματικού Παγώματος. Υπάρχουν όμως ουσιαστικές διαφορές. Στην προηγούμενη εφαρμογή κάθε γέμισμα των planes εξαρτιόταν από τον χρήστη ενώ εδώ έχει οριστεί ένας χρονιστής επιπέδου DSP rate να υπολογίζει κάθε στιγμή που πρέπει να ξεκινά ένα frame του φασματογράφου και ειδοποιεί το αντικείμενο καταγραφής jit.poke~ για την μετάβαση από frame σε frame. Το μέγεθος που ορίζεται είναι το πόσα frames θα υπολογίζει. Η ταχύτητα υπολογισμού εξαρτάται από το μέγεθος ανάλυσης του FFT παραθύρου και το sampling rate. Παράδειγμα για 2048 ανάλυση και 44100 sampling rate, αν ο χρονιστής θέλουμε να υπολογίσει 200 frames θα είναι  $(200 \text{ frames} * 1024 \text{ bins}) / 44100 \text{ sampling rate}$ , δηλαδή η καταγραφή θα είναι 4,6 δευτερόλεπτα για τα 200 frames ή 23 millisecond για το 1 frame. Ταχύτητα που προσεγγίζει τα 43 fps(frames per second).

Το αντικείμενο αποθήκευσης `jit.matrix` ενεργοποιείται από έναν δεύτερο χρονιστή ο οποίος είναι σε `control rate` και κρατά το αντικείμενο σε συνεχή θέση λειτουργίας για να προβάλλονται συνεχώς οι νέες μεταβολές της κατάστασης του `plane` κατά την καταγραφή. Ο σκοπός που εξυπηρετεί ο δεύτερος χρονιστής είναι για την συνεχή ανανέωση της κατάστασης του `plane` και όχι για την διαμόρφωση του περιεχομένου του. Χρησιμοποιείται είτε πρόκειται για καταγραφή, είτε πρόκειται για επεξεργασία.

Το αντικείμενο αποθήκευσης `jit.matrix` αρχικοποιείται να έχει δύο τρισδιάστατα `planes` διαστάσεων μισού μεγέθους ανάλυσης επί το πλήθος των `frames` που ορίζεται από τον χρήστη και με βάθος τιμών `float32`. Οι πληροφορίες που καταχωρούνται στο αντικείμενο παραμένουν αποθηκευμένες στην μνήμη του υπολογιστή προσωρινά. Το αντικείμενο έχει την δυνατότητα να εξάγει και να εισάγει τα `planes`. Έτσι τα φασματογραφήματα μπορούν να αποθηκευτούν μόνιμα και να χρησιμοποιηθούν ξανά. Το εξαγόμενο αρχείο έχει οριστεί από το πακέτο `Jitter`, έχει κατάληξη `.jxf` και η μορφή του βασίζεται στον τύπο αρχείων `IFF(Interchange File Format)`. Σε ότι αφορά στα δεδομένα που περιέχει για τη χρήση ως φασματογράφημα στην συγκεκριμένη εφαρμογή, αυτά είναι τύπου: `big-endian, 32-bit float, 2 planes, 2 dims`.

Ο αλγόριθμος `FFT` λειτουργεί συνεχόμενα για τα σήματα που δέχεται στην είσοδο του. Το πότε θα ξεκινήσει η καταχώριση στα `Jitter` αντικείμενα εξαρτάται από τον χρήστη. Μέσω ενός διακόπτη ανοίγει και κλείνει η ροή `FFT` δεδομένων προς το αντικείμενο καταγραφής. Εφόσον ενεργοποιηθεί, η καταγραφή θα ξεκινήσει από το πρώτο `frame`. Όταν φτάσει ο δείκτης του χρονιστή στο τελευταίο `frame`, παράδειγμα στο 200ο `frame`, θα επιστρέψει ξανά στην αρχή και η εγγραφή θα αντικαταστήσει τις προηγούμενες καταχωρήσεις.

Μετά από μια επιτυχημένη καταχώριση δεδομένων στα `planes`, τα αντικείμενα ανάγνωσης αναπαράγουν το περιεχόμενο καθενός. Η ταχύτητα για κάθε `frame` μπορεί να είναι ίδια με την ταχύτητα εγγραφής, αλλά μπορεί να χρησιμοποιηθεί και άλλη συνάρτηση που να ορίζει πως θα διαβάζονται τα `frames`. Το σύστημα αναπαραγωγής των `planes` μπορεί να είναι σε ανεξάρτητο `patch`, εφόσον πρόκειται για ήδη καταγεγραμμένα δεδομένα σε αρχείο `.jxf` που φορτώνεται ξανά μόνο για αναπαραγωγή.

### 2.2.2 Επεξεργασία και Αναπαραγωγή.

Η διαδικασία της επεξεργασίας μπορεί να εφαρμοστεί είτε πρόκειται για ένα εγγεγραμμένο plane είτε για ένα αποθηκευμένο που επαναφορτώνεται. Όπως έχει αναφερθεί πιο πριν το ένα plane αφορά συχνοτικά πλάτη και το άλλο φάση. Κατά την επεξεργασία προβάλλεται και χρησιμοποιείται μόνο εκείνο το plane που αφορά τις συχνότητες όπου η τροποποίηση του έχει άμεση επίπτωση στο αποτέλεσμα. Το plane της φάσης χρησιμοποιείται μόνο για την ανασύνθεση και αναπαραγωγή.

Η βασική λειτουργία της επεξεργασίας όπως είναι και ο χαρακτήρας της εφαρμογής, είναι η επανατοποθέτηση τμημάτων του φάσματος, με αποτέλεσμα αυτό να αναδομείται. Τα μερίσματα που εφαρμόζεται αυτή η μεταβολή είναι ως προς το χρόνο και τη συχνότητα, δηλαδή για κάθε ένα frame και για κάθε ένα συχνοτικό bin.

Η μέθοδος ελέγχου για την μεταβολή των μερισμάτων χρόνου/συχνότητας υλοποιείται μέσω δύο συναρτήσεων. Κάθε συνάρτηση αποθηκεύεται σε δισδιάστατο πίνακα τύπου buffer, το μέγεθος του οποίου ορίζεται σύμφωνα με το μέγεθος της αντίστοιχης διάστασης του χρόνου και της συχνότητας που έχει το πλαίσιο του φασματογραφήματος. Παράδειγμα για 200x1024 ανάλυση πλαισίου, το buffer για τα frames είναι  $200/44,1=4,53$  ms και για το buffer των συχνοτήτων είναι  $1024/44,1=23,21$  ms.

Αφού σχεδιαστεί και καταχωρηθεί η κάθε συνάρτηση στα buffer, εκτελούνται πράξεις στις τιμές κάθε sample για να οριστεί το σωστό εύρος τιμών. Αυτό τροποποιείται από -1 έως 1 που αρχικά λειτουργεί το buffer, σε -x έως x. Η μέγιστη κατά απόλυτο τιμή του x είναι ένα πολλαπλάσιο του μεγέθους της διάστασης του πλαισίου και η ελάχιστη το 0. Έτσι όταν η μεταβλητή x είναι 0 δεν υπάρχει μεταβολή. Όταν το x παίρνει τιμή ίση με το μέγεθος της διάστασης, η μεταβολή είναι ίση κατά ένα πλαίσιο. Το πρόσημο ορίζει την κατεύθυνση της μεταβολής. Η τιμές που παίρνει η μεταβλητή είναι δεκαδικές και σύμφωνα με το βάθος τιμών που ορίζει το DSP. Έτσι η ακρίβεια μπορεί να είναι μεγάλη και οι συναρτήσεις σύνθετες.

Η καταχωρημένη και τροποποιημένη στο εύρος τιμών συνάρτηση, διαβάζεται από το αντικείμενο του πακέτου γραφικής επεξεργασίας Jitter jit.scanoff, το οποίο εκτελεί την διεργασία μετακίνησης των cells του plane για τις συχνοτικές πληροφορίες, κατά όσο ορίζει η συνάρτηση. Αυτό εφαρμόζεται και για τις δύο συναρτήσεις των bins και των frames, γι' αυτό χρησιμοποιούνται δύο τέτοια αντικείμενα μετακίνησης. Παράδειγμα, για 200x1024 ανάλυση, αν η συνάρτηση για μεταβολή των συχνοτήτων ορίζει τιμή 0,5 στα πρώτα 512 bins και τιμή 0 στα υπόλοιπα 512 bins, τότε το μισό φάσμα θα έχει μετακινηθεί στον άξονα των συχνοτήτων κατά μισό πλαίσιο και το άλλο μισό φάσμα θα παραμείνει εκεί που ήταν.

Μια σειρά από αντικείμενα της γραφικής επεξεργασίας με έτοιμα εργαλεία, χρησιμοποιούνται για διευκόλυνση στις επεξεργασίες γενικότερα (jit.unpack jit.xfade jit.rota jit.op). Οι βασικότερες λειτουργίες της επεξεργασίας που εκτελούνται είναι η μετακίνηση του plane ομαδοποιημένα, η γραμμική παραμόρφωση κάθε άξονα του plane και η αναδίπλωση – επέκταση του plane για κάθε διάσταση. Ενώ κάθε λειτουργία του αντικειμένου έχει συγκεκριμένο σκοπό και απλή εφαρμογή, ο συνδυασμός πολλών λειτουργιών μαζί και με τις υπόλοιπες επεξεργασίες, καταλήγουν σε ένα πιο σύνθετο σύστημα.

Ένα τμήμα του αλγόριθμου επεξεργασίας είναι ένα σύστημα ανάδρασης αρχιτεκτονικής IIR. Εφαρμόζεται πάνω στο τρέχον νήμα των φασματικών πληροφοριών ανατροφοδοτώντας ένα υποπολλαπλάσιο του εαυτού του. Έτσι όταν πραγματοποιηθεί η επόμενη μεταβολή λόγω μιας επεξεργασίας στο περιεχόμενο, θα ξεκινήσει μια σταδιακή μετάβαση από την παρελθοντική κατάσταση στην τρέχουσα, με χρόνο απόσβεσης που εξαρτάται από το ποσοστό του ανατροφοδοτούμενου νήματος. Για ποσοστό 0% δεν υφίσταται μεταβολή ενώ για 99% η μεταβολή γίνεται με τη μέγιστη χρονική διάρκεια.

Κατά το τελικό στάδιο, στον αλγόριθμο επεξεργασίας πραγματοποιείται μια εξομάλυνση της μεταβολής από frame σε frame. Ενώ κάθε ένα frame δεν έχει περιβάλλουσα έντασης ως προς το χρόνο, η μετακίνηση του δείκτη ανάγνωσης στα διαδοχικά frames αποκαλύπτει διαφορές ενέργειας και έντασης. Γι' αυτό το λόγο γίνεται μια ποσοστιαία πρόσθεση – μίξη του παρόντος frame ( $v$ ) εκεί που βρίσκεται ο δείκτης, με το διαδοχικό του ( $v+1$ ). Το αποτέλεσμα που προκύπτει είναι ένα νέο frame που δεν υπάρχει στο αρχικό plane. Έτσι χρησιμοποιείται ένα νέο plane για να το αποθηκεύσει, όπου στην διάσταση των frames αποτελείται από ένα μόνο frame και στη διάσταση των bins είναι ίδιο με αυτό του αρχικού plane.

Με αυτή τη μέθοδο αναπαραγωγής του φάσματος δεν αναπαράγεται το αυθεντικό περιεχόμενο, αλλά μια ανασύνθεσή του. Η FFT ανασύνθεση διαβάζει ένα μόνο plane, που το περιεχόμενο του αλλάζει κατά την επιλογή του χρήστη.



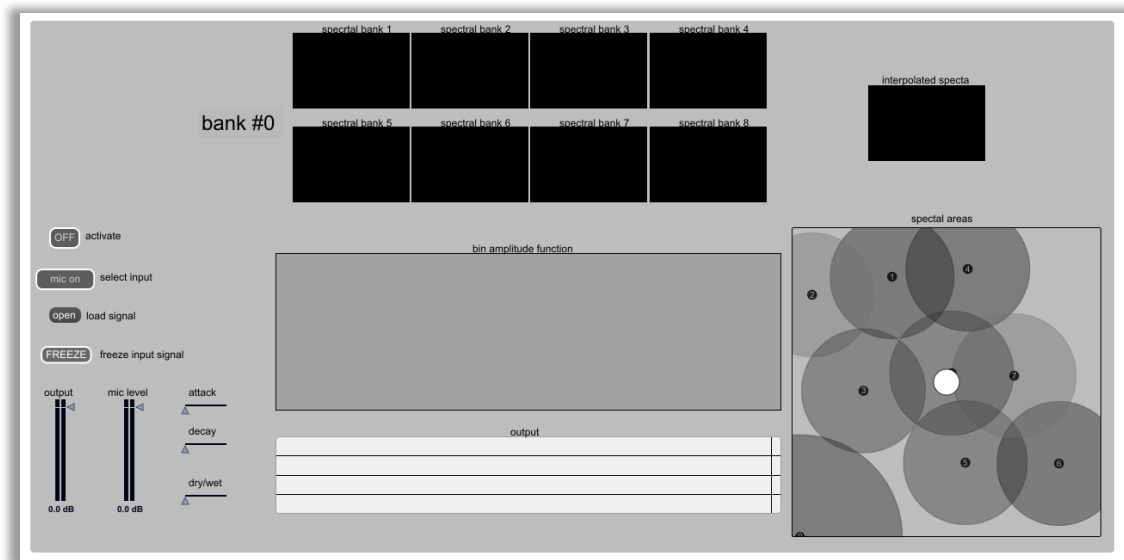
### 3. Τεχνική ανάλυση των εφαρμογών.

Σε αυτή την ενότητα θα εξηγηθεί για κάθε εφαρμογή, η διεπαφή, οι παράμετροι ελέγχου των λειτουργιών και οι αλγόριθμοι όπως αυτά υλοποιήθηκαν στο περιβάλλον Max/MSP Jitter.

#### 3.1 Φασματικό πάγωμα.

Η εφαρμογή του Φασματικού παγώματος (spectral freeze) αποτελείται από το κυρίως patch της διεπαφής (main\_spectralfreeze.maxpat), αυτό της φασματικής επεξεργασίας (fftbinfx4.maxpat) και από άλλα subpatches που είναι ενθυλακωμένα μέσα στα κυρίως patch.

##### 3.1.1 Η διεπαφή και οι παράμετροι του Φασματικού παγώματος.



Εικόνα14: Η διεπαφή της εφαρμογής του Φασματικού Παγώματος.

Όταν ανοίξει η εφαρμογή εμφανίζεται το παράθυρο της διεπαφής, μέσα στο οποίο υπάρχουν όλες οι παράμετροι που απαιτούνται για να χρησιμοποιηθεί η εφαρμογή. Τα μέσα που χρησιμοποιούνται είναι ο κέρσορας και το πληκτρολόγιο για συντομεύσεις.

Στα αριστερά του παραθύρου υπάρχει μια σειρά από αντικείμενα εντολών και ελέγχου. Πρώτο από πάνω είναι το αντικείμενο που στέλνει την εντολή για ενεργοποίηση και απενεργοποίηση όλης της εφαρμογής, το οποίο είναι τύπου διακόπτη (toggle) με δύο καταστάσεις ON και OFF. Ένα δεύτερο αντικείμενο τύπου διακόπτη (toggle) με δύο καταστάσεις mic on και playback, στέλνει την εντολή για την επιλογή του σήματος εισόδου, επιλέγοντας έτσι αν θα είναι από εξωτερική συσκευή μικροφώνου η καλωδίου και αν θα είναι για δείγμα ήχου που θα φορτώνεται. Αμέσως μετά υπάρχει αντικείμενο τύπου button το οποίο στέλνει την εντολή για να φορτωθεί το δείγμα ήχου από τον υπολογιστή.

Το επόμενο αντικείμενο FREEZE είναι επίσης τύπου button και στέλνει την εντολή για να καταγραφεί η στιγμιαία φασματική κατάσταση του επιλεγμένου σήματος εισόδου. Η χρήση της συγκεκριμένης εντολής έχει δεσμευτεί με το πλήκτρο space του πληκτρολογίου για λόγους πρακτικότητας.

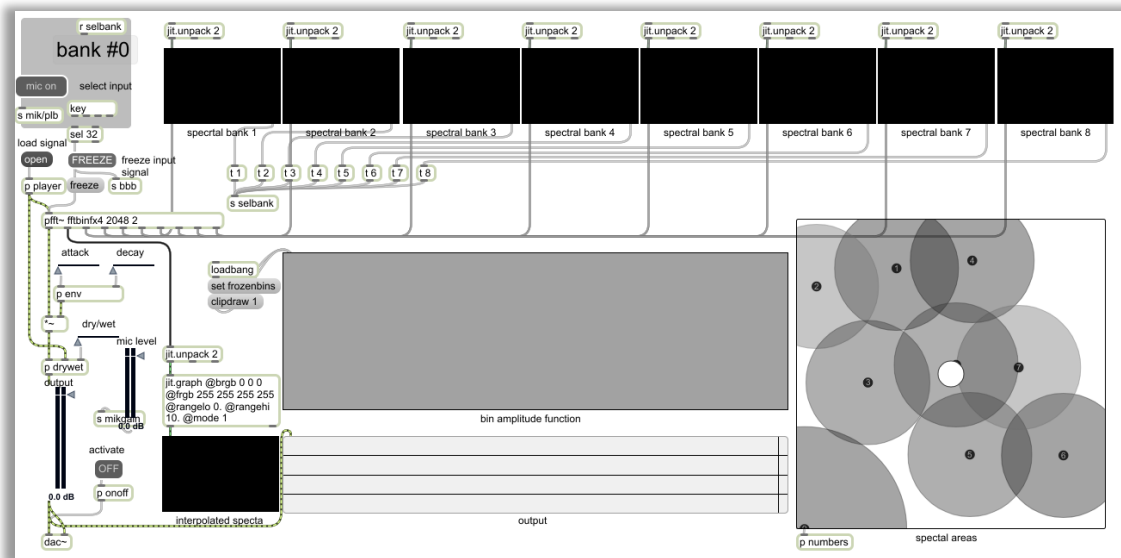
Στο κάτω μέρος αριστερά του παραθύρου υπάρχουν τα faders για το σήμα εξόδου και για την στάθμη του σήματος εισόδου όταν πρόκειται για εξωτερική συσκευή. Δίπλα υπάρχουν τρία οριζόντια sliders. Τα δύο πρώτα ορίζουν τις παραμέτρους attack και decay της περιβάλλουσας έντασης που παίρνει κατά την δημιουργία του ένα φασματικό παράθυρο και το τρίτο ελέγχει το ποσοστό των σημάτων που θα οδηγούνται στην έξοδο, dry/wet.

Κεντρικά στο πάνω μέρος του παραθύρου της εφαρμογής υπάρχει μια συστοιχία 8 πλαισίων. Κάθε πλαίσιο αποτελεί μια μνήμη προσωρινής αποθήκευσης τύπου bank κάθε φασματικής κατάστασης που δημιουργείται με την εντολή FREEZE. Ο χρήστης επιλέγει το πλαίσιο πατώντας τον κέρσορα μια φορά πάνω του. Μόλις καταγραφεί η επόμενη φασματική πληροφορία θα εμφανιστεί στο επιλεγμένο πλαίσιο η συχνοτική απόκριση.

Πιο κάτω στο κέντρο του παραθύρου, υπάρχει το πλαίσιο στο οποίο σχεδιάζεται η συνάρτηση τύπου φίλτρου συχνοτήτων, διαμορφώνοντας έτσι την τιμή κάθε συχνοτικού bin του σήματος εισόδου πριν την καταγραφή τις στιγμιαίας φασματικής του κατάστασης. Από κάτω ακριβώς υπάρχει το πλαίσιο ενός φασματογραφήματος που αποτυπώνει το συνολικό αποτέλεσμα της εξόδου του ήχου της εφαρμογής.

