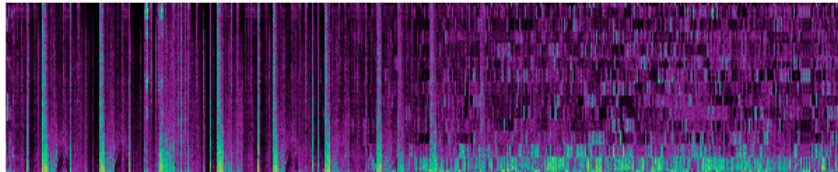


Τ.Ε.Ι. ΚΡΗΤΗΣ

Τμήμα Μουσικής Τεχνολογίας και Ακουστικής

Πτυχιακή εργασία

**«Φασματικές επεξεργασίες ήχου
μέσω ανάλυσης/ανασύνθεσης»**



Στεφάτος Γεώργιος

AM 460

**Επιβλέπων καθηγητής
Βαλσαμάκης Νικόλαος**

Περιεχόμενα

1.Εισαγωγή.....	5
2. Θεωρητική ανάλυση των τεχνικών που χρησιμοποιούνται στις εφαρμογές.....	6
2.0. Η ανάλυση-ανασύνθεση μέσω fft τεχνικής	6
2.1. spectral delay 92.1.1. Φαινόμενα σταθερού χρόνου καθυστέρησης (FIR, comb filtering, delay lines)	11
2.1.2 Φαινόμενα ανατροφοδότησης σταθερού χρόνου καθυστέρησης(IIR).....	13
2.1.3 Φαινόμενα μεταβλητού χρόνου καθυστέρησης	13
2.1.4.random walk	14
2.2. spectral gate	16
2.2.1.gate	16
2.3. spectral waveshaping.....	17
2.3.1.waveshaping	17
2.4. spectral spatialization.....	18
2.4.1.Panning ανάμεσα σε δύο κανάλια	18
2.4.3 Η οκταφωνία	19
3. Θεωρητική ανάλυση των εφαρμογών.....	21
3.1. spectral delay	21
2. spectral gate	24
3.3. spectral waveshaping	25
3.4. spectral spetialization	26
4. Ανάλυση των παραμέτρων.....	27
4.1 spectral delay	27
4.2 Spectral gate	32
4.3 spectral waveshaping	34
4.4 spectral spatialization	36

5. Ανάλυση των αλγορίθμων	38
5.1. spectral delay.....	
5.1.1 Η δημιουργία καθυστέρησης στα μερίσματα	39
5.1.2 Το random walk	45
5.1.3 Το αντικείμενο multislider δύο σειρών	54
5.2 spectral gate	57
5.3 spectral waveshaping	58
5.4 spectral spatialization	60
6. Παραγόμενα ηχητικά φαινόμενα	62
6.1 spectral delay	62
6.1.1. Η παράμετρος delay	63
6.1.2. Η παράμετρος delay όταν μεταβάλλεται από το random walk	65
6.1.2. Η παράμετρος feedback	66
6.3. spectral gate	67
6.3. spectral waveshaping	67
6.4. spectral spatialization	67
7. Συμπεράσματα	68
Βιβλιογραφία	69

1. Εισαγωγή

Η παρούσα εργασία έχει ως βασικό σκοπό να παρουσιάσει και να επεκτείνει της δυνατότητες που δίνει στον σχεδιαστή ήχου η ανάλυση/ανασύνθεση μέσω fft.

Η τεχνική ανάλυσης ανασύνθεσης έχει μακρά ιστορία στην σύνθεση και επεξεργασία ήχου στα πλαίσια εξειδικευμένων ερευνητικών και καλλιτεχνικών εφαρμογών. Τα τελευταία χρόνια έχουν αρχίσει και εμφανίζονται εμπορικές εφαρμογές μουσικής τεχνολογίας που εφαρμόζουν τη παραπάνω τεχνική.

Κάποιες ενδεικτικές εμπορικές μηχανές ανασύνθεσης είναι το CUBE της VirSyn(VirSyn 2004) και το Spectral Delay της Native Instruments(Native Instruments 2001).

Στη παρούσα εργασία έχουν δημιουργηθεί αλγόριθμοι που επεκτείνουν τις τρέχουσες και προϋπάρχουσες εφαρμογές.