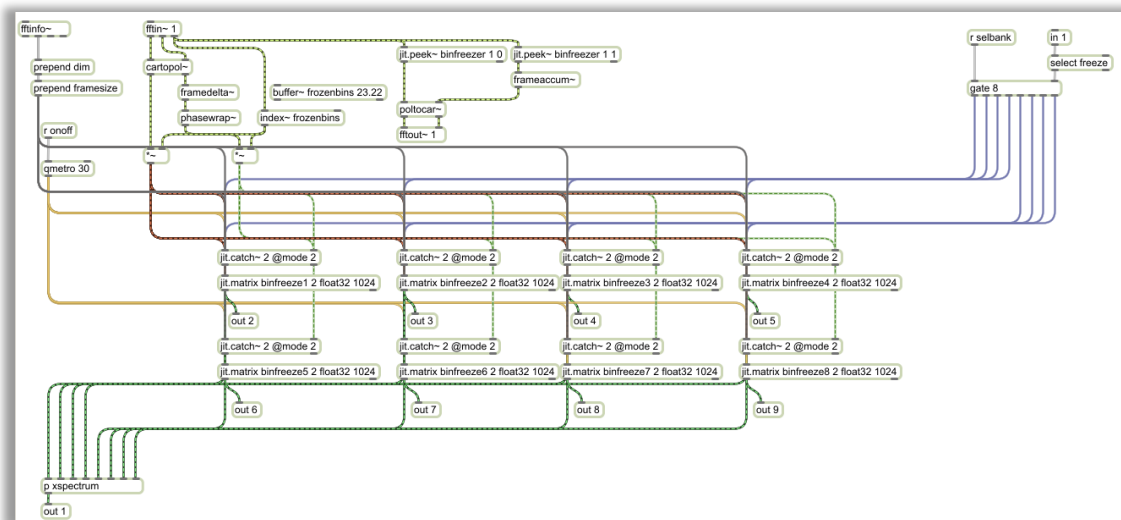
Πάνω δεξιά στο παράθυρο υπάρχει ένα πλαίσιο στο οποίο αποτυπώνεται απλώς η συχνοτική απόκριση από το άθροισμα των προσωρινά αποθηκευμένων φασμάτων. Κάτω υπάρχει ένα αντικείμενο που ορίζει μέσω γραφημάτων, τις φασματικές περιοχές, το πόσο και που επικαλύπτονται μεταξύ τους. Κάθε κύκλος αντιπροσωπεύει μια μνήμη προσωρινής αποθήκευσης και η ακτίνα κάθε κύκλου την ένταση που θα ακούγεται η φασματική πληροφορία κάθε μνήμης, όπου το κέντρο είναι η μέγιστη ένταση και η περιφέρεια, η ελάχιστη. Κάθε γραφικό αντικείμενο τοποθετείται ελεύθερα μέσα στο πλαίσιο. Η θέση του κέρσορα ορίζει το αποτέλεσμα στη συνακρόαση όλων των φασματικών περιοχών.

### 3.1.2 Ανάλυση των αλγορίθμων του Φασματικού Παγώματος.



Εικόνα15: Ο κεντρικός αλγόριθμος της εφαρμογής.

Ο κεντρικός αλγόριθμος συγκροτεί τη διεπαφή, τις παραμέτρους και τους υπο-αλγόριθμους, σε ένα παράθυρο. Το τμήμα της επεξεργασίας και σύνθεσης του αποτελέσματος, πραγματοποιείται στο αντικείμενο 'pfft~ fftbinfx4 2048 2', που αποτελεί το σύστημα του FFT αλγορίθμου και τα συστήματα κατασκευής και σύνθεσης των φασματικών πληροφοριών. Ο αλγόριθμος φορτώνεται ως πηγαίος κώδικας από το αρχείο `fftbinfx4.maxpat` και αρχικοποιούνται τα στοιχεία του FFT, δηλαδή το μέγεθος παραθύρου ίσο με 2048 και το overlap factor ίσο με 2.

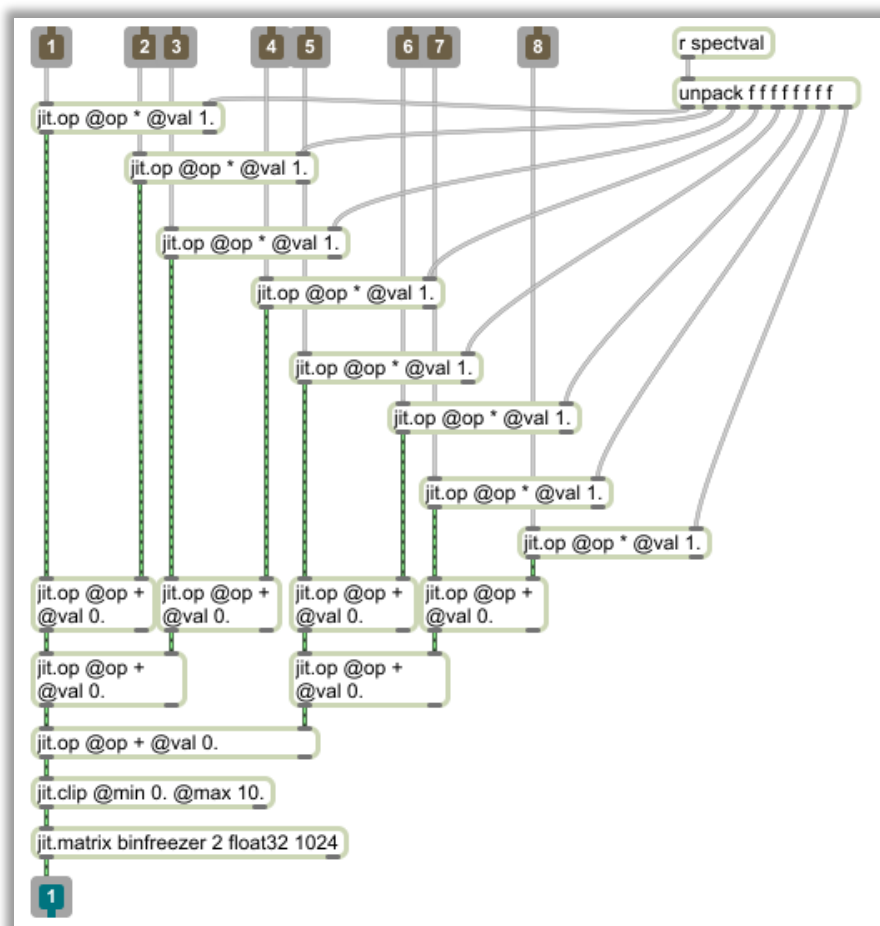


Εικόνα16: Ο αλγόριθμος FFT.

Στην είσοδο του, το αντικείμενο `rfft~` δέχεται το σήμα εισόδου και την εντολή “freeze”. Το παράθυρο ανάλυσης που δημιουργείται, δέχεται τροποποιήσεις, όπως έχει αναφερθεί και στη θεωρία. Το αντικείμενο `cartopol~` μετατρέπει τις φασματικές πληροφορίες σε συχνότητα/φάση. Η φάση υπολογίζεται και αναδιπλώνεται με τα `framedelta~ phasewrap~`. Στη συνέχεια διαμορφώνονται οι τιμές των bins μέσω της συνάρτησης που έχει οριστεί ως ηχητικό φίλτρο ελευθέρου σχηματισμού και η οποία έχει φορτωθεί στο ‘buffer~ frozenbins 23.22’. Η διαμόρφωση πραγματοποιείται μέσω πολλαπλασιασμού του `index~` του `buffer~` με τα bins.

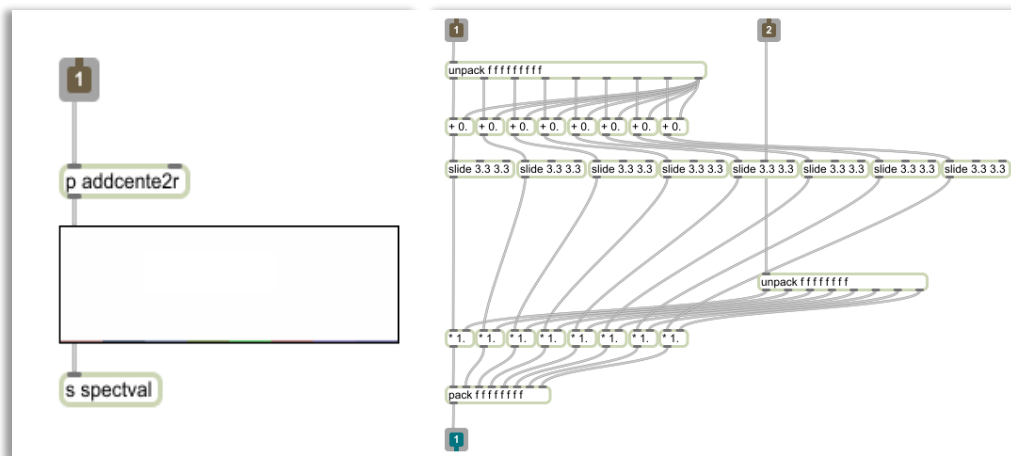
Αφού έχει οριστεί η επιθυμητή τιμή, τα bins στέλνονται ταυτόχρονα σε οκτώ ζεύγη αντικειμένων `jit.catch~` και `jit.matrix`. Επάνω δεξιά υπάρχει το αντικείμενο `gate` που διαλέγεται σύμφωνα με την επιλεγμένη προσωρινή μνήμη που θα δρομολογήσει την εντολή “freeze”. Κάθε φορά δρομολογείται μόνο μια εντολή σε ένα `jit.catch~`.

Όταν δοθεί με επιτυχία η εντολή, το `jit.catch~` καλεί το `jit.matrix` να καταγράψει την κατάσταση των bins τη στιγμή που έλαβε την εντολή. Το περιεχόμενο κάθε `jit.matrix` προωθείται στην έξοδο του συστήματος και στον υπο-αλγόριθμο `xspectrum`.



Εικόνα17: Εντός του `xspectrum`.

Για κάθε ένα jit.matrix πραγματοποιείται ένας πολλαπλασιασμός που διαμορφώνει το ποσοστό συμμετοχής του στο τελικό αποτέλεσμα. Ο κάθε πολλαπλασιαστής προκύπτει από το μέσω διεπαφής των φασματικών περιοχών.



Εικόνα18: Οι μεταβλητές των φασματικών περιοχών.

Για να μπορεί να πραγματοποιείται ανά πάσα στιγμή ο πολλαπλασιασμός, υπάρχει συνεχής ροή του περιεχομένου κάθε jit.matrix. Αυτό πραγματοποιείται με το αντικείμενο qmetro που ενημερώνει συνεχώς την έξοδο του jit.matrix.

Όταν οριστεί και το ποσοστό κάθε φασματικής περιοχής, γίνεται πρόσθεση και των οκτώ jit.matrix σε ένα τελικό αντικείμενο jit.matrix, που θα αποτελεί το φάσμα του ήχου κατά την ανασύνθεση.

Επομένως μέσα στον αλγόριθμο rfft~ (Εικ.16) καλούνται τα αντικείμενα jit.peek~ να διαβάζουν το οποιοδήποτε περιεχόμενο προκύπτει στο τελικό jit.matrix. Τα δεδομένα που επιστρέφουν από την ανάγνωση είναι τα καινούργια bins που δημιούργησε η πρόσθεση, τα οποία δέχονται τις αντίστοιχες τροποποιήσεις roltoacar~ και frameaccum~ και περνάνε στην έξοδο του rfft~ για την ανασύνθεσή τους σε σήμα ήχου στο πεδίο του χρόνου.

Η έξοδοι του rfft~ (Εικ.15) αποτελούνται από το ηχητικό αποτέλεσμα που είναι το σήμα του ήχου και τα οκτώ σήματα φασματικών πληροφοριών. Τα σήματα των φασματικών πληροφοριών συνδέονται σε μια σειρά από αντικείμενα jit.unpack, jit.graph, jit.pwindow, με σκοπό την προβολή του φάσματος. Το φάσμα προβάλλεται στα πλαίσια του κεντρικού αλγόριθμου με την μορφή γραφήματος συχνοτικής απόκρισης.

Ένα σήμα jitter όπως έχει κατασκευαστεί στον συγκεκριμένο αλγόριθμο, αποτελείται από την απόκριση των συχνοτήτων και φάσης. Το jit.unpack τα ξεχωρίζει και προωθεί μόνο το συχνοτικό φάσμα. Τα jit.graph και jit.pwindow ορίζουν πως θα προβάλλεται το φάσμα.

Το σήμα του ήχου που έχει δημιουργηθεί από την επεξεργασία δέχεται μια περιβάλλουσα έντασης μόνιμου sustain, με παραμέτρους μόνο attack και decay. Αυτές οι παράμετροι ενεργοποιούνται κάθε φορά που πρόκειται να υπάρξει δημιουργία σήματος φασματικών δεδομένων κατά την επεξεργασία και συγχρονίζονται με την εντολή 'freeze' που προκαλεί την δημιουργία αυτών των φασματικών δεδομένων.

Πριν το τελικό fader για την έξοδο του ήχου, το σήμα της επεξεργασίας δέχεται μια μίξη τύπου dry/wet με το σήμα εισόδου και με δυνατότητα επιλογής του ποσοστού του. Στη συνέχεια παίρνει την τελική στάθμη από το fader εξόδου και στέλνεται στο digital to analog και στο φασματογράφημα τις διεπαφής.

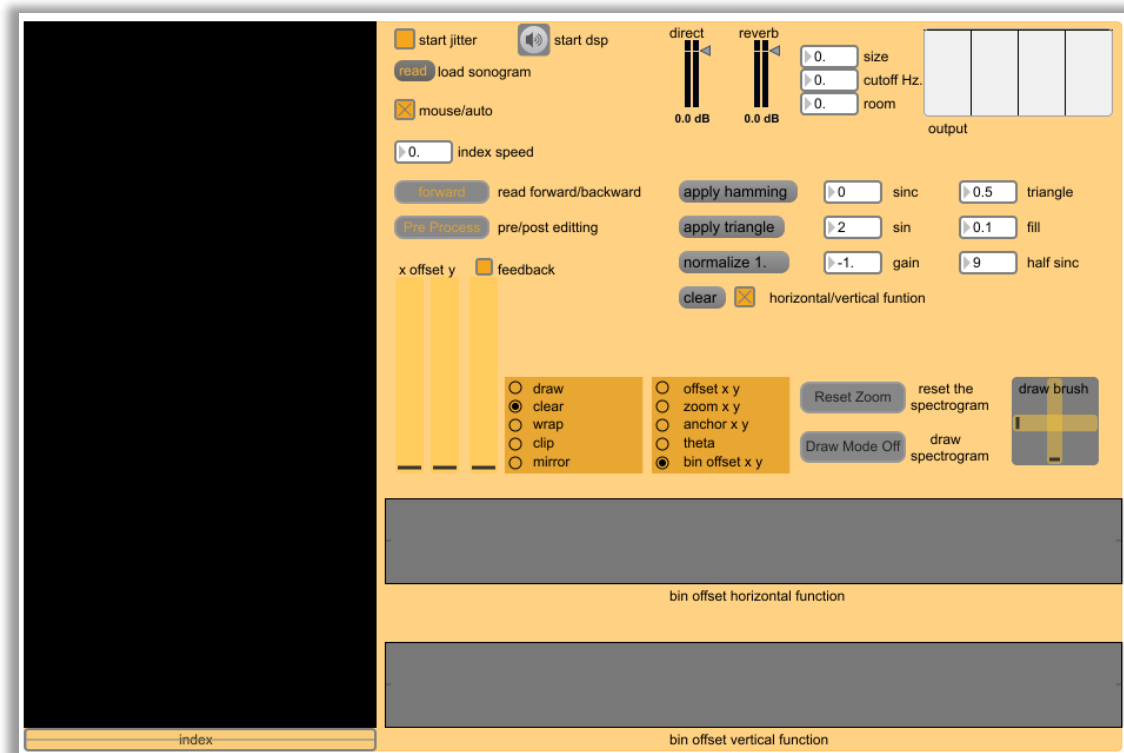
### **3.2 Χρονική/Συχνотική Αναδόμηση Φάσματος.**

Η εφαρμογή για τη Χρονική/Συχνотική Αναδόμηση Φάσματος (spectral offset) αποτελείται από τους αλγόριθμους για την καταγραφή και τους αλγόριθμους για την επεξεργασία.

Ο αλγόριθμος της καταγραφής είναι ένα απλούστερο σύστημα, με μόνη λειτουργία την εγγραφή ηχητικού σήματος σε plane φασματογραφήματος και την εξαγωγή του plane σε ανεξάρτητο αρχείο. Αποτελείται από το patch jrecplay.maxpat για τη διεπαφή και το jfftrecplay.maxpat για τον αλγόριθμο καταγραφής.

Ο αλγόριθμος της επεξεργασίας είναι το κυριότερο μέρος της εφαρμογής και εκτελεί μια ομάδα λειτουργιών που εφαρμόζουν τροποποιήσεις στο plane που εξήγαγε ο αλγόριθμος καταγραφής. Αποτελείται από το patch main\_spectraloffset.maxpat που είναι ο κεντρικός αλγόριθμος μαζί με την διεπαφή. Μέσα στον κεντρικό αλγόριθμο υπάρχουν ενθυλακωμένα subpatches με λειτουργίες επεξεργασίας σημάτων και μεταβλητών. Αυτά είναι το patch jitfft.maxpat για την ανασύνθεση και ένα patch που δημιουργεί εφέ reverb στο σήμα εξόδου.

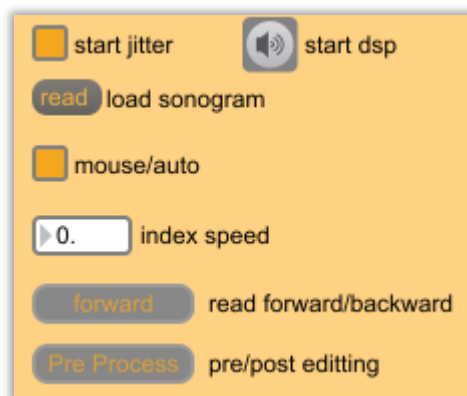
### 3.2.1 Η διεπαφή και οι παράμετροι της Χρονικής/Συχνοτικής Αναδόμησης Φάσματος.



Εικόνα19: Η διεπαφή της εφαρμογής.

Ανοίγοντας την εφαρμογή εμφανίζεται το παράθυρο της διεπαφής και οι παράμετροι των λειτουργιών των αλγορίθμων σε ένα σύνολο.

Στα αριστερά του παραθύρου υπάρχει ένα ορθογωνικό πλαίσιο που αντιπροσωπεύει το διαδραστικό φασματογράφημα. Το φασματογράφημα αποτελεί ταυτόχρονα είσοδο και έξοδο του συστήματος των αλγορίθμων, καθώς είναι το μέσο χειρισμού της άμεσης επεξεργασίας και το μέσο προβολής του αποτελέσματος αυτής. Το ηχητικό αποτέλεσμα στην έξοδο του συστήματος εξαρτάται από την θέση ανάγνωσης του φασματογραφήματος. Κάτω ακριβώς από το πλαίσιο του φασματογραφήματος υπάρχει ένα slider που ορίζει το τμήμα του, που θα υποβάλλεται σε ανάγνωση.



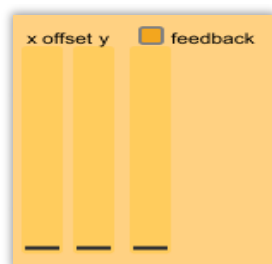
Εικόνα20: Παράμετροι ελέγχου αρχικοποίησης και επεξεργασίας.

Πιο κεντρικά στο παράθυρο της εφαρμογής υπάρχει μια ομάδα από εντολές, μεταβλητές και διακόπτες/κουμπιά, που ορίζουν τις παραμέτρους της επεξεργασίας που θα ελέγχεται από το πλαίσιο του φασματογραφήματος.

Οι διακόπτες 'start jitter', 'start dsp' ενεργοποιούν τη ροή των δεδομένων jitter και audio αντίστοιχα. Η εντολή 'read load sonogram' φορτώνει στην προσωρινή μνήμη της εφαρμογής, ένα plane φασματογραφήματος που έχει δημιουργηθεί από τον αλγόριθμο καταγραφής.

Ο διακόπτης 'mouse/auto' ορίζει αν θα γίνεται αυτόματη ανάγνωση ή ελεγχόμενη από τον κέρσορα, στο φασματογράφημα. Από κάτω βρίσκεται η μεταβλητή 'index speed' η οποία ορίζει πόσο γρήγορα θα διαβάζεται το επιλεγμένο τμήμα του φασματογραφήματος και η εντολή forward/backward που ορίζει την κατεύθυνση του αυτόματου δείκτη ανάγνωσης.

Η εντολή 'Pre Process/Post Process' ελέγχει το στάδιο που θα εφαρμόζεται η επεξεργασία στο φασματογράφημα. Συγκεκριμένα υπάρχουν δύο βασικές ομάδες επεξεργασίας. Αυτή της μετατόπισης χρονικά/συχνοτικά των bins των περιοχών του φασματογραφήματος και αυτή της αλλοίωσης και μετατόπισης του φασματογραφήματος ως συνόλου. Η εντολή διαλέγει ποιά επεξεργασία θα γίνεται πρώτη και ποιά δεύτερη.

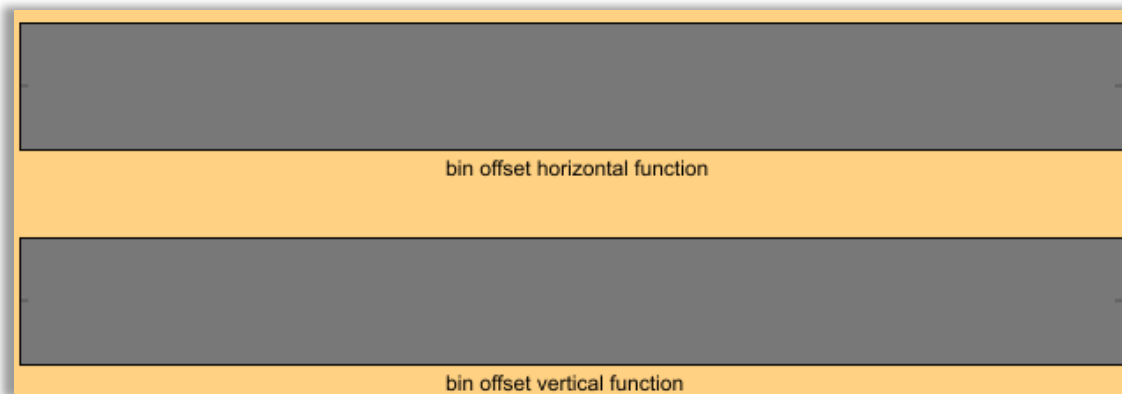


Εικόνα21: Παράμετροι ελέγχου μετατόπισης και ανάδρασης.

Τα δύο πρώτα sliders ορίζουν σε πιο ποσοστό θα εφαρμόζεται η μετατόπιση των bins. Όπου ελάχιστη τιμή σημαίνει καθόλου μετατόπιση και μέγιστη είναι η μεταβολή κατά πολλά πλαίσια. Όταν ένα bin φτάσει στο όριο του πλαισίου του φασματογραφήματος θα εμφανιστεί από την αντίθετη πλευρά έχοντας ήδη διανύσει το μήκος ενός πλαισίου και συνεχίζει σύμφωνα με τον χειρισμό από τον μουσικό/χρήστη, την πορεία του.

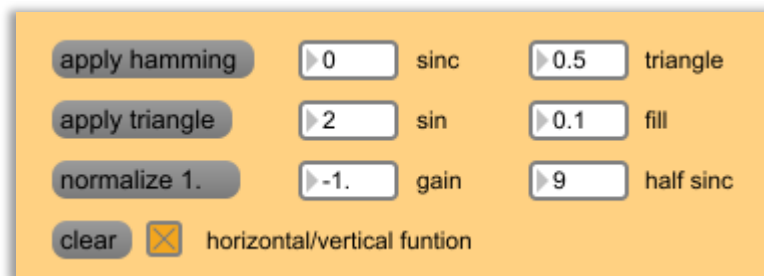
Το δεξί slider ορίζει το ποσοστό που θα ανατροφοδοτείται το jitter σήμα, κατά το τελικό στάδιο της επεξεργασίας, εφαρμόζοντας ένα είδος εφέ τύπου blur οπτικά. Ταυτόχρονα δημιουργεί και μία κατάσταση που δίνει χρόνο απόσβεσης σε κάθε τροποποίηση του περιεχομένου του φασματογραφήματος.





Εικόνα22: Οι συναρτήσεις μεταβολής των bins.

Για την μορφή που θα συμβεί μια μεταβολή στα bins, σχεδιάζεται μια συνάρτηση για κάθε άξονα(χρόνου/συχνότητων). Τα δύο πλαίσια της εικόνας 22 είναι οι μνήμες τύπου buffer που σχεδιάζεται και φορτώνεται η κάθε συνάρτηση.



Εικόνα23: Εντολές συναρτήσεων.

Υπάρχουν οι εντολές που γεμίζουν κάθε buffer με έτοιμες συναρτήσεις. Αυτές είναι ημιτονοειδούς τύπου, τριγωνικές και άλλες. Μπορούν να φορτώνονται οι έτοιμες συναρτήσεις και παράλληλα να σχεδιάζονται με τον κέρσορα επιπλέον συντεταγμένες στο buffer.



Εικόνα24: Εντολές επεξεργασίας του συνόλου του φασματογραφήματος.

Τα αντικείμενα που εφαρμόζουν τις επεξεργασίες δέχονται μια σειρά από εντολές. Για την ευκολότερη χρήση τους ομαδοποιήθηκαν κατά αυτόν τον τρόπο στην εικόνα 24.

Στο πρώτο πλαίσιο οι εντολές draw, clear, wrap, clip, mirror, ορίζουν το πώς θα αποτυπώνεται το φασματογράφημα. Εδώ κάθε εντολή έχει την ιδιότητα να ορίζει το περιβάλλον που θα περιέχει το φασματογράφημα και είναι στατικού τύπου.

Το 'draw' ορίζει πως κάθε μετατροπή θα παραμένει για πάντα αποτυπωμένη σε όλες της παρελθοντικές θέσεις της.

Το 'clear' ορίζει πως κάθε μετατροπή δεν θα αφήνει πληροφορία στις παρελθοντικές θέσεις και θα αποτυπώνεται μόνο η τρέχουσα θέση.

Το 'wrap' ορίζει την αναδίπλωση οριζόντια για κάθε frame, αποτυπώνοντας το φασματογράφημα ξανά μετά το τελευταίο frame.

Το 'clip' ορίζει την αποτύπωση αξονικά μόνο του τελευταίου frame/bin.

Το 'mirror' ορίζει την αποτύπωση αξονικά για κάθε frame/bin μετά το τέλος του φασματογραφήματος, αλλά με ανάποδη σειρά σαν να είναι καθρεπτισμένο.

Η δεύτερη ομάδα εντολών offset, zoom, anchor, theta, bin offset, ελέγχει την παραμόρφωση και αναδόμηση του φασματογραφήματος. Εδώ κάθε εντολή είναι το εργαλείο που αντιπροσωπεύει ο κέρσορας πάνω στο πλαίσιο του φασματογραφήματος και στο φασματογράφημα.

Το 'offset' ενεργοποιεί την μετατόπιση ολόκληρου του φασματογραφήματος στην κατάσταση που είναι.

Το 'zoom' ενεργοποιεί την παραμόρφωση του φασματογραφήματος δηλαδή την επιμήκυνση ή σμίκρυνση κάθε άξονα.

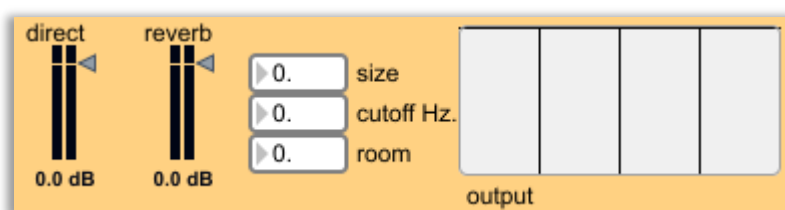
Το 'anchor' ενεργοποιεί την μετατόπιση του φασματογραφήματος εντός των ορίων της παραμόρφωσης.

Το 'theta' ενεργοποιεί την παραμόρφωση του φασματογραφήματος κατά γωνία ως προς τον άξονα των frames.

Το 'bin offset' ενεργοποιεί την εφαρμογή της συνάρτησης για την μεταβολή των bins.

Η εντολή 'reset zoom' επαναφέρει το φασματογράφημα στην αρχική του κατάσταση εφόσον αυτό είχε υποστεί παραμόρφωση.

Ο συνδυασμός της κατάστασης του περιβάλλοντος του φασματογραφήματος και του επιλεγμένου εργαλείου επεξεργασίας, μπορεί να αποτελεί μια τεχνική. Η εντολή 'draw mode' είναι ένα παράδειγμα που συνδυάζει την κατάσταση συνεχόμενης αποτύπωσης 'draw' όταν το φασματογράφημα έχει μικρύνει στο ένα εκατοστό του μεγέθους του και ως εργαλείο ορίζεται το 'offset' για ελεύθερη αποτύπωση σε όποιες θέσεις του πλαισίου του φασματογραφήματος επιθυμεί ο χρήστης. Τα δύο sliders (εικ.24) εφαρμόζουν μια μικρή μεταβολή του μεγέθους του φασματογραφήματος, μόνο όταν είναι ενεργοποιημένος ο συνδυασμός 'draw mode'.



Εικόνα25: Η έξοδος του ήχου της εφαρμογής.

Τέλος για την έξοδο ήχου του συστήματος υπάρχουν τα faders του απευθείας σήματος, που διαβάζεται από το φασματογράφημα και του ίδιου σήματος, αφού περάσει από μια μονάδα reverb. Οι παράμετροι του reverb ρυθμίζονται από τις τρεις μεταβλητές δίπλα στο fader του. Επίσης υπάρχει στην άκρη το πλαίσιο ενός άλλου φασματογράφου που αποτυπώνει το σύνολο του σήματος εξόδου.

### 3.2.2 Ανάλυση των αλγορίθμων της εφαρμογής Χρονική/Συχνотική Αναδόμηση Φάσματος.

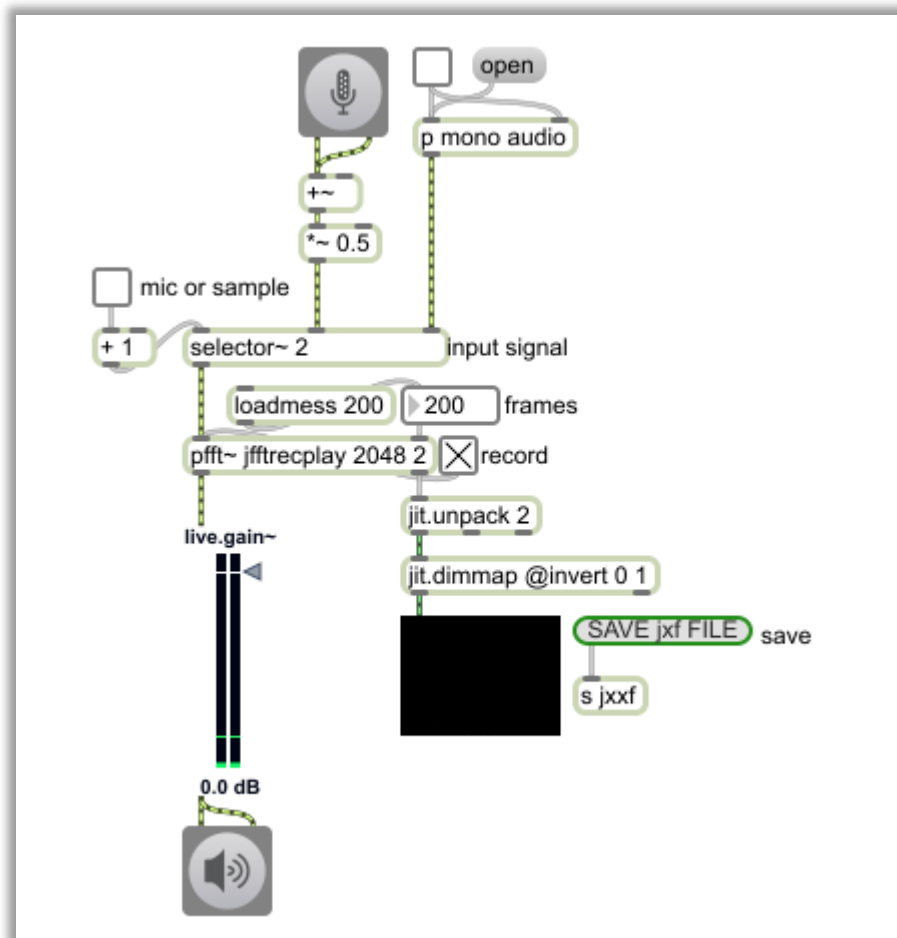
Όπως αναφέρθηκε και πριν ο αλγόριθμος επεξεργασίας είναι ξεχωριστός από τον αλγόριθμο καταγραφής. Αυτό γίνεται για δύο λόγους.

Πρώτον διότι πρακτικά η εφαρμογή είναι πραγματικού χρόνου εκτέλεσης ως προς το σήμα που παράγει κατά την έξοδο (τύπου synthesizer) και όχι του σήματος εισόδου(τύπου εφέ).

Ο δεύτερος λόγος είναι η εξοικονόμηση επεξεργαστικής ισχύος, καθώς ενώ ένας μέσος σύγχρονος επεξεργαστής CPU μπορεί να υποστηρίξει την εκτέλεση και των δύο αλγορίθμων ταυτόχρονα, όμως χάνει σε απόδοση σε ότι αφορά το αποτέλεσμα και αυτό είναι άμεσα αντιληπτό με καθυστερημένη ανταπόκριση στις εντολές και μείωση της ταχύτητας κατά την ανάγνωση των παραθύρων ανασύνθεσης και γενικότερα χαμηλότερη απόδοση.

Έτσι το σύστημα λειτουργεί πάντα σε ένα αλγόριθμο έχοντας μεγαλύτερη απόδοση. Πρώτα θα εκτελείται ο αλγόριθμος καταγραφής για να δημιουργεί το αρχείο και μετά θα εκτελείται ο αλγόριθμος επεξεργασίας για να τροποποιεί το αρχείο.

### 3.2.2.1 Ο αλγόριθμος καταγραφής.



Εικόνα26: Ο κεντρικός αλγόριθμος καταγραφής.

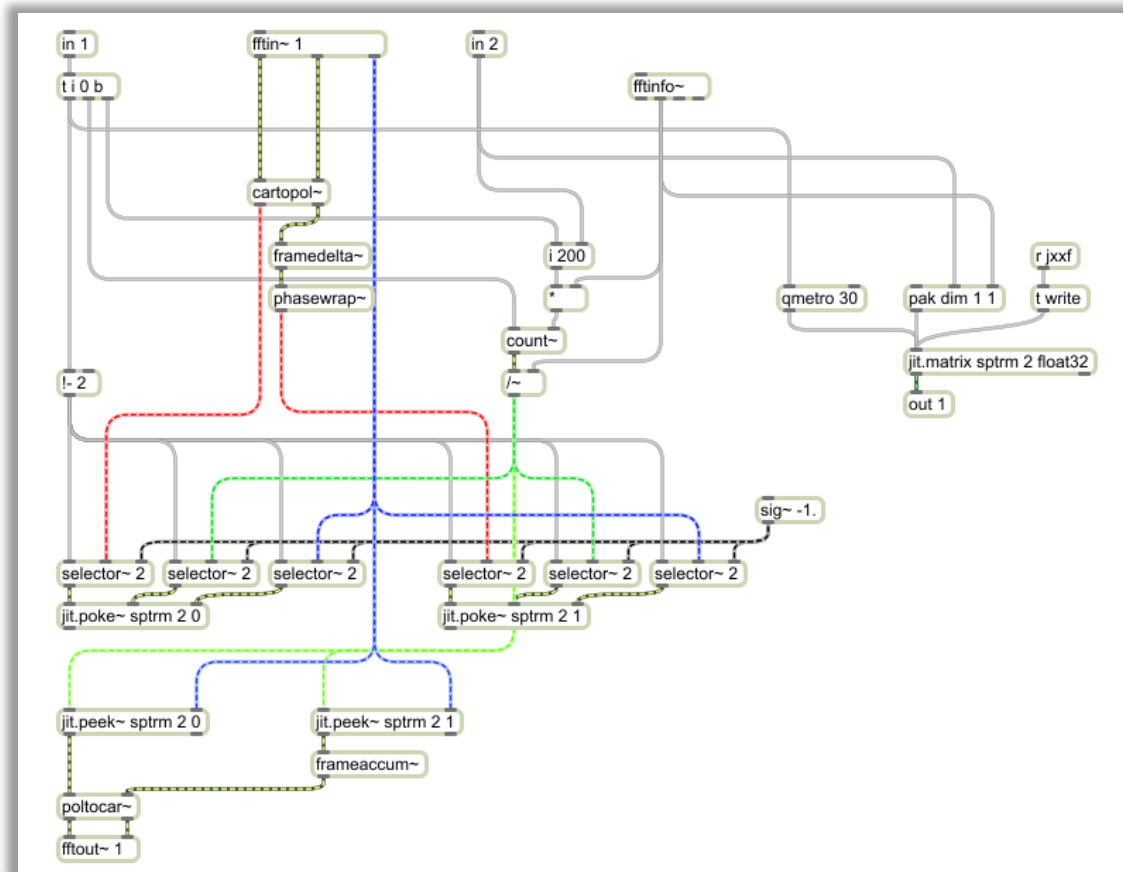
Ο κεντρικός αλγόριθμος καταγραφής φασματογραφήματος, δέχεται ως σήμα εισόδου ένα κανάλι είτε από μια εξωτερική μονάδα μικροφώνου ή ηχητικού συστήματος, είτε από ένα δείγμα ήχου που φορτώνεται.

Ο λόγος που χρησιμοποιείται ένα κανάλι είναι γιατί ο αλγόριθμος εξάγει ένα μόνο φασματογράφημα (μονοφωνικό) και αργότερα η επεξεργασία χειρίζεται ένα φασματογράφημα επίσης.

Τα σήματα εισόδου είναι σχεδιασμένα να μετατρέπονται από στερεοφωνικά σε μονοφωνικά. Ο αλγόριθμος δεν μπορεί να υποστηρίξει πάνω από δύο κανάλια.

Αφού επιλεγεί το επιθυμητό σήμα εισόδου, ορίζεται σε μία μεταβλητή το πλήθος των frames που θα καταγραφούν. Ως προεπιλεγμένη τιμή είναι τα 200 frames, καθώς τόσα υποστηρίζει ο αλγόριθμος επεξεργασίας. Το μέγεθος της ανάλυσης δεν μπορεί να αλλάξει και είναι σταθερά 2048.

Για να ξεκινήσει η εγγραφή είναι απαραίτητο να είναι ενεργοποιημένο το dsp και να ανοίξει ο διακόπτης record. Όσο η κατάσταση της εγγραφής είναι ενεργή, θα γράφονται frame by frame οι πληροφορίες που παράγονται από την ανάλυση του σήματος εισόδου. Όταν η εγγραφή φτάσει στο τελευταίο frame, θα συνεχιστεί ξανά από το πρώτο.



Εικόνα27: Το subpatch jfftreplay.maxpat.

Εντός του pfft~ πραγματοποιείται η διαδικασία της εγγραφής. Τα αντικείμενα jit.poke~ εισάγουν δεδομένα στη μνήμη jitter jit.matrix. Υπάρχουν δύο jit.poke~, ένα για συχνότητες και ένα για φάση. Κάθε αντικείμενο δέχεται τα ορίσματα της πληροφορίας της ανάλυσης(spectral data) του τρέχοντος frame εγγραφής(jitter frame index) και το βήμα του παραθύρου ανάλυσης(FFT frame index).

Τα φασματικά δεδομένα καθώς και το βήμα του παραθύρου ανάλυσης τα στέλνει το pfft~ αυτόματα. Για τον υπολογισμό των frames που θα εγγραφούν καλείται το αντικείμενο count~, που υπολογίζει σε επίπεδο dsp rate χρονικά διαστήματα τα οποία με τις κατάλληλες πράξεις κυμαίνονται σε εύρος τιμών από 1 έως 200.

Η μνήμη jitter jit.matrix είναι συνδεδεμένη με το αντικείμενο 'qmetro' με σκοπό να ανανεώνεται συνεχώς η αποτύπωση του περιεχομένου της μνήμης. Η ανανέωση γίνεται ανά ένα frame και προβάλλεται στον κεντρικό αλγόριθμο καταγραφής.