Συγκεκριμένα έχουν δημιουργηθεί αλγόριθμοι σύνθεσης ήχου που ασκούν:

- φασματικές σειρές καθυστέρησης (spectral delay)
- φασματική δυναμική επεξεργασία (spectral gate)
- φασματικό μετασχηματισμό των εντάσεων (spectral waveshaping)
- φασματική χωροθέτηση ήχου (spectral spatialization)

2. Θεωρητική ανάλυση των τεχνικών που χρησιμοποιούνται στις εφαρμογές

Προτού αναλύσουμε την λειτουργία των εφαρμογών που αναπτύξαμε στη παρούσα εργασία, πρέπει να καλύψουμε τα θεωρητικά θέματα στα οποία έχουν βασιστεί.

2.0. Η ανάλυση-ανασύνθεση μέσω fft τεχνικής

Η ανάπτυξη όλων των εφαρμογών της εργασίας βασίζεται στην ανάλυση/ανασύνθεση μέσω fft (fast fourier transform).

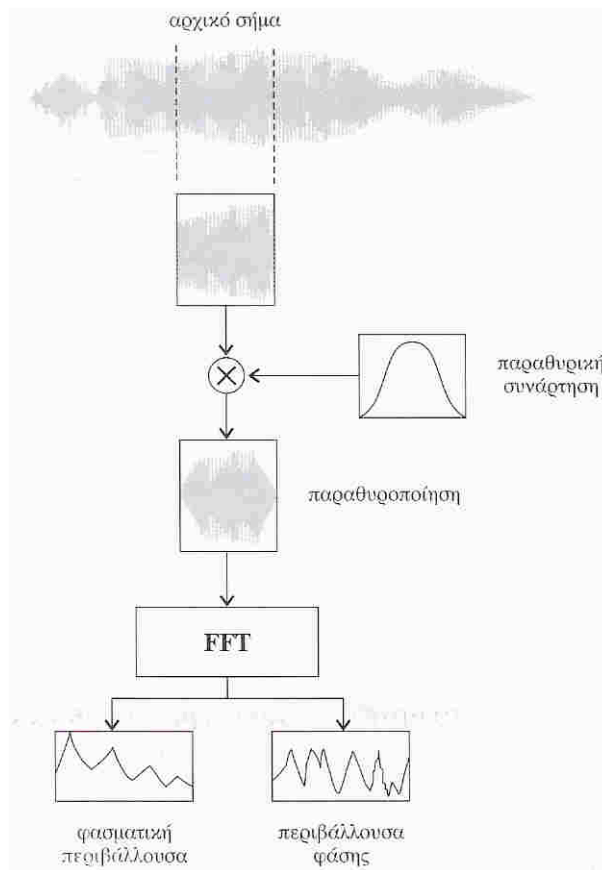
Σύμφωνα με το θεώρημα fourier οποιαδήποτε περιοδική συνάρτηση μπορεί να αναπαρασταθεί ως άθροισμα απλών ημιτόνων. Το θεώρημα fourier βρίσκει εφαρμογή στην ανάλυση-ανασύνθεση μέσω fft στην οποία η περιοδική ή η μη περιοδική κυματομορφή του σήματος χωρίζεται σε μικρά τμήματα τα οποία θεωρούνται περιοδικά και στα οποία εφαρμόζεται το θεώρημα fourier ώστε να εξαχθούν συμπεράσματα για το φάσμα που αντιστοιχεί σε αυτά (Miranda 1994).

Η fft ανάλυση και η fft ανασύνθεση (inverse fast fourier transform) είναι δύο διαφορετικές διαδικασίες. Στην fft ανάλυση το σήμα αναλύεται σε μία σειρά ημιτόνων για το καθένα από τα οποία επιστρέφει μία τιμή πλάτους και μία τιμή φάσης. Στην fft ανασύνθεση τα δεδομένα της ανάλυσης ανασυντίθεται στη κυματομορφή που αντιστοιχεί στη σειρά αυτή των ημιτόνων.

Στο πρώτο στάδιο της ανάλυσης του σήματος η κυματομορφή του σήματος χωρίζεται σε τμήματα ανά ένα σταθερό αριθμό δειγμάτων.

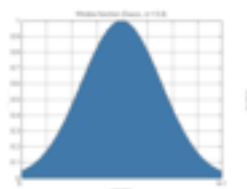
Τα τμήματα αυτά ονομάζονται παράθυρα (frames) και ο αριθμός των δειγμάτων που μπορεί να περιέχει μπορεί να είναι μόνο δύναμη του 2 .

Το κάθε ένα από αυτά τα παράθυρα προτού αναλυθεί πολλαπλασιάζεται με μία παραθυρική συνάρτηση (window function) για να αποφευχθεί η είσοδος επιπλέον αρμονικών στο σήμα.

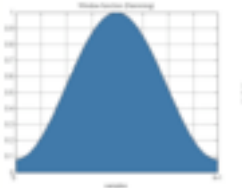


Σχήμα 1 παραθυροποίηση του σήματος (Διαμαντόπουλος 2005, σελ 232)

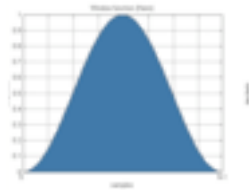
Παρακάτω παρατίθενται μερικοί τύποι παραθυρικών συναρτήσεων που χρησιμοποιούνται συχνά κατά τη διαδικασία της παραθυροποίησης:



gaus



hamming



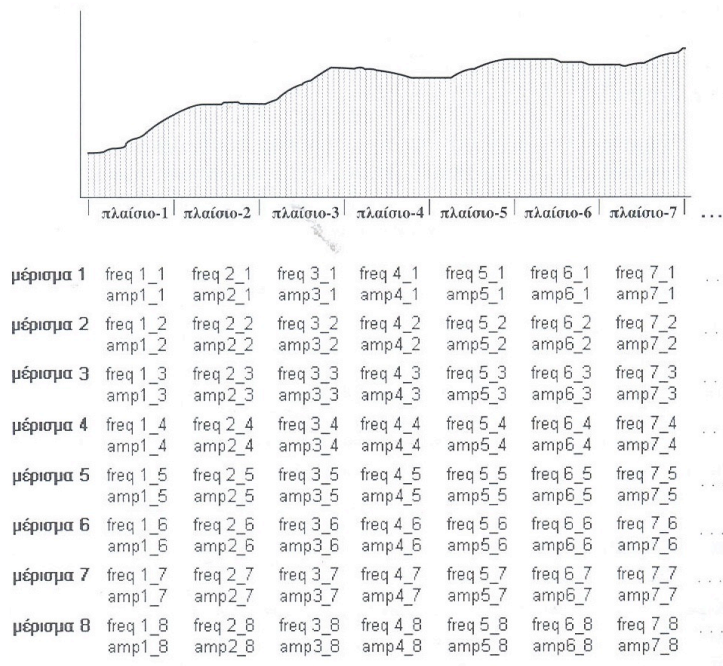
hanning



lanczos

Σχ.2 παραθυρικές συναρτήσεις (wikipedia)

Ένα παράθυρο n αριθμού δειγμάτων θα αναλυθεί σε $n/2$ αριθμό περιοχές φάσματος που θα βρίσκονται στη φασματική περιοχή 0 Hz έως $sr/2$ Hz (sr συχνότητα Nyquist - όπου sr θα θεωρούμε από εδώ και στο εξής τη τιμή 44100) και άλλες $n/2$ περιοχές φάσματος περιοχές οι οποίες θα βρίσκονται στη φασματική περιοχή $sr/2$ Hz έως sr Hz αλλά δε θα χρησιμοποιούνται.



Σχήμα 3 . ανάλυση κυματομορφής σε μερίσματα (Διαμαντόπουλος 2005, σελ 234)

Οι περιοχές αυτές ονομάζονται μερίσματα (bins) και αντιστοιχούνται σε ημίτονα μίας συγκεκριμένης συχνότητας για το καθένα από τα οποία, τα αποτελέσματα της ανάλυσης fourier

επιστρέφουν μία τιμή πλάτους και μία τιμή φάσης(Roads Pope,Piccialli,Depdi 1997).

Η μέγιστη τιμή πλάτους για κάθε μέρισμα δίνεται από το τύπο:

$$\frac{\text{μέγεθος παραθύρου}}{(\text{σύνολο σημείων στη παραθυρική συνάρτηση/ μέγεθος μεταπήδησης})^{1/2}}$$

Στη παρούσα εφαρμογή που το μέγεθος του παραθύρου είναι 1024 και το μέγεθος μεταπήδησης 2, το μέγιστο πλάτος του κάθε μερίσματος ισούται με 650.

Το κάθε μέρισμα συμβολίζει ένα ημίτονο διαφορετικής συχνότητας η οποία βρίσκεται από το τύπο

*sr/window-size*αύξοντα-αριθμό-μερίσματος,*

διαφορετικού πλάτους και φάσης.

Η τιμή πλάτους και φάσης που επιστρέφει η ανάλυση για κάθε μέρισμα εκφράζεται μέσω ενός μιγαδικού αριθμού.