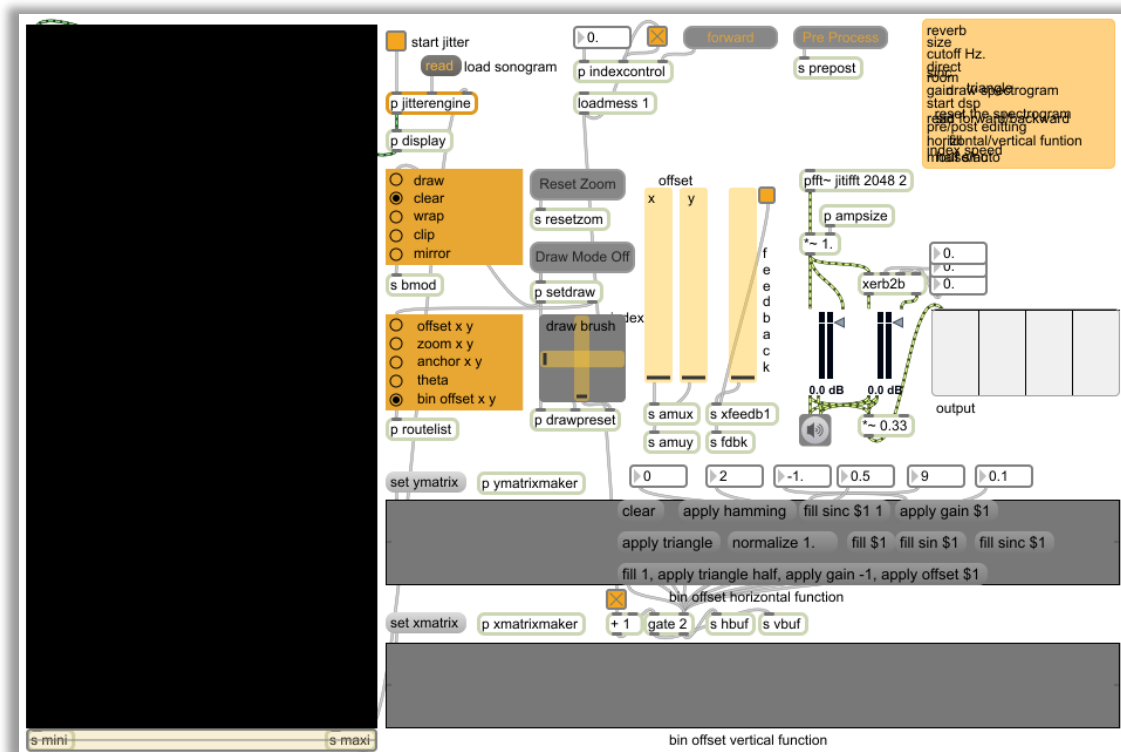
Τα αντικείμενα jit.peek~ διαβάζουν το περιεχόμενο της μνήμης jit.matrix ανά ένα frame, με την ίδια ταχύτητα που πραγματοποιείται η εγγραφή και στην ίδια θέση με αυτή της εγγραφής. Για να υπάρξει όμως περιεχόμενο και πληροφορία θα πρέπει να είναι γεμάτο το frame που διαβάζεται με δεδομένα. Επειδή τα αντικείμενα διαβάζουν με καθυστέρηση ίση με το πλήθος των frames, κατά την πρώτη ανάγνωση/εγγραφή του plane δεν θα υπάρξει επιστροφή δεδομένων.

Τα δεδομένα του jit.peek~ περνούν στο παράθυρο της ανασύνθεσης και έπειτα στην έξοδο του pfft~ και στην έξοδο του ήχου.

Όταν γεμίσουν με επιτυχία τα planes του φασματογραφήματος, ο χρήστης/μουσικός μπορεί να επιλέξει να σταματήσει την εγγραφή. Το περιεχόμενο θα παραμείνει στην μνήμη προσωρινά. Με την εντολή 'SAVE jxf FILE' μπορεί να αποθηκευτεί στον δίσκο το περιεχόμενο της μνήμης για να μπορεί να φορτώνεται στον αλγόριθμο επεξεργασίας και αναπαραγωγής.

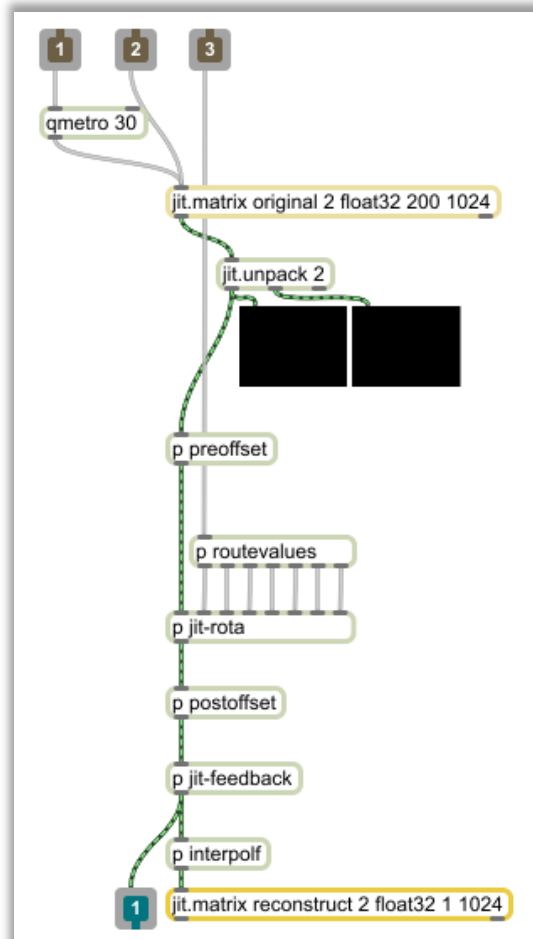
Το συνολικό σύστημα καταγραφής είναι φτιαγμένο σύμφωνα με την ίδια τεχνική που έχει αναπτυχθεί και στο παρελθόν από άλλους, όπως αναφέρεται και στη θεωρία και αποτελεί βοηθητικό εργαλείο και μέρος της συνολικής εφαρμογής.

### 3.2.2.2 Ο αλγόριθμος επεξεργασίας.



Εικόνα28: Ο κεντρικός αλγόριθμος επεξεργασίας.

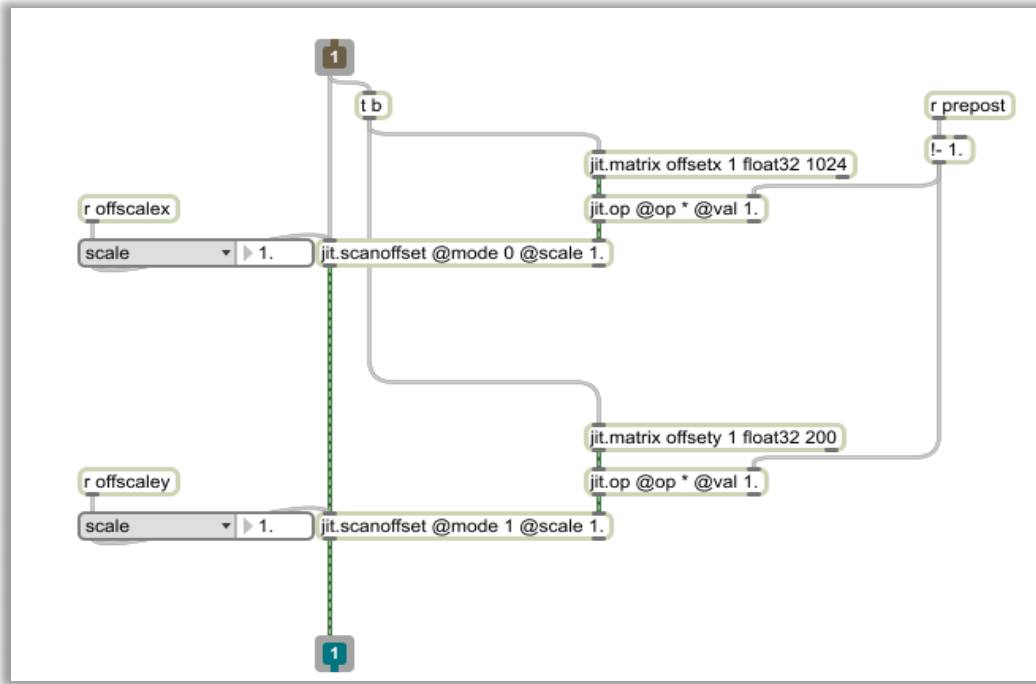
Ο κεντρικός αλγόριθμος επεξεργασίας δέχεται το αρχείο jxf χαρακτηριστικών 200x1024 dim, 32-bit float, 2 planes. Το subpatch που ελέγχει το αρχείο είναι το 'p jitterengine'. Αυτό αποθηκεύει το αρχείο στην προσωρινή μνήμη 'jit.matrix original' και αναλύει τα planes του.



Εικόνα29: ο υποαλγόριθμος jitterengine.

Το αντικείμενο 'jit.unpack 2' αποκαλύπτει το plane του πλάτους συχνοτήτων και το plane του πλάτους της φάσης. Επεξεργασία μπορεί να εφαρμοστεί και στα δύο plane. Η σημαντικότερη πληροφορία όμως βρίσκεται στο plane των συχνοτήτων και η χρησιμότητα της φάσης σύμφωνα και με τη θεωρία είναι στην ανασύνθεση. Εφαρμόζοντας επεξεργασίες στις συχνότητες μόνο, υπάρχει κέρδος από πλευράς επεξεργαστικής ισχύος και μικρή απώλεια ακρίβειας κατά την ανασύνθεση, χωρίς να υπάρχει ιδιαίτερη αλλοίωση του αποτελέσματος.

Το jitter σήμα προωθείται στο subpatch 'p preoffset'. Από κάτω υπάρχει ένα subpatch με το όνομα 'p postoffset'. Αυτά τα δύο subpatches εφαρμόζουν την ίδια επεξεργασία αλλά σε κάθε περίπτωση μόνο το ένα από τα δύο λειτουργεί. Ο σκοπός τους είναι να μπορεί να εφαρμόζεται η επεξεργασία σε διαφορετικά στάδια. Η εντολή pre/post process από την διεπαφή ορίζει αν θα είναι ενεργοποιημένο το πρώτο ή το δεύτερο subpatch.



Εικόνα30: Ο υποαλγόριθμος preoffset.

Η επεξεργασία πραγματοποιείται από το αντικείμενο `jit.scanoffset`. Υπάρχουν δύο τέτοια αντικείμενα στη σειρά, ένα για κάθε διάσταση. Στην πρώτη είσοδο κάθε αντικείμενου περνάει το σήμα προς επεξεργασία ενώ στη δεύτερη είσοδο φορτώνεται η συνάρτηση της μετατόπισης από την αντίστοιχη μνήμη `jit.matrix` για κάθε άξονα.

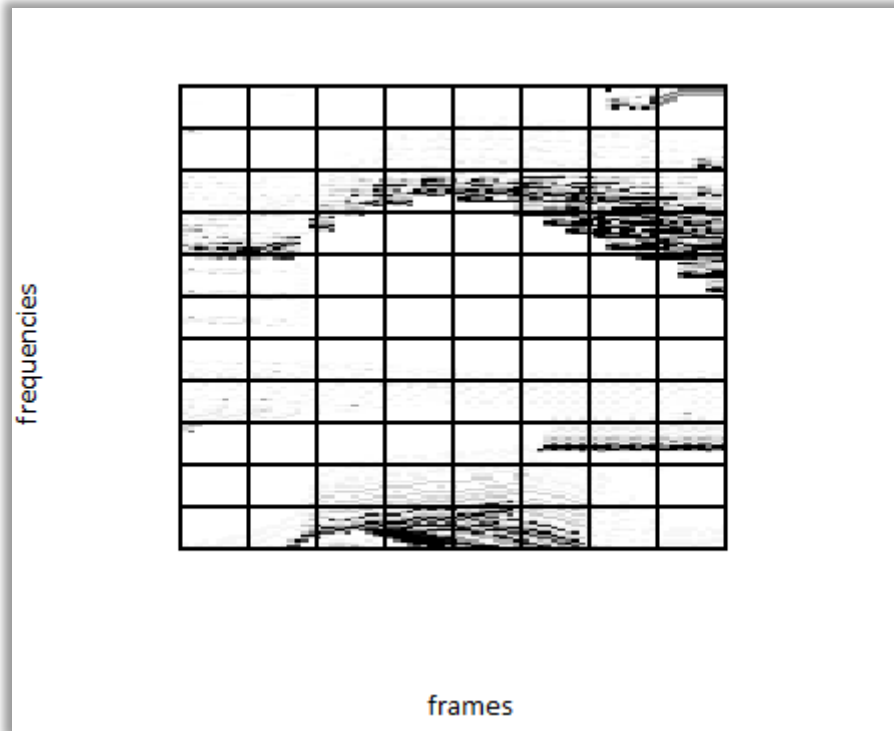
Κάθε μνήμη `jit.matrix` γεμίζει από τα πλαίσια για τον σχεδιασμό των συναρτήσεων που έχει η διεπαφή. Τα πλαίσια αυτά πραγματοποιούν την αποθήκευση σε μνήμες τύπου `buffer~` και έτσι πραγματοποιείται μία μετατροπή από `buffer` σε `jit.matrix`.



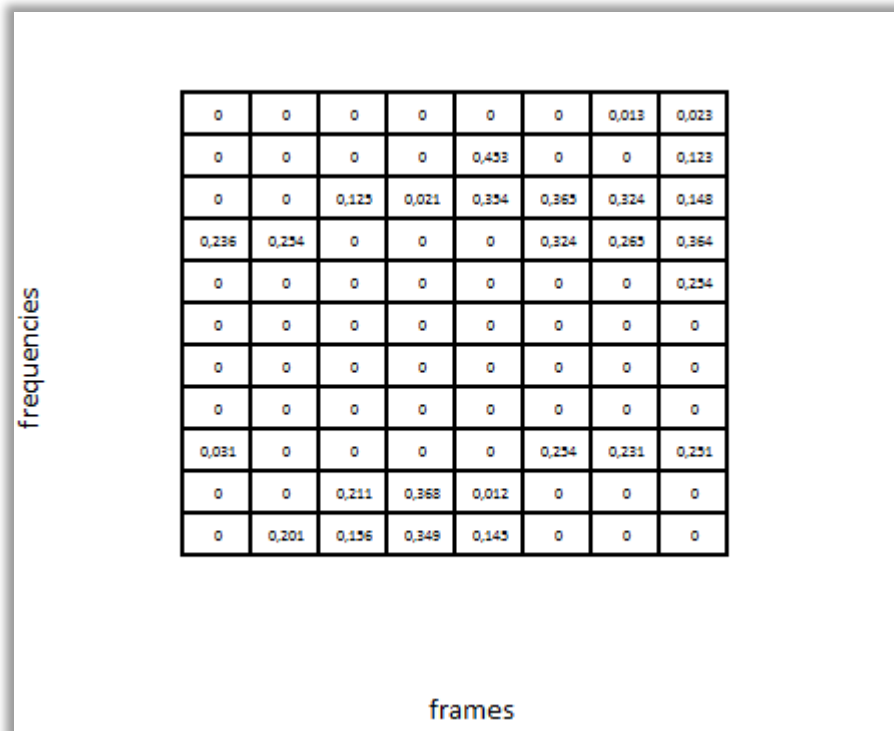
Εικόνα31: Η εγγραφή της συνάρτησης από το `buffer~` στο `jit.matrix` (πάνω για των άξονα των frames, κάτω για των άξονα των συχνοτήτων).



Το αποτέλεσμα από την εφαρμογή της επεξεργασίας είναι να μετατοπιστούν τα cells από κάθε plane σύμφωνα με τις συναρτήσεις. Το μέγεθος της μετατόπισης ορίζεται από την μεταβλητή scale σε κάθε άξονα και η τιμή που θα λάβει προέρχεται από τον κέρσορα που επιδρά με το φασματογράφημα.

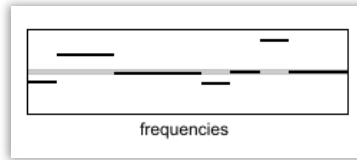


Εικόνα32: Προσομοίωση του plane των συχνοτήτων γραφικά.



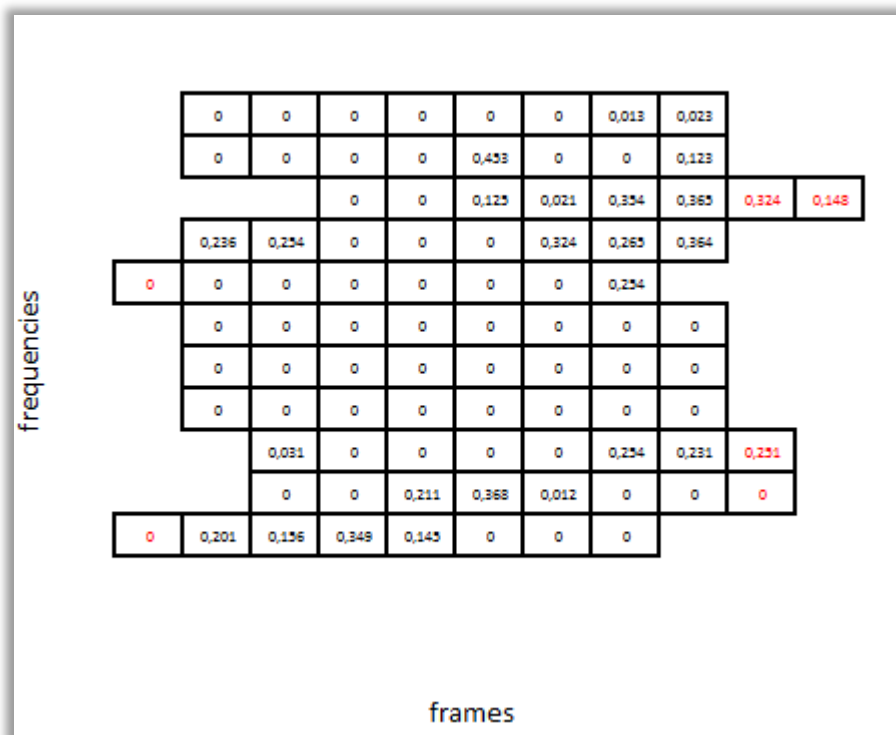
Εικόνα33: Προσομοίωση του plane των συχνοτήτων αριθμητικά.

Στις παραπάνω εικόνες αποτυπώνεται ο τρόπος που αντιλαμβάνεται ο αλγόριθμος επεξεργασίας τα planes. Όταν δοθεί πληροφορία διαφορετική από το μηδέν από την συνάρτηση μετατόπισης, ο αλγόριθμος του αντικειμένου jit.scanoffset επανατοποθετεί κάθε cell σύμφωνα και με το ποσοστό που έχει οριστεί.



Εικόνα34: Προσομοίωση συνάρτησης για τις συχνότητες.

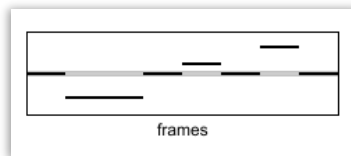
Παρότι η συνάρτηση των συχνοτήτων σχεδιάζεται οριζόντια στη διεπαφή (εικ.22) λόγω των περιορισμών των διαδραστικών αντικειμένων, ο αλγόριθμος την εφαρμόζει στον κάθετο άξονα.