Ο μιγαδικός εκφράζεται μαθηματικά μέσω του τύπου

$$y = \alpha * i + \beta$$

(Γαγάλης,Γλαμπεδάκης,Κωτσάκης,Κομισόπουλος,Λευκαδίτης, Τραχάλιος 1992)

όπου y ο μιγαδικός,

$\alpha * i$ το φανταστικό μέρος του μιγαδικού και

β το πραγματικό μέρος

το πλάτος του μερίσματος δίνεται από τη μαθηματική έκφραση

$$|y| = (\alpha^2 + \beta^2)^{1/2}$$

και η φάση του θ από το τύπο :

$$\eta\mu\theta = \beta / (\alpha^2 + \beta^2)^{1/2}$$

Κατόπιν μέσω της αντίστροφης διαδικασίας (ifft) το άθροισμα των μερισμάτων θα πάρει τη μορφή της αντίστοιχης κυματομορφής.

Ο συνολικός ήχος επιτυγχάνεται μέσω της διαδοχής των στατικών αυτών φασμάτων όπως σε μία ταινία όπου η οπτική αίσθηση της κίνησης δίνεται από τη διαδοχή στατικών εικόνων-παραθύρων.

Εάν το χρονικό παράθυρο είναι πολύ μεγάλο, το σήμα θα αναλυθεί σε περισσότερες ζώνες που σημαίνει ότι θα έχουμε πιο πιστή αναπαράσταση του φάσματος αλλά θα χάσει τη «χρονική λεπτομέρεια».

Αντιθέτως εάν το σήμα αναλυθεί σε λιγότερες ζώνες ώστε το χρονικό παράθυρο να είναι μικρότερο, θα κερδίσουμε σε χρονική λεπτομέρεια και θα χάσουμε τη φασματική λεπτομέρεια (Roads 1999).

Η ασυμβατότητα αυτή ανάμεσα στη φασματική και τη χρονική λεπτομέρεια λύνεται μέσω της αλληλοεπικάλυψης των χρονικών παραθύρων (κάτι το οποίο οδηγεί σε περισσότερους μαθηματικούς υπολογισμούς στο χρόνο δηλαδή περισσότερη υπολογιστική ισχύ).

Όταν τα παράθυρα αλληλεπικαλύπτονται, το καθένα από αυτά δεν καλύπτει αποκλειστικά τα δείγματα της δικής του περιοχής, αλλά και άλλων περιοχών που ανήκουν σε γειτονικά παράθυρα. Το μέγεθος της περιοχής που καλύπτει από το γειτονικό του παράθυρο, ονομάζεται μέγεθος μεταπήδησης (hop-size) και μπορεί να δεχτεί τιμές ίσες με το μέγεθος παραθύρου/2^v.

Εάν για παράδειγμα μέγεθος μεταπήδησης ισούται με 256 δείγματα, το πρώτο παράθυρο θα καλύπτει τη πρώτη περιοχή με τα αντίστοιχα χρονικά δείγματα 0-1023, το αμέσως επόμενο αντί να ξεκινήσει από το δείγμα 1024 και να καταλήξει στο 2047, θα καλύπτει τη περιοχή που αντιστοιχεί στα δείγματα 256 – 1279, το τρίτο τη περιοχή δειγμάτων 512 – 1536, το τέταρτο τη περιοχή 768 – 1791 κοκ καταλήγοντας να πληρούμε και τις δύο προϋποθέσεις, καλύτερης συχνότητας (με μέγεθος παραθύρου ανάλυσης 1024) αλλά και καλύτερης ανάλυσης χρόνου (ίση με 256 δείγματα).

Το μεγάλο πλεονέκτημα που δίνει στο σχεδιαστή ήχων η ανάλυση/ανασύνθεση fourier είναι ότι αφού το σήμα αναλυθεί κατά την είσοδό του στη μηχανή διαμόρφωσης σε (μέγεθος παραθύρου/2) μερίσματα, προτού ανασυντεθεί, υπάρχει η δυνατότητα να επέμβουμε στα δεδομένα του κάθε μερίσματος ξεχωριστά δηλαδή να επέμβουμε στη φασματική πληροφορία του σήματος.

2.1. spectral delay

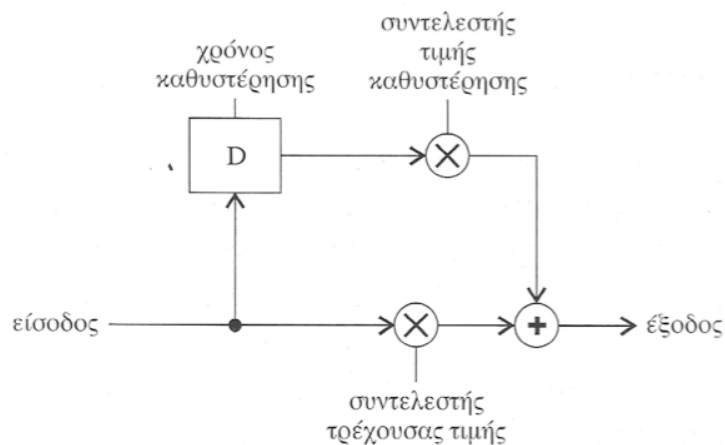
Η εφαρμογή spectral delay είναι ένας συνδυασμός των τεχνικών fft ανάλυσης-ανασύνθεσης και των φαινομένων που παρουσιάζονται κατά την είσοδο του σήματος σε κύκλωμα σταθερού ή μεταβλητού χρόνου καθυστέρησης.

2.1.1. Φαινόμενα σταθερού χρόνου καθυστέρησης (FIR, comb filtering, delay lines)

Τα φαινόμενα σταθερού χρόνου καθυστέρησης προέρχονται από τη τεχνική της αποθήκευσης του σήματος και της εκ νέου αναπαραγωγής και πρόσθεσής του στο πρωτότυπο μετά από προκαθορισμένο χρόνο.

Καθ' όλη τη διάρκειά του φαινομένου ο χρόνος παραμένει σταθερός.

Το κύκλωμα σταθερής καθυστέρησης χρόνου έχει την ακόλουθη διάταξη:



Σχήμα 4 κύκλωμα σταθερής καθυστέρησης χρόνου(Διαμαντόπουλος 2005, σελ 336)

Τα φαινόμενα που παρουσιάζονται, προσδιορίζονται ανάλογα με τη χρονική καθυστέρηση(d) και χωρίζονται σε τρεις κύριες κατηγορίες(Διαμαντόπουλος 2005) :

Για $d < 10$ msec

Όταν ο χρόνος καθυστέρησης d ισούται με ένα μικρό αριθμό δειγμάτων, τότε παρουσιάζονται φαινόμενα διέλευσης του σήματος από χαμηλοπερατό φίλτρο (lowpass filter).

Για μεγαλύτερες τιμές παρουσιάζονται φαινόμενα τύπου κτένας (comb filter).

10 msec < d < 50 msec

Στη περίπτωση αυτή δημιουργείται ένα ψυχοακουστικό φαινόμενο (φαινόμενο Haas) σύμφωνα με το οποίο η ανθρώπινη αντίληψη αδυνατεί να διαχωρίσει τα δύο όμοια σήματα λόγω της μικρής χρονικής τους διαφοράς και τα αντιλαμβάνεται ως ένα ενιαίο το οποίο δίνει τη ψευδαίσθηση του μεγαλύτερου όγκου στο σήμα.

Σε αυτό το ψυχοακουστικό φαινόμενο βασίζεται η κατασκευή ψηφιακών προσομοιωτών χώρων.

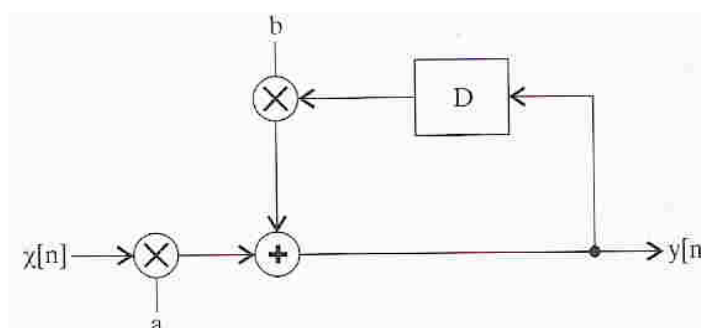
$d > 50$ msec

Ο εγκέφαλος αντιλαμβάνεται δύο ξεχωριστά σήματα δηλαδή επανάληψη του ίδιου σήματος (όπως ο αντίλαλος)

Όπως γίνεται αντιληπτό με την ίδια διάταξη μπορούμε να πετύχουμε μία πληθώρα ηχητικών φαινομένων.