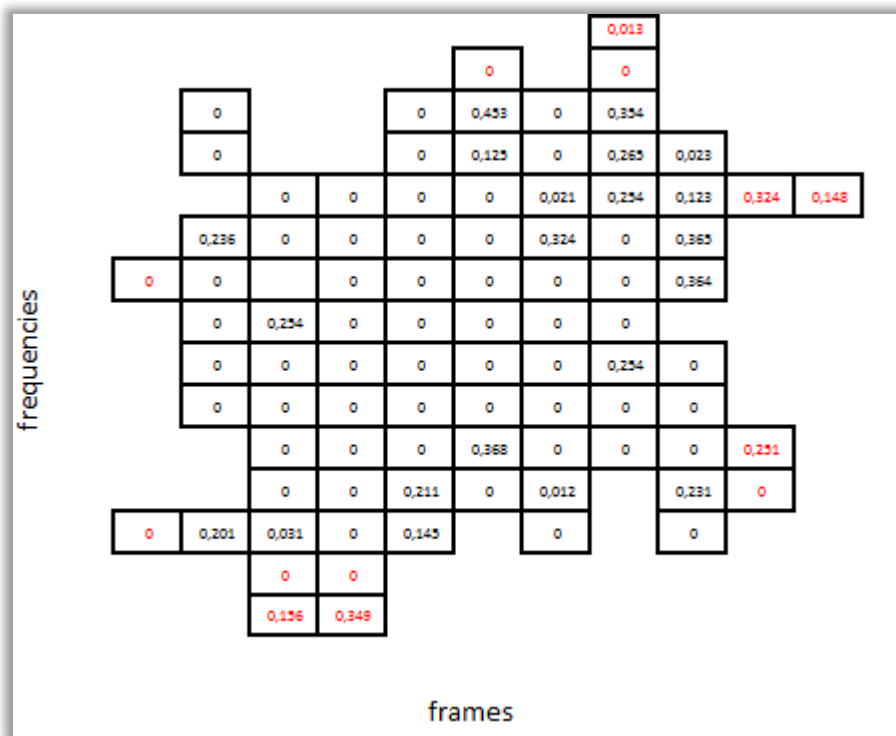
Εικόνα35: Προσομοίωση της εφαρμογής της συνάρτησης συχνοτήτων στα cells.

Το παράδειγμα της εικόνας αποτυπώνει πως θα είναι το plane μετά την εφαρμογή της συνάρτησης των συχνοτήτων. Με κόκκινο χρώμα είναι οι τιμές που έχουν υπερβεί τα όρια του plane στο φασματογράφημα.

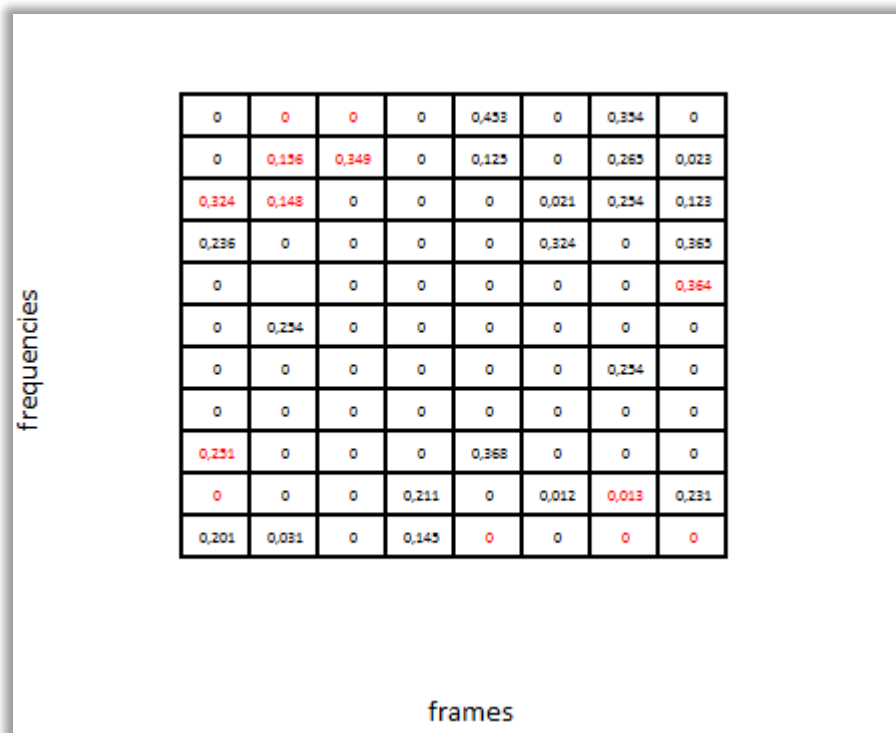
Ταυτόχρονα υπάρχει η δυνατότητα να εφαρμόζεται και μετατόπιση στον άξονα των frames. Στη περίπτωση αυτή, η συνάρτηση και ο άξονας είναι σε οριζόντια θέση.



Εικόνα36: Προσομοίωση συνάρτησης των frames.



Εικόνα37: Προσομοίωση του plane και με τις δυο συναρτήσεις.

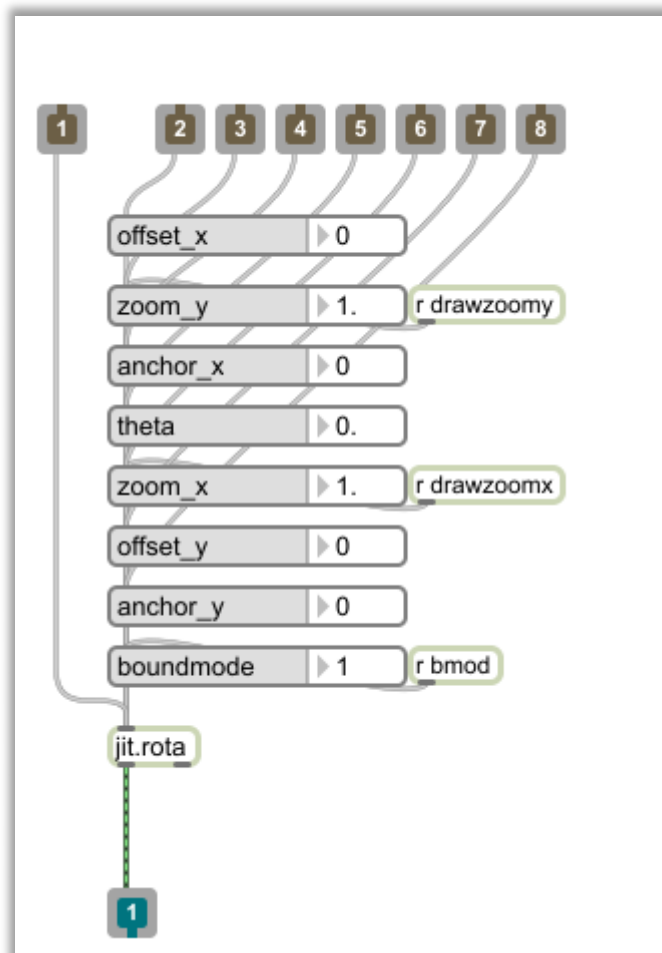


Εικόνα38: Προσομοίωση του plane όπως θα έχει ανακατασκευαστεί

Έχοντας και τις δύο συναρτήσεις, υπάρχουν cells που μετακινήθηκαν εντός των ορίων του plane και cells που εξείχαν. Ο αλγόριθμος έχει την ιδιότητα να τοποθετεί τα cells που έχουν μείνει απέξω από την αντίθετη μεριά πάνω στην διάσταση τους, επαναλαμβάνοντας το φασματογράφημα από την αρχή με νέο όμως περιεχόμενο. Ουσιαστικά ο αλγόριθμος βλέπει κάθε διάσταση ενωμένη κυκλικά με τον εαυτό της.

Όλη η διαδικασία μπορεί να πραγματοποιείται πριν η μετά το στάδιο της διαμόρφωσης του περιβάλλοντος του φασματογραφήματος και της εφαρμογής του εργαλείου επεξεργασίας. Με άλλα λόγια ένα ανακατασκευασμένο plane μπορεί να δέχεται περαιτέρω επεξεργασία, ή μπορεί να παραμείνει στην αρχική του μορφή όταν τροποποιείται το περιβάλλον του και να ανακατασκευάζεται μετά. Αυτός είναι και ο χαρακτήρας pre/post process της εφαρμογής.

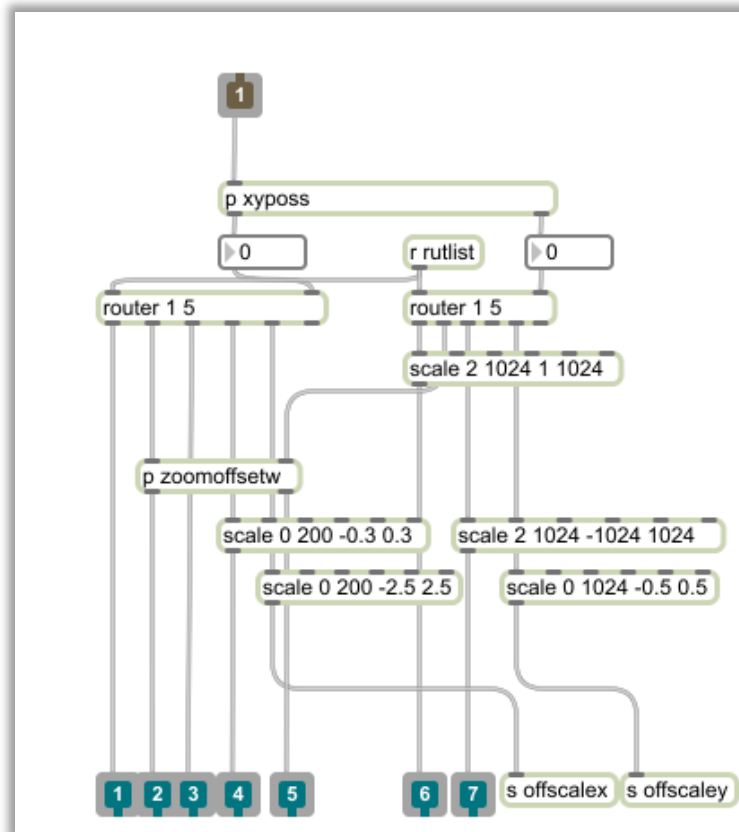
Στην εικόνα 29 του αλγόριθμου 'jitterengine', το σήμα jitter περνάει από το subpatch 'p jit-rota' το οποίο είναι συνδεδεμένο με το subpatch 'p routevalues'.



Εικόνα39: Το subpatch jit-rota.

Το αντικείμενο `jit.rota` εφαρμόζει τα έτοιμα εργαλεία επεξεργασίας και τις εντολές που αναφέρονται στην διεπαφή. Στην εικόνα φαίνονται με τα ίδια ονόματα οι μεταβλητές της διεπαφής με εξαίρεση το `'boundmode'` που ορίζει τις καταστάσεις του περιβάλλοντος του φασματογραφήματος (`draw`, `clear`, `wrap`, `clip`, `mirror`).

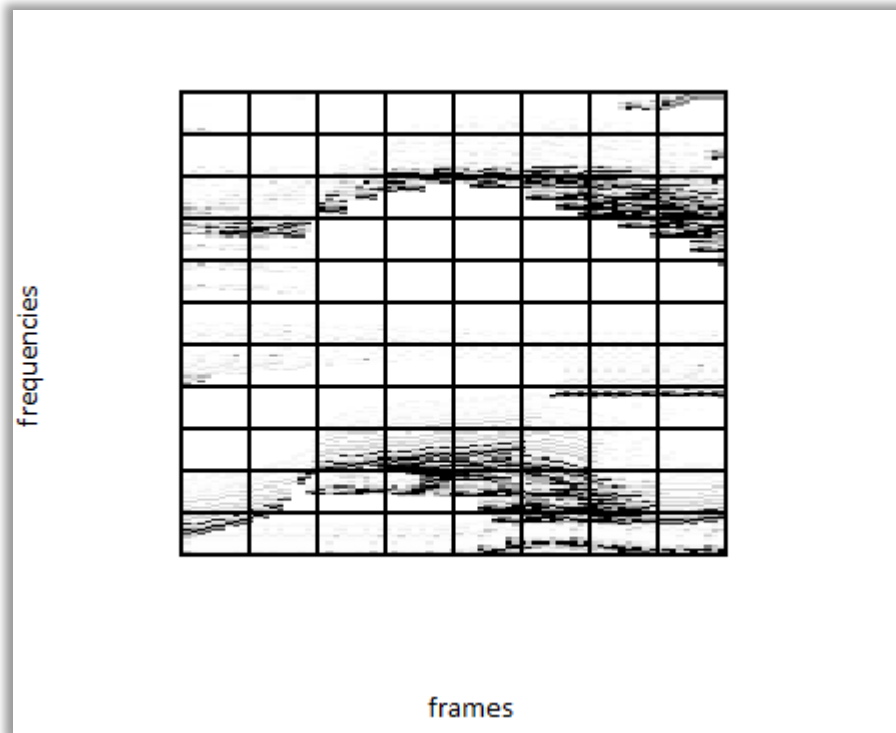
Όλες οι τιμές των μεταβλητών που εισέρχονται στο `jit.rota` βγαίνουν από το subpatch `'routevalues'`, που είναι ένα κέντρο δρομολόγησης και μετατροπής τιμών των περισσότερων μεταβλητών που χρησιμοποιούνται στην επεξεργασία.



Εικόνα40: Το subpatch `routevalues`.

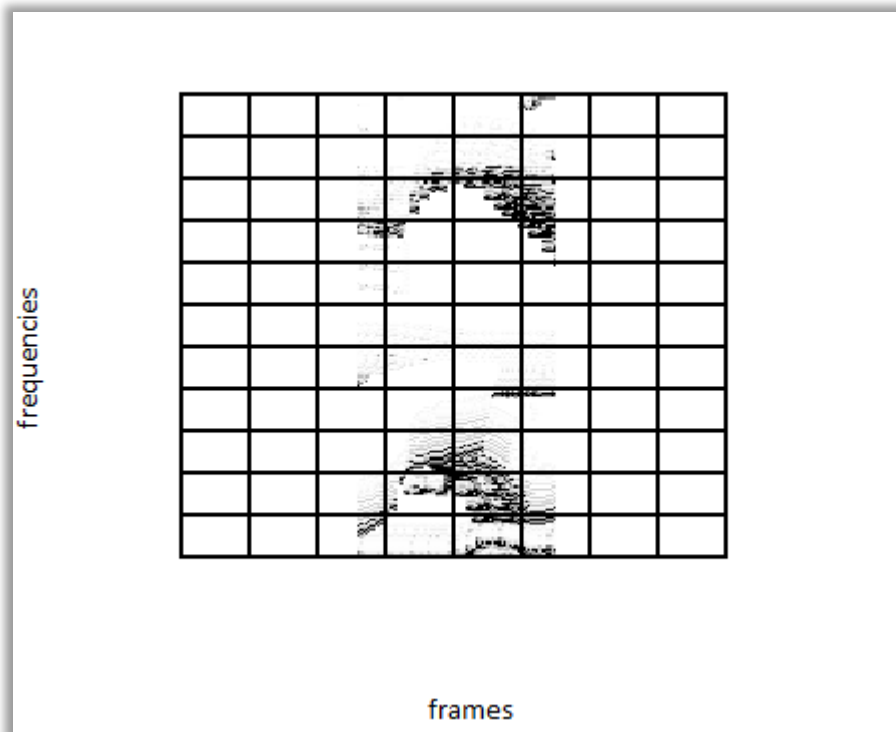
Σαν είσοδο των τιμών, το `'routevalues'` δέχεται την θέση του κέρσορα πάνω στο πλαίσιο του φασματογραφήματος όταν ενεργοποιείται. Με μία μικρή τροποποίηση ο κέρσορας επιστρέφει πάντα τη θέση ενός cell του plane, προωθώντας αυτή την τιμή με την αντίστοιχη κλίμακα. Με αυτόν τον τρόπο μπορεί να ελέγχεται οποιαδήποτε μεταβλητή της εφαρμογής με βάθος ακρίβειας ενός cell η αλλιώς ενός bin.

Ο τρόπος που αντιλαμβάνεται ο αλγόριθμος το σήμα επεξεργασίας παραμένει ο ίδιος και σε αυτό το στάδιο. Αυτό που αλλάζει είναι η μέθοδος τροποποίησης των cells.



Εικόνα41: Προσομοίωση του plane όπως εισάγεται στο jit.rota.

Το αντικείμενο jit.rota εφαρμόζει ομαδοποιημένα της μεταβολές στο οποίο plane δέχεται στην είσοδο του.

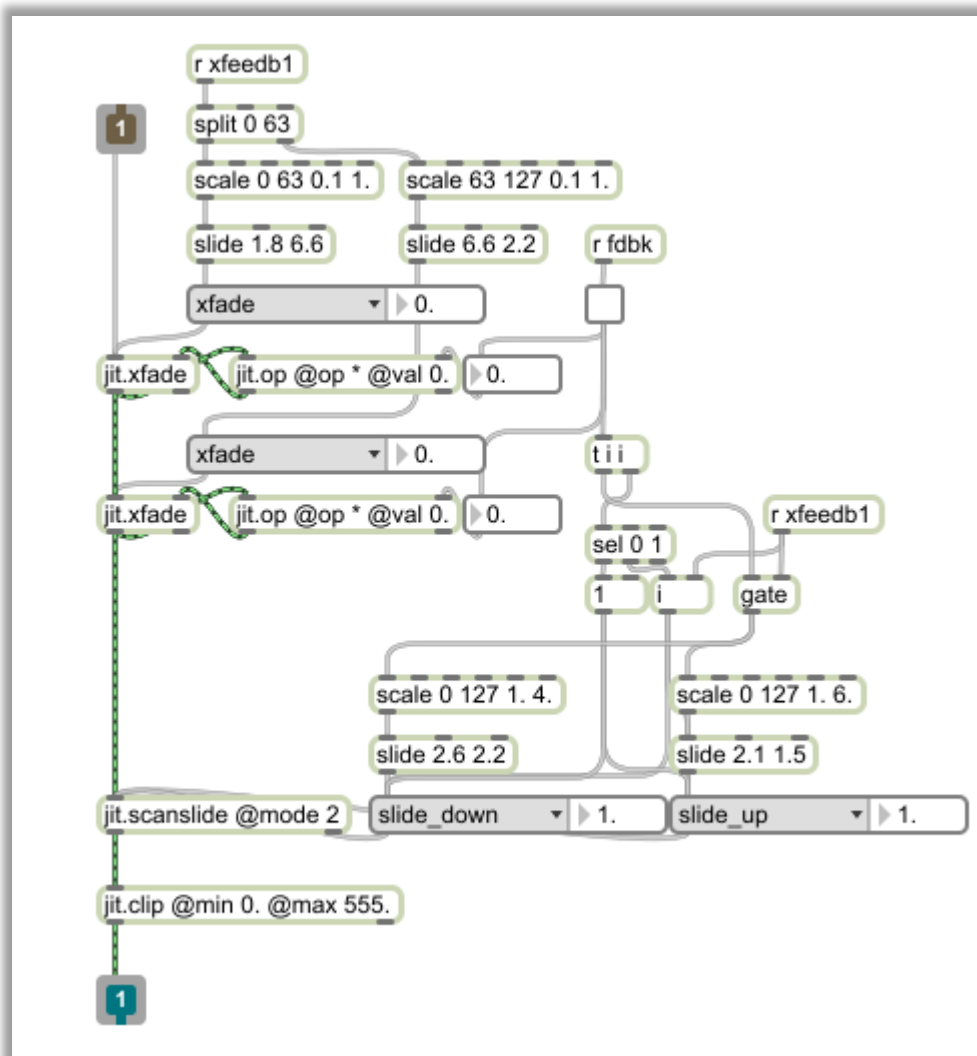


Εικόνα42: Προσομοίωση παραμόρφωσης του άξονα των frames.

Στην εικόνα 42 αποτυπώνεται πως θα σχηματιζόταν το plane αν δεχόταν συνολική παραμόρφωση στον άξονα των frames από την εντολή 'zoom' του 'jit.rota'. Αντίστοιχα με ομαδοποιημένο τρόπο εφαρμόζονται και οι υπόλοιπες επεξεργασίες.

Σε αυτό το στάδιο έχει ολοκληρωθεί η διεργασία του αλγορίθμου με τροποποιήσεις πάνω στα bins/cells του φασματογραφήματος/plane και η περιεχόμενη πληροφορία περνά σε ένα στάδιο προετοιμασίας για να βγει σε μορφή σήματος ήχου ξανά.

Πίσω στο 'jitterengine' (εικ.29) το σήμα jitter εισάγεται στο subpatch 'p jit-feedback'. Ο αλγόριθμος που εκτελείται μέσα σε αυτό, ελέγχεται από τις παραμέτρους της διεπαφής ενός διακόπτη ενεργοποίησης και ενός slider (εικ.21). Ο αλγόριθμος έχει επιπρόσθετο χαρακτήρα λειτουργίας και είναι στην επιλογή του χρήστη το αν θα εκτελείται και σε ποιο βαθμό.



Εικόνα43: Το subpatch jit-feedback.

Το αντικείμενο jit.xfade πραγματοποιεί μίξη μεταξύ δύο matrix. Η μεταβλητή ελέγχου είναι ένας δεκαδικός που ορίζει το ποσοστό που θα έχει το πρώτο σε σχέση με το δεύτερο matrix αφού προστεθούν στην έξοδο του αντικειμένου.

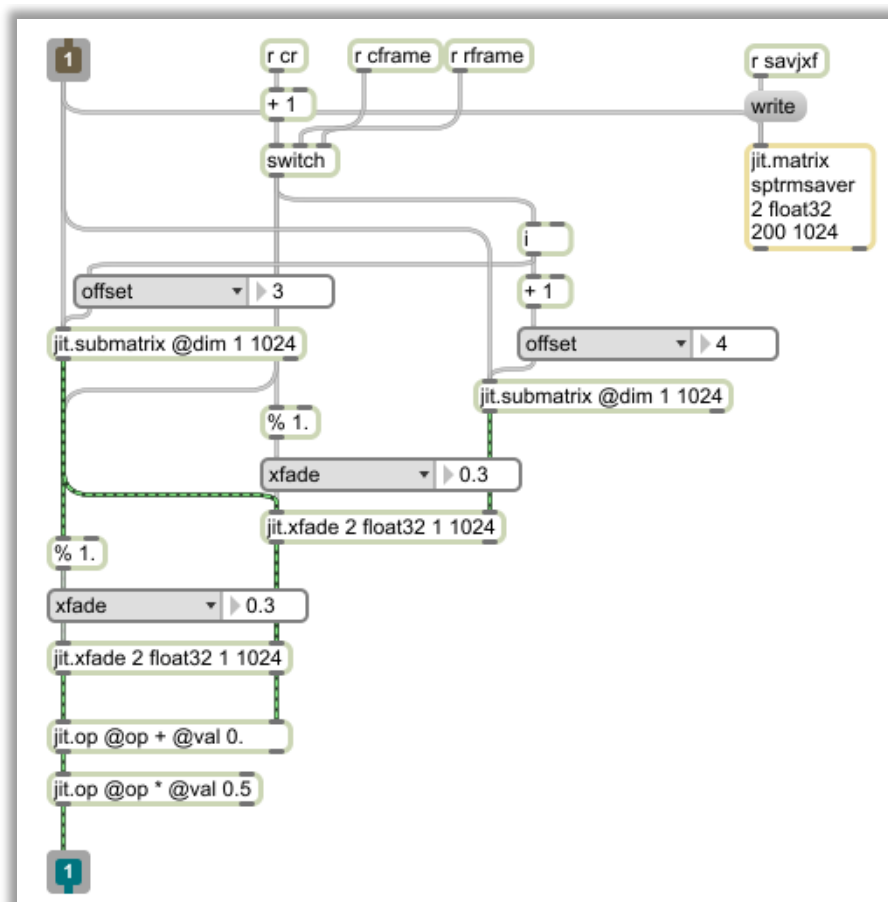
Στην προκειμένη περίπτωση ως δεύτερο σήμα εισόδου είναι η έξοδος του αντικειμένου. Σαν αποτέλεσμα αυτού δημιουργείται ανάδραση και το ποσοστό που ορίζει η μεταβλητή, πλέον αφορά το μέγεθος της απόσβεσης του συστήματος.

Όταν στο σύστημα τροφοδοτείται το ίδιο σήμα σταθερά, δε θα παρατηρηθεί κάποια ιδιαίτερη αλλαγή. Όταν υπάρξει μία μεταβολή λόγω των επεξεργασιών που αναφέρονται παραπάνω, τότε θα ξεκινήσει μια μετάβαση από την παρελθοντική κατάσταση πριν την επεξεργασία, στην τρέχουσα κατάσταση μετά την επεξεργασία. Ο χρόνος της μετάβασης ορίζεται από την μεταβλητή όπου μεγαλύτερη τιμή σημαίνει πιο αργή μετάβαση.

Δύο αντικείμενα `jit.xfade` είναι τοποθετημένα σειριακά για τον καλύτερο έλεγχο της ανάδρασης και μεγαλύτερη απόδοση στο τελικό αποτέλεσμα.

Πριν την έξοδο από το `subpatch` υπάρχει το αντικείμενο `jit.scanslide` που δημιουργεί ένα εφέ τύπου `blur` στον άξονα των συχνοτήτων για βελτίωση του τελικού αποτελέσματος μετά την εφαρμογή του `feedback`.

Πίσω στην εικόνα του `jitterengine`, το σήμα `jitter` οδηγείται σε μια έξοδο προς αποτύπωση της συνολικής του κατάστασης μετά από όλες τις επεξεργασίες και ταυτόχρονα στο `subpatch 'p interpolf'`.



Εικόνα44: Το `subpatch interpolf`.



Αυτό το στάδιο του αλγόριθμου μετατρέπει από το σύνολο του φασματογραφήματος, το τελικό frame που θα στέλνεται στην ανασύνθεση FFT και στην έξοδο του ήχου του συστήματος.

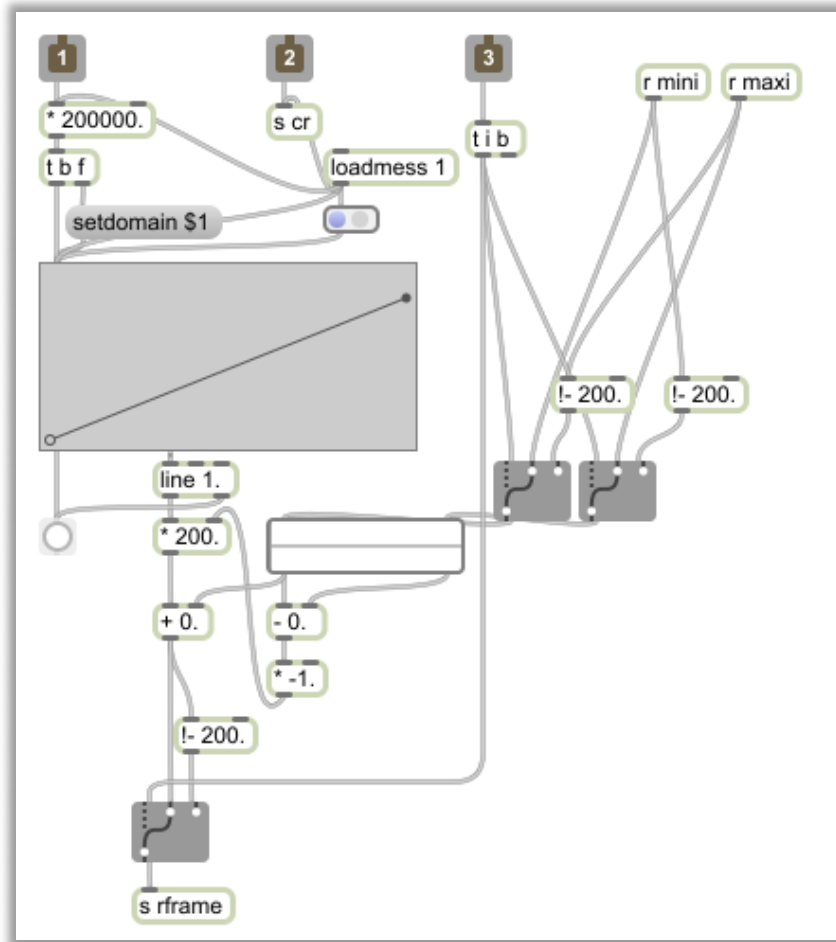
Όπως εισάγεται το jitter σήμα, αποθηκεύεται σε μία νέα προσωρινή μνήμη `jit.matrix` με σκοπό να μπορεί να εξάγεται το επεξεργασμένο φασματογράφημα σε αντίστοιχο αρχείο `jxf` όπως αυτό που τροφοδότησε τον αλγόριθμο αρχικά.

Όμως το αντικείμενο `jit.matrix` όπως και το αρχείο που δέχεται ο αλγόριθμος αποτελείται από δύο `plane`. Σε περίπτωση που για λόγους εξοικονόμησης επεξεργαστικής ισχύος επιλέχτηκε να συνδεθεί μόνο το `plane` των συχνοτήτων στον αλγόριθμο επεξεργασίας, το `jit.matrix` θα αποθηκεύσει μόνο το `plane` των συχνοτήτων και στις δύο θέσεις τους. Διαφορετικά αν επεξεργαστεί ολόκληρο το σήμα θα αποθηκεύσει κανονικά και το `plane` της φάσης.

Το σήμα `jitter` περνά στο αντικείμενο `jit.submatrix` όπου αλλάζει τελείως η κατάσταση του και μετατρέπεται από ένα `plane` τριών διαστάσεων `frequency/frame/bin` σε δυσδιάστατο `frequency/bin`. Το `jit.submatrix` ουσιαστικά διαβάζει πάντα ένα τμήμα του συνολικού `matrix` του σήματος εισόδου του και στην προκομμένη περίπτωση διαβάζει πάντα ένα `frame`.

Το `frame` που διαβάζει το `jit.submatrix` είναι η θέση ανάγνωσης του φασματογραφήματος που θα ακούγεται στην έξοδο του συστήματος όπως θα συνέβαινε με μία κυματομορφή. Ουσιαστικά είναι το `buffer` ανάγνωσης του σήματος.

Ο δείκτης ανάγνωσης ορίζει την θέση και στην προκειμένη περίπτωση το `frame` που θα διαβάζει η ανασύνθεση. Υπάρχουν δύο τρόποι ελέγχου του δείκτη ανάγνωσης, ο ένας είναι μέσω του φασματογραφήματος, επιλέγοντας με τον κέρσορα το `frame` με μεγαλύτερη ελευθερία κινήσεων και ο δεύτερος τρόπος είναι μέσω μιας συνάρτησης.



Εικόνα45: To subpatch indexcontrol, συνάρτηση ανάγνωσης των frames.

Πίσω στον κεντρικό αλγόριθμο (εικ.28) υπάρχει το subpatch 'p indexcontrol' που δημιουργεί τη συνάρτηση για τη μεταβίβαση από frame σε frame καθώς και τη μέθοδο που θα συμβαίνει αυτό. Οι παράμετροι που ρυθμίζονται είναι το εύρος των frames, η φορά του δείκτη ανάγνωσης και η ταχύτητα του.

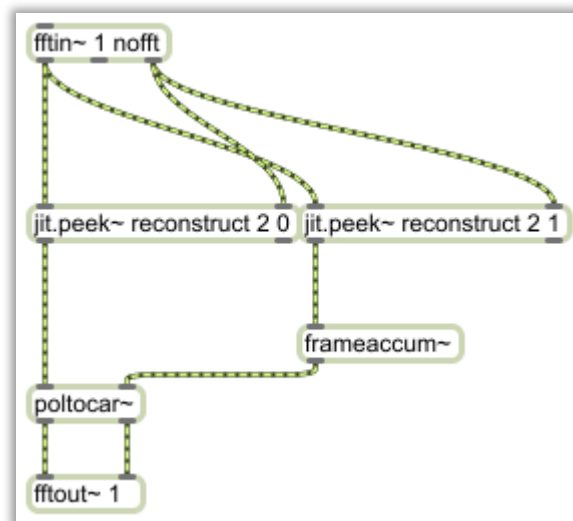
Πάντοτε εφαρμόζεται μία μέθοδος ανάγνωσης των frames. Γι' αυτό στην διεπαφή της εφαρμογής υπάρχει ο διακόπτης που επιλέγει τη μέθοδο.

Επιλέγοντας μια από τις παραπάνω μεθόδους, το jit.submatrix λαμβάνει ένα frame από το συνολικό plane. Ένα δεύτερο jit.submatrix (εικ.44) έχει σχεδιαστεί ώστε να διαβάζει το αμέσως επόμενο frame από αυτό που διαβάζει το πρώτο jit.submatrix. Με τη χρήση του αντικειμένου jit.xfade γίνεται μια ποσοστιαία μίξη μεταξύ των δύο jit.submatrix και κατά συνέπεια μεταξύ των διαδοχικών frames.

Η τεχνική αυτή όπως αναφέρεται και στη θεωρία λέγεται frame interpolation και έχει χρησιμοποιηθεί και στο παρελθόν.

Όπως φαίνεται στον υποαλγόριθμο της εικόνας 44, γίνεται μίξη μεταξύ του N και N+1 frame και στη συνέχεια γίνεται μίξη του αποτελέσματος με το N frame ξανά. Κάθε frame αποτελεί μία ακέραια θέση ή τιμή. Καθώς ο δείκτης ανάγνωσης μεταβαίνει από frame σε frame, κάνει προσπέλαση και των ενδιάμεσων δεκαδικών τιμών. Αυτές οι δεκαδικές τιμές ορίζουν τη μεταβλητή του jit.xfade για την μίξη των frames στο συγκεκριμένο κομμάτι.

Το αποτέλεσμα της εξόδου του subpatch 'r interpolf'(εικ.44) που φαίνεται στο τέλος του 'jitterengine' (εικ.29), αποθηκεύεται στο τελικό 'jit.matrix reconstruct' που αποτελείται από ένα μόνο frame, το τελικό frame της ανασύνθεσης.



Εικόνα46: Ο αλγόριθμος της ανασύνθεσης jitiff.maxpat.

Όπως φαίνεται και στον κεντρικό αλγόριθμο(εικ.28), στην είσοδο του pfft~ αντικείμενου δεν υπάρχει κάποιο σήμα. Έτσι μέσα στο pfft~ δεν δημιουργείται παράθυρο ανάλυσης καθώς τα δεδομένα της ανάλυσης είναι ήδη κατασκευασμένα.

Το αντικείμενο jit.peek~ διαβάζει τα περιεχόμενα plane του τελικού jit.matrix και επιστρέφει τις πληροφορίες πλάτους συχνοτήτων και φάσεων. Μετά την τροποποίηση των τιμών τους, τα δεδομένα στέλνονται απευθείας στο παράθυρο ανασύνθεσης.

Παρότι το pfft~ δεν εκτελεί κάποια ανάλυση, είναι σημαντικό να αρχικοποιηθούν οι παράμετροι του μεγέθους ανάλυσης και του overlap factor, καθώς τα χρησιμοποιεί κατά την ανασύνθεση.

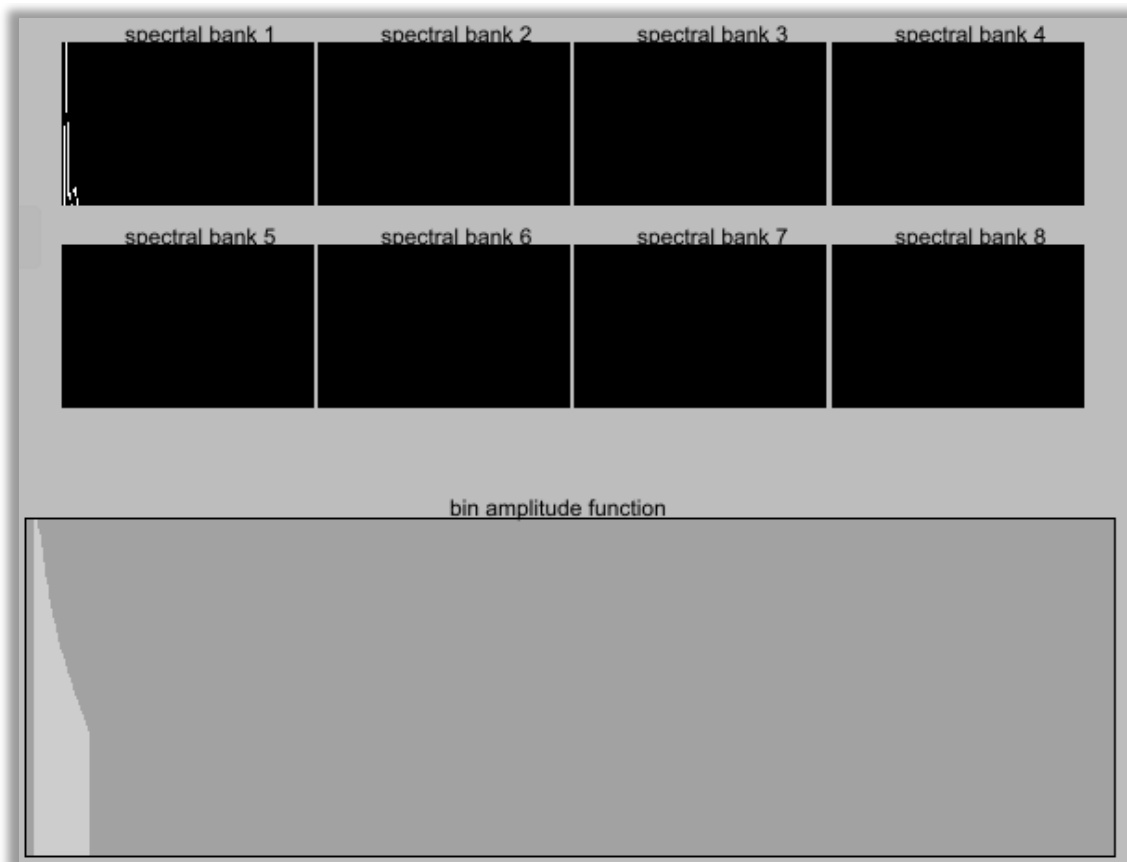
Ο ήχος που βγαίνει πλέον από το pfft~ οδηγείται στο fader του απευθείας σήματος και στη μονάδα reverb και στη συνέχεια από τη μονάδα reverb στο fader του σήματος με βάθος. Τα δύο fader συνδέονται στην έξοδο του dsp και σε ένα φασματογράφο για να παρακολουθείται το αποτέλεσμα ευκολότερα.

#### 4. Τρόπος λειτουργίας των εφαρμογών και ηχητικά παραδείγματα.

Σε αυτή την ενότητα θα δοθούν παραδείγματα και οδηγίες για την λειτουργία των εφαρμογών.

##### 4.1 Η εφαρμογή του Φασματικού Παγώματος.

Ενεργοποιώντας την εφαρμογή το σήμα εισόδου (dry) θα πρέπει να ακούγεται στην έξοδο ως ένδειξη ότι η εφαρμογή έχει αρχικοποιηθεί σωστά και μπορεί να ξεκινήσει η συλλογή των φασματικών πληροφοριών. Με την ρύθμιση του slider dry/wet επιλέγεται το ποσοστό που θα έχει το απευθείας σήμα εισόδου, από το σήμα που δημιουργεί ο αλγόριθμος.

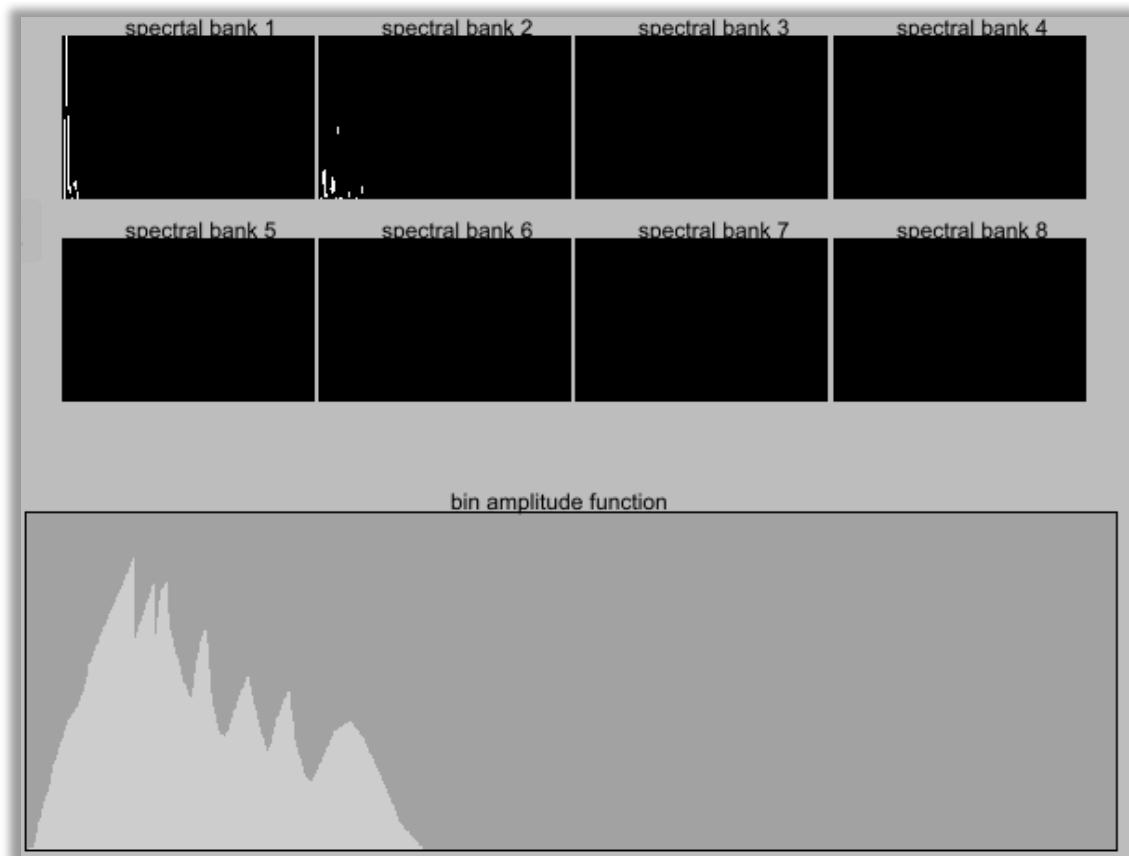


Εικόνα47: Συνάρτηση πλάτους των συχνοτικών bins και οι μνήμες αποθήκευσης.