2.1.2 Φαινόμενα ανατροφοδότησης σταθερού χρόνου καθυστέρησης (IIR)

Επεκτείνοντας την παραπάνω διάταξη, δημιουργούμε μία πιο σύνθετη, προσθέτοντας στην είσοδο του σήματος την έξοδό του πολλαπλασιασμένη με ένα συντελεστή μικρότερο του ένα (καθώς στην αντίθετη περίπτωση θα επέλθει υπεροδήγηση-απειρισμός του πλάτους του σήματος καθώς θα έχουμε ένα συνεχή πολλαπλασιασμό).



Σχήμα 5 ανατροφοδοτούμενο κύκλωμα σταθερής καθυστέρησης χρόνου
(Διαμαντόπουλος 2005, σελ 338)

Όπου $x(n)$ το σήμα εισόδου και a ο συντελεστής με τον οποίο πολλαπλασιάζεται $y(n)$ το σήμα εξόδου και b ο συντελεστής με τον οποίο πολλαπλασιάζεται και D ο χρόνος καθυστέρησης.

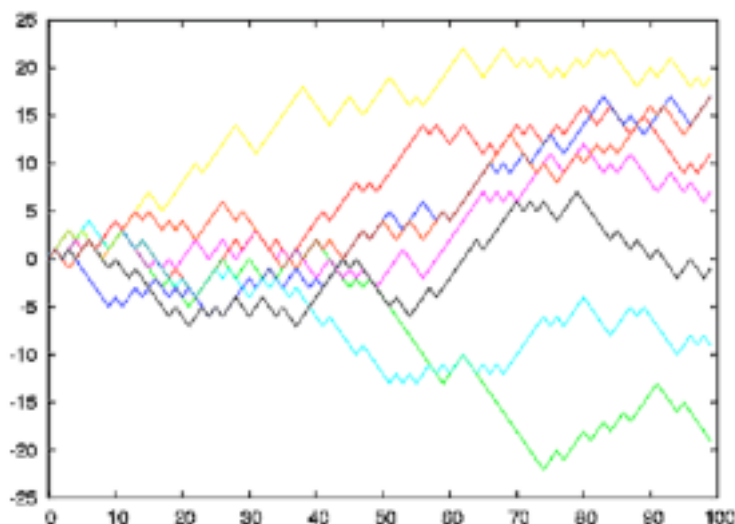
2.1.3 Φαινόμενα μεταβλητού χρόνου καθυστέρησης

Σε αυτή τη διάταξη, ο χρόνος καθυστέρησης δεν είναι σταθερός αλλά μεταβάλλεται με ποικίλες διαδικασίες.

Για παράδειγμα όταν μεταβάλλουμε το χρόνο καθυστέρησης με μία συνάρτηση τότε πετυχαίνουμε ηχητικά φαινόμενα που είναι γνωστά με την ονομασία «flanger».

2.1.4. random walk

Το random walk είναι μία μαθηματική έννοια η οποία στην πιο ευρεία της μορφή περιγράφει τη πορεία μίας τροχιάς όπου αλλάζει διαδοχικά τη διεύθυνσή της με τυχαίο τρόπο (Serra 1993).



Σχ.6 η κίνηση 8 διαφορετικών τροχιών random walk (wikipedia)

Τη μαθηματική αυτή έννοια τη συναντάμε σε ένα ευρύ πεδίο επιστημών από τη φυσική και τη μικροβιολογία, μέχρι την οικολογία και την οικονομία.

Ο Ξενάκης τη χρησιμοποίησε ως μέσο για τη σύνθεση ήχου στο έργο του GENDY.

Στην πιο απλή της μορφή, μία τροχιά ξεκινάει από μία αρχική τιμή και αλλάζει τη διεύθυνσή της επιλέγοντας ανάμεσα σε δύο αντιδιαμετρικές διευθύνσεις, με ισοπίθανο τρόπο (Serra 1993).

Το βήμα με το οποίο διανύει τη κάθε απόσταση μεταξύ δύο διαδοχικών αλλαγών διεύθυνσης, αποκαλείται βήμα (step) του random walk.

Επειδή θέλουμε συνήθως η τροχιά αυτή να δέχεται τιμές μέσα σε ένα προκαθορισμένο σύνολο τιμών περιορίζουμε τη κίνησή της θέτοντας δύο όρια, το πάνω και το κάτω, τα οποία αποκαλούνται barriers. Όταν μία τιμή πάει να ξεπεράσει το όριο, ανακλάται από αυτό αλλάζοντας τη διεύθυνσή του.