Αμέσως μετά επιλέγεται η μνήμη που θα καταχωρηθεί το FFT frame τη στιγμή που θα δοθεί η εντολή "freeze". Η επιλογή της μνήμης γίνεται όταν ο κέρσορας πατήσει μία φορά πάνω στο αντίστοιχο πλαίσιο. Κάθε πλαίσιο μνήμης είναι αριθμημένο και ο αριθμός του πλαισίου που έχει επιλεγεί αναγράφεται στα αριστερά (εικ.14).

Το επόμενο στάδιο είναι να σχεδιαστεί η συνάρτηση που θα ορίζει πως θα κυμαίνεται το πλάτος κάθε συχνοτικού bin του FFT frame που πρόκειται να καταγραφεί.

Όταν οριστούν όλα τα παραπάνω, μπορεί να δοθεί ανά πάσα στιγμή η εντολή “freeze”, είτε με τον κέρσορα πατώντας πάνω στο κουμπί της εντολής (εικ.14), είτε με το πλήκτρο “spacebar” στο πληκτρολόγιο.



Εικόνα48: Η αποθήκευση σε διαφορετικές μνήμες.

Το αποτέλεσμα θα είναι να αποτυπωθεί το FFT frame διαμορφωμένο με την συνάρτηση μέσα στο πλαίσιο της μνήμης. Η εντολή που γεμίζει τη μνήμη μπορεί να επαναλαμβάνεται μέχρι να δημιουργηθεί ένα επιθυμητό αποτέλεσμα.

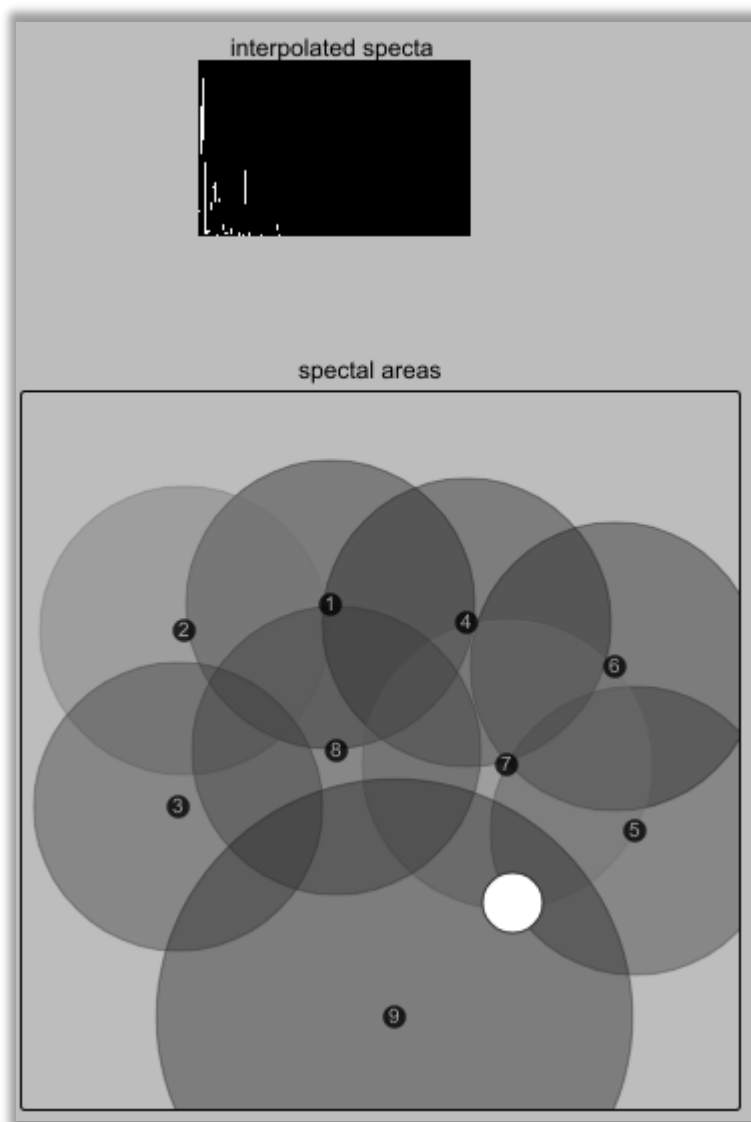
Όταν το FFT frame που έχει η μνήμη θέλουμε να το διατηρήσουμε, επιλέγουμε με τον κέρσορα μία άλλη κενή μνήμη και έτσι αποδεσμεύεται η μνήμη με το FFT frame που θέλουμε και τη θέση της παίρνει η καινούργια επιλεγμένη μνήμη.

Ολόκληρη η διαδικασία μπορεί να γίνει μέχρι να γεμίσουν και οι οκτώ μνήμες οι οποίες μπορούν να έχουν διαφορετικές συναρτήσεις και τελείως διαφορετικά σήματα που θα δημιουργούν FFT frames.



Εικόνα49: Όλες οι μνήμες περιέχουν φασματικά δεδομένα.

Το κεντρικό μέρος της εφαρμογής είναι ο έλεγχος της μίξης κάθε μνήμης σε ένα τελικό frame το οποίο θα μετατρέπεται σε σήμα ήχου και θα ακούγεται στην έξοδο.



Εικόνα50: Ο μηχανισμός μίξης των frames.

Το πλαίσιο “interpolated spectra” αποτυπώνει το περιεχόμενο του τελικού frame. Από κάτω το πλαίσιο με τα κυκλικά αντικείμενα είναι ο χώρος μίξης των frames που αποθηκεύτηκαν με την διαδικασία που αναφέρθηκε πιο πάνω.

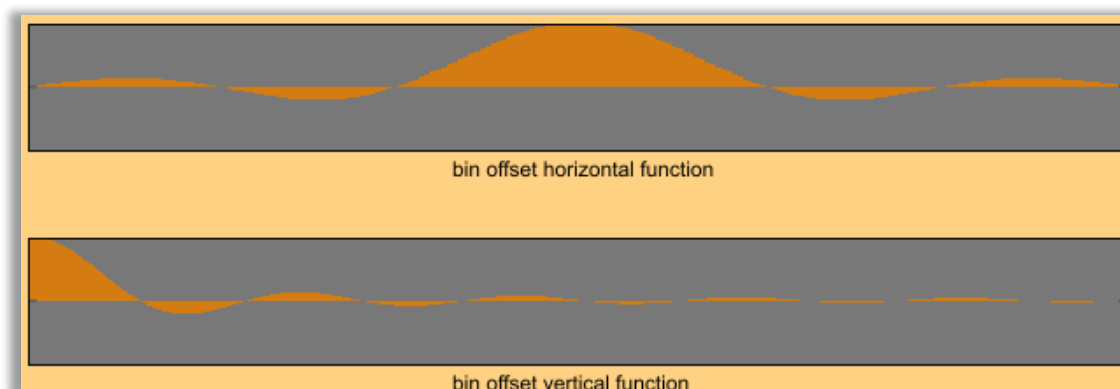
Κάθε κύκλος είναι αριθμημένος και αντιστοιχεί σε μία μνήμη. Ο κύκλος 9 είναι το άθροισμα όλων των frame με αναλογία 1:1.

Αντίθετα, η αναλογία του ποσοστού κάθε frame στην μίξη, ορίζεται από τις επικαλυπτόμενες περιοχές σκίασης μεταξύ των υπολοίπων κύκλων.

Όταν τοποθετηθούν οι περιοχές στις θέσεις που επιλέγει ο χρήστης, με τον κέρσορα πραγματοποιείται μια κίνηση επάνω σε αυτές. Το αποτέλεσμα της μεταβολής από την μίξη των φασμάτων αποτυπώνεται ζωντανά στο τελικό frame και περνά στην έξοδο του συστήματος ως το “wet” σήμα. (ηχητικά παραδείγματα: sample\_freeze.wav)

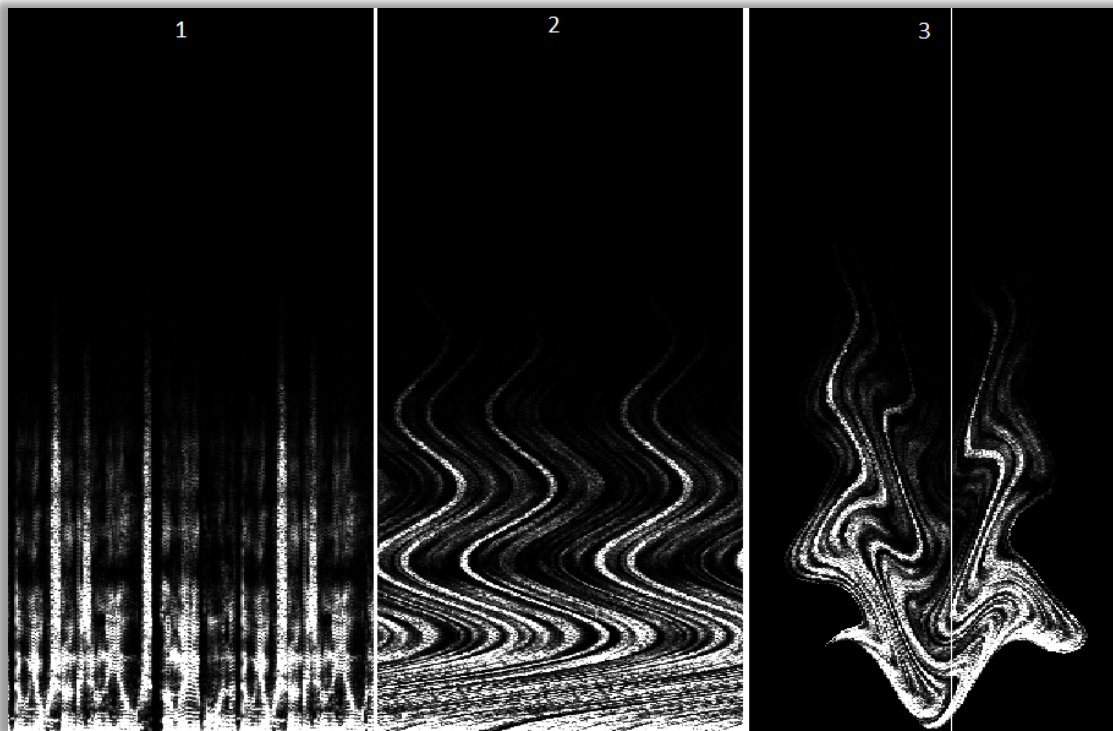
#### 4.2 Η εφαρμογή για τη Χρονική/Συχνотική Αναδόμηση Φάσματος.

Ενεργοποιώντας την εφαρμογή, θα ακούγεται από την έξοδο ήχου το αυθεντικό φάσμα που φορτώθηκε. Ξεκινώντας την διαδικασία επεξεργασίας του φασματογραφήματος, σχεδιάζονται οι συναρτήσεις μετατόπισης.



Εικόνα51: Συναρτήσεις μετατόπισης.

Ο σχεδιασμός μπορεί να γίνει είτε με τις εντολές που αναφέρονται στη διεπαφή, είτε με τον κέρσορα απευθείας πάνω στο πλαίσιο της συνάρτησης. Έπειτα ορίζεται το ποσοστό που θα εφαρμόζει η κάθε συνάρτηση στον κάθε άξονα, μέσω των αντιστοίχων sliders στη διεπαφή.

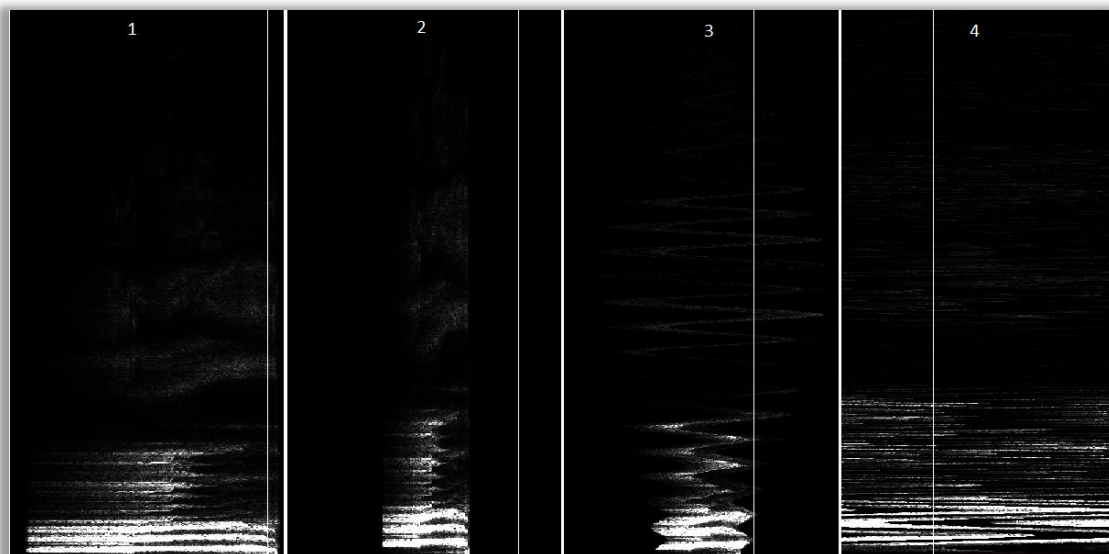


Εικόνα52: 1. Το αυθεντικό φασματογράφημα 2. Η εφαρμογή συνάρτησης στον κατακόρυφο άξονα 3. Η εφαρμογή συνάρτησης και στους δυο άξονες.

Ο κέρσορας όταν επιδρά με το πλαίσιο του φασματογραφήματος, πατώντας επάνω του, εφαρμόζει το εργαλείο που είναι επιλεγμένο στη διεπαφή. Στην πάνω εικόνα εμφανίζεται το αποτέλεσμα όταν ο κέρσορας αποτελεί το εργαλείο εφαρμογής της συνάρτησης μετατόπισης των bins/frames.

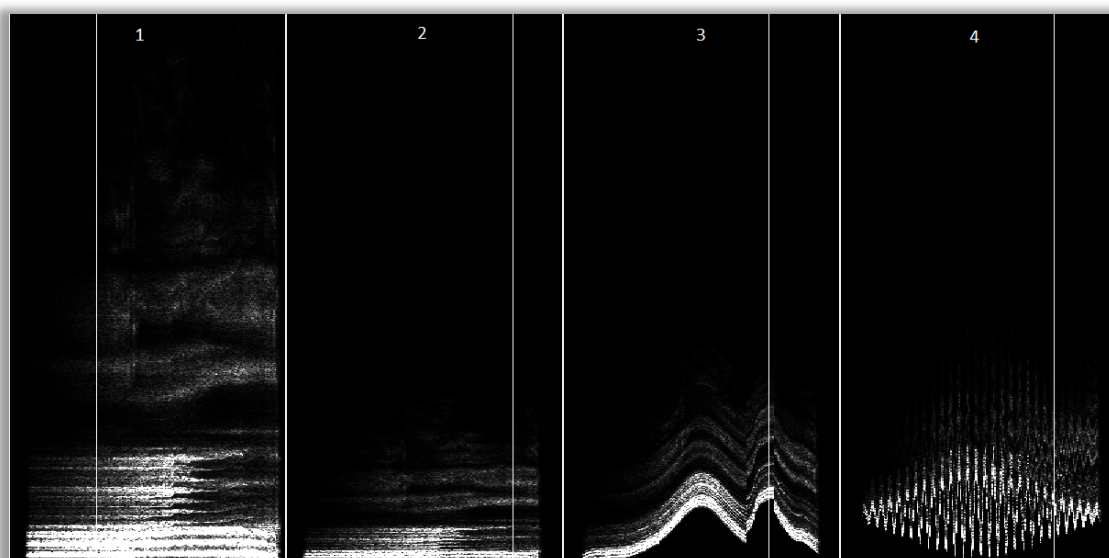
Ο κέρσορας μπορεί να αποτελεί ταυτόχρονα με το εργαλείο επεξεργασίας και τον δείκτη ανάγνωσης του φασματογραφήματος. Διαφορετικά το φασματογράφημα μπορεί να διαβάζεται από έναν αυτόματο δείκτη του οποίου ορίζεται το εύρος η ταχύτητα και η κατεύθυνση ανάγνωσης.





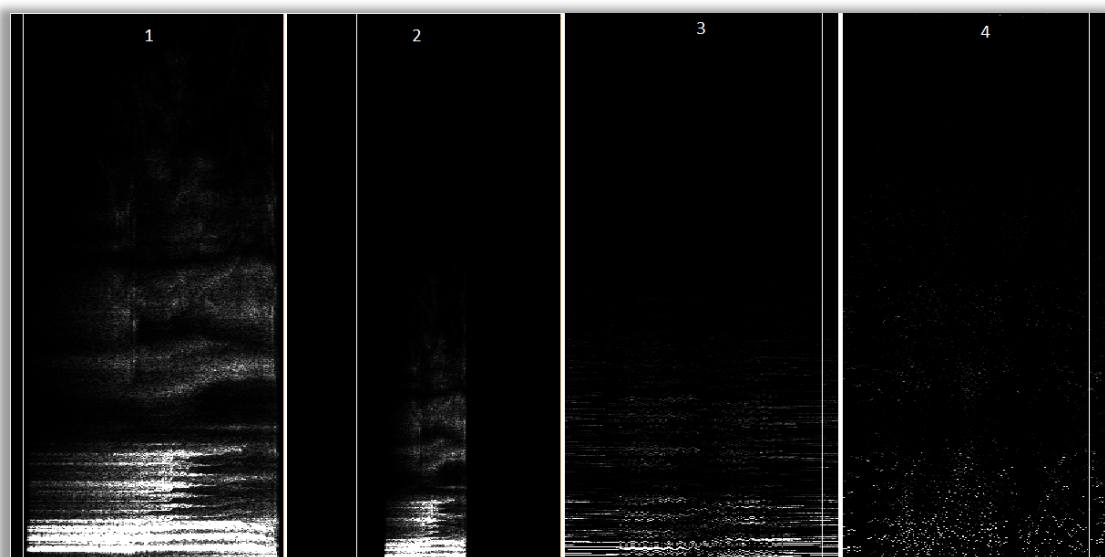
Εικόνα 53: Στάδια επεξεργασίας στον άξονα των συχνοτήτων.

Στην εικόνα 53 φαίνεται η σταδιακή προσαρμογή του φασματογραφήματος στις επεξεργασίες που δέχεται στον άξονα των συχνοτήτων. Στο στάδιο 1 είναι το αρχικό φασματογράφημα. Στο στάδιο 2 το φασματογράφημα έχει παραμορφωθεί με σκοπό να διακρίνονται οι μεταβολές του εντονότερα κατά το τελικό αποτέλεσμα. Στην στάδιο 3 εφαρμόζεται αρχικά σε ένα μικρό ποσοστό η μεταβολή του άξονα των συχνοτήτων και η μορφοποίηση στο φάσμα γίνεται αντιληπτή και οπτικά και ακουστικά. Στο στάδιο 4 το φασματογράφημα έχει αναδομηθεί σε μεγάλο βαθμό και το αποτέλεσμα έχει διαφοροποιηθεί τελείως από το αρχικό στάδιο. (ηχητικό παράδειγμα: sample\_offset1.wav)



Εικόνα 54: Στάδια επεξεργασίας στον άξονα των frames.

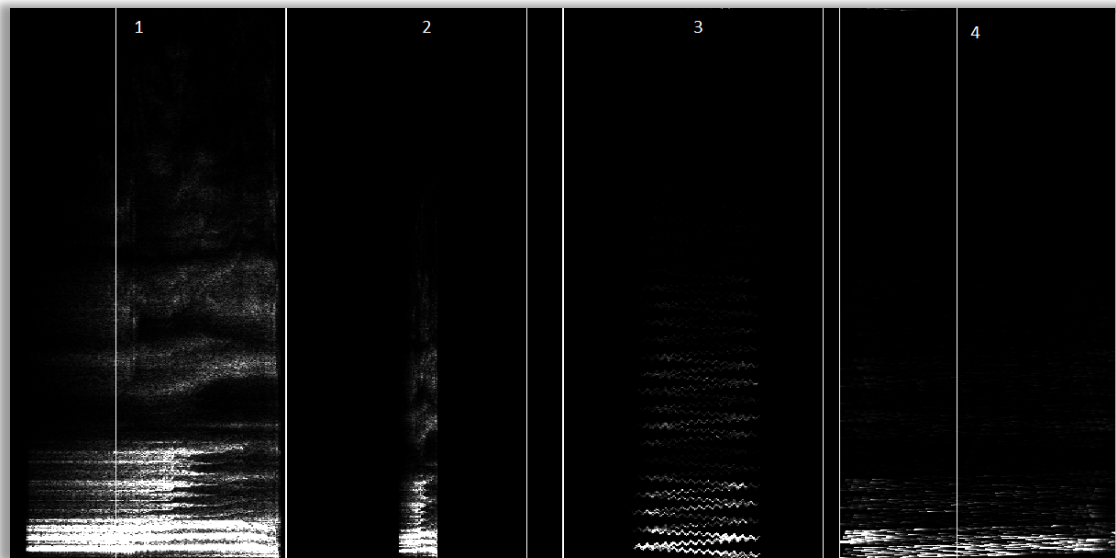
Στην εικόνα 54 φαίνονται τα στάδια που προσαρμόζεται το φασματογράφημα όταν εφαρμοστούν επεξεργασίες στον άξονα των frames. Το στάδιο 1 δείχνει το αρχικό περιεχόμενο. Στο στάδιο 2 το φασματογράφημα παραμορφώνεται στον άξονα των συχνοτήτων για να γίνεται εντονότερα αντιληπτή η μεταβολή που ακολουθεί. Στο στάδιο 3 εφαρμόζεται η συνάρτηση στον άξονα των frames, σαν αποτέλεσμα το φασματογράφημα αναδομείται και οι διακυμάνσεις γίνονται αντιληπτές τόσο οπτικά όσο και ακουστικά. Στο στάδιο 4 έχει εφαρμοστεί σε πολλαπλάσιο βαθμό η συνάρτηση που αναδομεί το φασματογράφημα με εντονότερα αποτελέσματα. (ηχητικό παράδειγμα: sample\_offset2.wav)



Εικόνα 55: Στάδια επεξεργασίας και στους δύο άξονες.

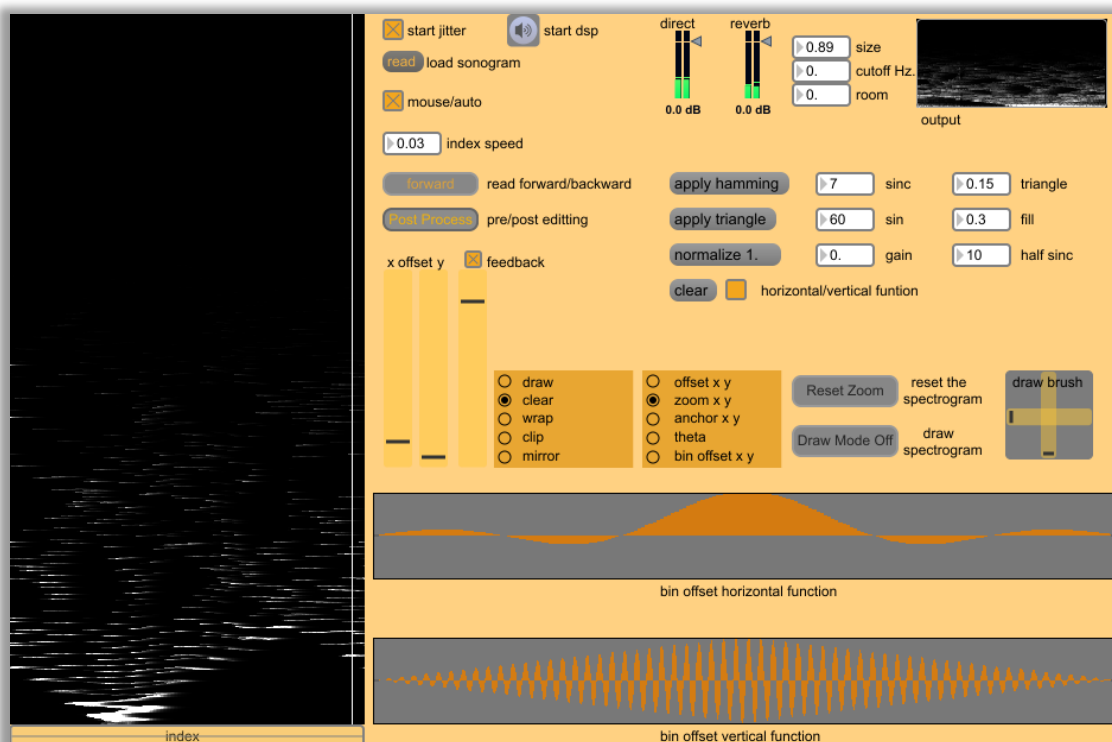
Στην εικόνα 55 το φασματογράφημα αρχικά παραμορφώνεται σε όλες τις διαστάσεις για να είναι ευκολότερη η επεξεργασία του. Με την εφαρμογή ταυτόχρονα και των δύο συναρτήσεων το φασματογράφημα αναδομείται σε μεγάλο βαθμό. Το αποτέλεσμα είναι πρωτότυπο και τα χαρακτηριστικά του αρχικού φασματογραφήματος έχουν μετασχηματιστεί σε κάτι τελείως νέο. (ηχητικό παράδειγμα: sample\_offset3.wav)

Για την εφαρμογή των επεξεργασιών και την δημιουργία φασματογραφημάτων με οποιονδήποτε βαθμό τροποποίησης, χρειάζεται να αρχικοποιηθούν και να συνδυαστούν αρκετές παράμετροι από την διεπαφή όπως φαίνεται στις παρακάτω εικόνες.



Εικόνα 56: Στάδια πολλαπλών επεξεργασιών.

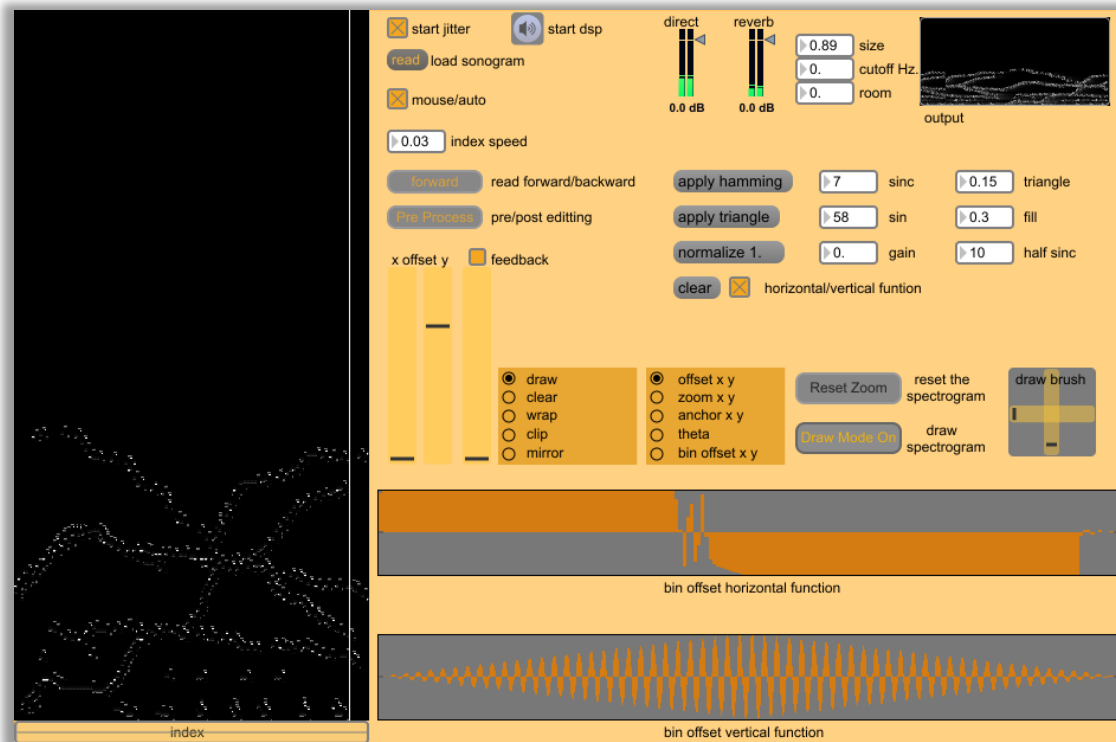
Στην εικόνα 56 προβάλλονται τα στάδια επεξεργασίας του φασματογραφήματος με πολλαπλές εφαρμογές εργαλείων. Οι μεταβολές αφορούν τους άξονες χρόνου και συχνότητας, το συνολικό μήκος του φασματογραφήματος καθώς και την ενεργοποίηση του συστήματος ανάδρασης πάνω στα bins του φασματογραφήματος. (ηχητικό παράδειγμα: sample\_offset4.wav)



Εικόνα57: Εφαρμογή πολλαπλών επεξεργασιών στο φασματογράφημα.

Στην εικόνα 56 φαίνεται το φασματογράφημα με μετατοπισμένα bins, συρρικνωμένα frames και ενεργοποιημένο το σύστημα ανάδρασης που σχηματίζει ίχνη από τις μεταβολές των bins.

Εδώ η επεξεργασία εφαρμόζεται σε ολόκληρο το περιβάλλον του φασματογραφήματος (post process) και όχι μόνο εντός των διαστάσεων του. Με τον τρόπο αυτό δημιουργούνται κενές περιοχές φασματικών δεδομένων λόγω της μετατόπισης.



Εικόνα58: Η εφαρμογή της αποτύπωσης.

Έχοντας επιλέξει την εντολή “Draw Mode” και ρυθμίζοντας το μέγεθος των διαστάσεων του φασματογραφήματος παραμορφώνοντας το, δίνεται η δυνατότητα αποτύπωσης του, σε οποιοδήποτε σημείο του πλαισίου του, αφήνοντας τα παρελθοντικά δεδομένα αποτυπωμένα και δημιουργώντας έτσι ένα τελείως διαφορετικό περιεχόμενο που μπορεί να ανασυντεθεί σε ήχο. (ηχητικό παράδειγμα: sample\_offset5.wav)

Με τις εντολές clear και draw μπορεί να καθαριστεί το πλαίσιο από όλες τις παρελθοντικές αποτυπωμένες θέσεις και να σχεδιαστεί ξανά το καθαρό πλαίσιο.

## 5. Συμπεράσματα.

Ο στόχος που αρχικά αναφέρθηκε για το περιεχόμενο και το αποτέλεσμα των εφαρμογών έχει επιτευχθεί σε σημαντικό βαθμό. Οι εφαρμογές λειτουργούν και επεξεργάζονται σύμφωνα με τις αρχές και τις μεθόδους που εξηγήθηκαν στην θεωρεία. Το ηχητικό αποτέλεσμα έχει σίγουρα ένα βαθμό ενδιαφέροντος και πρωτοτυπίας και η χρήση των εφαρμογών μπορεί να είναι για ζωντανή μουσική εκτέλεση πέρα της ηχογράφησης.

Η εφαρμογή του Φασματικού Παγώματος (spectral freeze) δημιουργεί πρωτότυπους ήχους μέσω συλλογής χροιών. Οι χροιές χρησιμοποιούνται για σύνθεση ενός εξολοκλήρου νέου ήχου με διαφορετικό ηχόχρωμα και δυναμική.

Η εφαρμογή της Χρονικής/Συχνοτικής Αναδόμησης Φάσματος (spectral offset) παρέχει ένα είδος υβριδικής τεχνικής στην επεξεργασία δεδομένων ήχου. Τα εργαλεία που χρησιμοποιούνται συνδυάζονται με ένα πρωτότυπο τρόπο. Οι δυνατότητες δημιουργίας ηχητικών συμβάντων είναι στα πλαίσια πρωτότυπης σύνθεσης ήχου και ηχητικού σχεδιασμού.

Κατά τα στάδια της ανάπτυξης και δοκιμής των εφαρμογών, παρατηρήθηκαν ορισμένες ιδιαιτερότητες από πλευράς συστήματος, περιβάλλοντος και εφαρμογής. Αυτές οι ιδιαιτερότητες οδήγησαν σε κάποιους περιορισμούς. Και κατά συνέπεια για την αντιμετώπιση αυτών των περιορισμών χρειάστηκε να τροποποιηθούν κάποιες πιο ανεπτυγμένες και απαιτητικές λειτουργίες για να μην επηρεαστεί η λειτουργικότητα της συνολικής εφαρμογής.

Συγκεκριμένα στην εφαρμογή της χρονικής/συχνοτικής αναδόμησης φάσματος, ο κυριότερος περιορισμός είναι στο μέγεθος της ανάλυσης FFT και στο μέγεθος της ανάλυσης του plane αποθήκευσης και επεξεργασίας. Το ένα μέρος του ζητήματος είναι ότι η εφαρμογή δεν μπορεί να προσαρμόσει πλήρως και αυτόματα τα μεγέθη της για διάφορες αναλύσεις. Το δεύτερο μέρος είναι ότι για μεγάλες αναλύσεις η ομαλότητα της ροής της εφαρμογής αλλοιώνεται και η επιδόσεις είναι δραματικά χαμηλότερες. Έτσι για να αποφευχθούν θραύσεις και καθυστερήσεις στον πραγματικό χρόνο εκτέλεσης της εφαρμογής, έχει οριστεί ως βέλτιστη ανάλυση plane τα 200x1024 bins. Η εφαρμογή δεν θα λειτουργήσει σωστά με διαφορετική είσοδο.

Στην ίδια εφαρμογή επίσης, μπορεί να παρατηρηθεί ελαφριά πτώση της απόδοσης και θραύσεις στο ηχητικό αποτέλεσμα όταν έχει ενεργοποιηθεί το σύστημα ανάδρασης και τροφοδοτείται με πληροφορίες από ολόκληρο το μέγεθος του plane. Σε έναν ταχύτερο επεξεργαστή θα παρατηρηθεί σε χαμηλότερο βαθμό. Η προτεινόμενη χρήση σε περίπτωση που είναι έντονο το φαινόμενο της θραύσης, είναι να τροφοδοτούνται μικρότερα μερίσματα φάσματος στο σύστημα ανάδρασης.

Είναι χαρακτηριστικό του περιβάλλοντος Max/MSP Jitter η εκτέλεση του γραφικού μέρους είναι προεπιλεγμένο να λαμβάνει χώρα πάνω στον επεξεργαστή της CPU. Αντίθετα με άλλα γραφικά περιβάλλοντα που δεσμεύουν τον επεξεργαστή της GPU και έχουν βελτιστοποιηθεί για πολλές αρχιτεκτονικές τέτοιων επεξεργαστών, το περιβάλλον Jitter προεπιλέγει να λειτουργεί μόνο με CPU.

Η χρήση της GPU προαπαιτεί ένα διαφορετικό και επιπλέον αλγοριθμικό υπόβαθρο που όμως η εργασία δεν επικεντρώθηκε γύρω από αυτό και δεν έγινε η ανάπτυξη του. Με την βελτιστοποίηση των εφαρμογών πάνω στην GPU σε ένα επόμενο στάδιο, οι επιδόσεις θα αυξανόντουσαν σημαντικά και κάποιοι από τους περιορισμούς θα ξεπερνιόντουσαν. Οι εφαρμογές είναι ανοιχτές για περαιτέρω ανάπτυξη, βελτιστοποίηση και προσαρμογή σε νέα δεδομένα.

Και στις δύο εφαρμογές η μέθοδος ελέγχου και χρήσης είναι το ποντίκι και το πληκτρολόγιο. Σε ένα επόμενο στάδιο οι εφαρμογές θα μπορούσαν να προσαρμοστούν σε περισσότερο σύγχρονα συστήματα διεπαφής, για παράδειγμα εναλλακτικοί controllers ή οθόνες αφής ή πιο ανεπτυγμένα διαδραστικά γραφικά περιβάλλοντα, με σκοπό την ευκολότερη, πιο απλοποιημένη και αποτελεσματικότερη χρήση τους.

## Παράρτημα

Το κείμενο συνοδεύεται από cd με τα περιεχόμενα αρχεία των εφαρμογών και των δειγμάτων ήχου.

Τα αρχεία είναι:

main\_spectralfreeze.maxpat  
fftbinfx4.maxpat  
sample.wav  
sample\_freeze.wav

main\_spectraloffset.maxpat  
jrecplay.maxpat  
jitiff.maxpat  
jfftreplay.maxpat  
xerb2b.maxpat  
sample\_offset1.wav  
sample\_offset2.wav  
sample\_offset3.wav  
sample\_offset4.wav  
sample\_offset5.wav  
sample\_matrix.jxf

Οι εφαρμογές είναι συμβατές με την έκδοση του προγράμματος Max/MSP Jitter 6.0 η μεγαλύτερο. Οι προτεινόμενες απαιτήσεις συστήματος για της εφαρμογές είναι διπύρηνος επεξεργαστής (Intel ή AMD) με τουλάχιστον 2 GHz και το περιβάλλον Max Runtime 6 η μεγαλύτερο. Πληκτρολόγιο. Ποντίκι.

## **Βιβλιογραφία/Αναφορές.**

Charles, J. F. (2008), A tutorial on Spectral Sound Processing Using Max/MSP and Jitter. Computer Music Journal 32(σ. 87-102).

Cycling74 : Share, <http://www.cycling74.com/share.html> ( Μαΐου 2013).

Droljc, T. (2011), STFT Analysis Driven Sonographic Sound Processing in Real-Time using Max/MSP and Jitter. University of Hull.

Dudas, R., Lippe, C. (2006), The Phase Vocoder – part i: <https://cycling74.com/2006/11/02/the-phase-vocoder-%E2%80%93-part-i/> (10 Νοεμβρίου 2014).

Διαμαντόπουλος, Τ. (2004), Προγραμματισμός και Σύνθεση Ήχου, Αθήνα: Εκδόσεις Ελλην.

Dodge, C., Jerse T. (1997), Computer Music Synthesis, Composition and Performance second edition. Broadway, New York: Schirmer Books. Boston, Massachusetts: Carnage Learning.

Forums Cycling 74, Jitter Fractals, <http://cycling74.com/project/jitter-fractals/> (18 Αυγούστου 2014).

Ircam Anasynth: AudioSculpt, <http://anasynth.ircam.fr/home/english/software/audiosculpt> (25 Νοεμβρίου 2014).

Max 5 API Reference, JXF file specification, [https://cycling74.com/sdk/MaxSDK-5.1.6/html/chapter\\_\\_jit\\_\\_jxf.html](https://cycling74.com/sdk/MaxSDK-5.1.6/html/chapter__jit__jxf.html) (23 Οκτωβρίου 2014).

Puckette, M. (2007), The Theory and Technique of Electronic Music. Singapore: World Scientific Publishing Co. Pte. Ltd.

Roads, C. (1996), The Computer Music Tutorial. Cambridge, Massachusetts: The MIT Press.

SpecDraw , <http://iem.at/~eckel/science/SpecDraw/SpecDraw.html> (25 Νοεμβρίου 2014).