Εάν y_n είναι η τρέχουσα τιμή του random walk τότε η επόμενη αριθμητική τιμή y_{n+1} θα δίνεται από το τύπο:

$$y_{n+1} = y_n * (\pm 1) * \text{τιμή βήματος}$$

Εάν υπήρχε ως κάτω όριο το 20 και η αρχική τιμή ήταν το 22 τότε θα μπορούσε να επιλέξει ως μία σειρά τιμών:

Βήμα		+1	-1	+1	-1	-1	+1	+1
Random walk	22	23	22	21	20	21	22	21...

Όπως παρατηρούμε η 5^η τιμή που είχε τη τιμή 20 και κατόπιν μειώθηκε κατά 1, δε πήρε τη τιμή 19 αλλά τη τιμή 21 επειδή ανακλάται από το όριο.

2.2. spectral gate

2.2.1. δυναμική επεξεργασία σήματος

Η επεξεργασία των διακυμάνσεων της δυναμικής ενός σήματος μας επιτρέπει να ελαττώσουμε τα ανεπιθύμητα φαινόμενα που εμφανίζονται κατά την αναπαραγωγή του (υπεροδήγηση του σήματος, θόρυβος).

Το gate είναι ένας επεξεργαστής που αποσκοπεί στην αποκοπή των χαμηλών δυναμικών

(Roads Rope, Piccialli, Depdi 1997).

Στην πιο απλή του μορφή ρυθμίζεται μέσω μόνο μίας παραμέτρου, το όριο του gate, η οποία ρυθμίζει πότε το σήμα θα περάσει στην έξοδο του επεξεργαστή και πότε όχι.

Αν το πλάτος του σήματος σε κάθε χρονική στιγμή είναι μεγαλύτερο από τη τιμή του ορίου τότε η τιμή του πλάτους δρομολογείται στην έξοδο του επεξεργαστή τη συγκεκριμένη χρονική στιγμή.

Εάν είναι μικρότερη από τη τιμή του ορίου τότε αντικαθίσταται από τη τιμή 0 στην έξοδο του επεξεργαστή.

Με αυτό το τρόπο έχουμε τη δυνατότητα να αποκόψουμε χαμηλές δυναμικές όταν αυτές είναι υπεύθυνες για την είσοδο θορύβου στο σήμα.

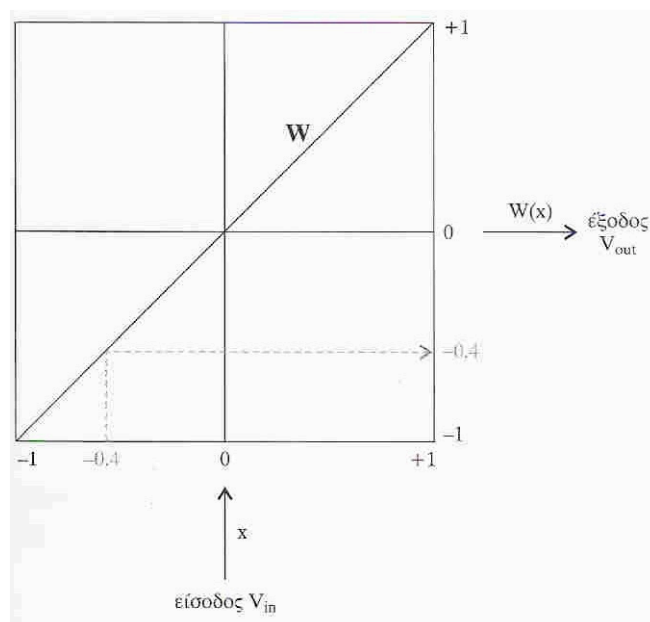
2.3. spectral waveshaping

2.3.1. waveshaping

Η τεχνική μετασχηματισμού των τιμών δυναμικής ενός σήματος χρησιμοποιείται ως μέθοδος για την δημιουργική παραμόρφωση του.

Ο μετασχηματισμός τιμών γίνεται σε μία διάταξη που έχει την ιδιότητα για κάθε τιμή της δυναμικής του σήματος εισόδου να δίνει μία νέα τιμή δυναμικής (Miranda 1998).

Ο μετασχηματισμός τιμών γίνεται βάση μίας συνάρτησης μετασχηματισμού (transfer function) που για κάθε τιμή εισόδου x , δίνει μία τιμή εξόδου $\psi=f(x)$.



Σχήμα 7 γραμμική συνάρτηση μετασχηματισμού (Διαμαντόπουλος 2005, σελ 309)

Εάν η συνάρτηση είναι γραμμική $\psi=k \cdot x$

τότε το σήμα θα αλλάξει το πλάτος.

Μία συχνή περίπτωση μετασχηματισμού κυματομορφής δημιουργείται όταν το σήμα υπεροδηγείται με αποτέλεσμα ένα μεγάλο μέρος του δυναμικού εύρους του σήματος να αντικαθίσταται με τη μέγιστη δυναμική του, δημιουργώντας παραμόρφωση (distortion).

2.4. spectral spatialization

2.4.1. Panning ανάμεσα σε δύο κανάλια

Για να μπορέσουμε να αναπαραστήσουμε τον ήχο που παράγει ένα ηχητικό αντικείμενο καθώς κινείται στον οκταφωνικό χώρο πρέπει πρώτα να κατανοήσουμε πως επιτυγχάνεται η κίνησή του στον στερεοφωνικό χώρο.

Θα μελετήσουμε την ιδανική περίπτωση όπου ο ακροατής βρίσκεται στο κέντρο ενός νοητού κύκλου, σε μία ακτίνα με γωνία μικρότερη από 90° του οποίου έχουν τοποθετηθεί δύο ηχεία.

Για να δημιουργήσουμε την ψευδαίσθηση της κίνησης του ηχητικού αντικειμένου, πρέπει όταν η ένταση του πρώτου ηχείου αυξάνεται, η ένταση του δεύτερου να μειώνεται.

Η τεχνική αυτή ονομάζεται panning και εξαρτάται από το είδος της αυξομείωσης της έντασης ανάμεσα στα δύο κανάλια (Roads 1999).

Εάν αυτή είναι γραμμική, τότε όταν η ένταση του πρώτου ηχείου ισούται με $\alpha \text{ amp}$ η ένταση στο δεύτερο θα ισούται με $(1-\alpha) \text{ amp}$ όπου $\alpha \leq 1$.

Ένα πρόβλημα που εμφανίζεται στη γραμμική μετάβαση έντασης είναι ότι ανάλογα με τη γωνία που θέλουμε να εμφανίσουμε το ηχητικό αντικείμενο στο στερεοφωνικό χώρο, αλλάζει η τιμή της συνολικής έντασης, γεγονός που ο ακροατής το αντιλαμβάνεται σαν μία απομάκρυνση του ηχητικού αντικειμένου.

Προκειμένου να διατηρήσουμε τη συνολική ένταση που αντιλαμβάνεται ο ακροατής σταθερή, εφαρμόζουμε ημιτονοειδείς μεταβάσεις ανάμεσα στα δύο ηχεία οι οποίες ονομάζονται και μεταβάσεις σταθερής ενέργειας.

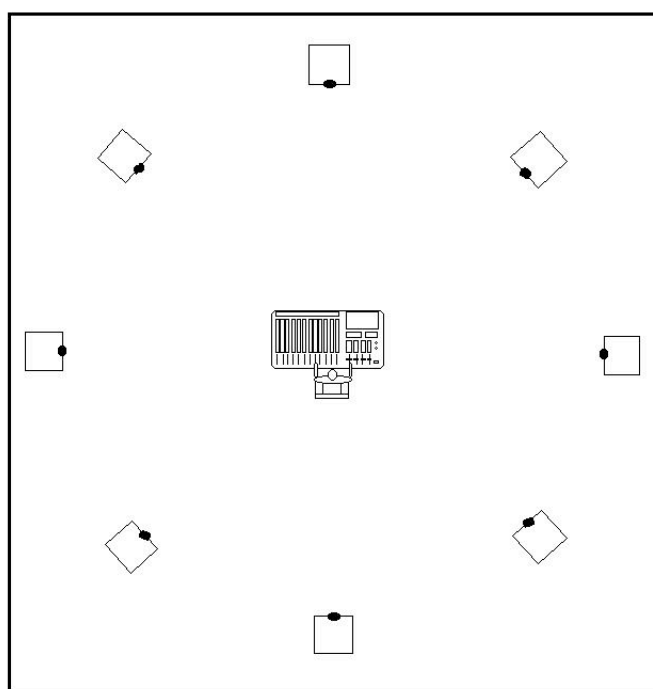
Για να δημιουργηθεί ημιτονοειδής μετάβαση ανάμεσα στα δύο ηχεία, πολλαπλασιάζεται το πλάτος του σήματος που οδηγείται σε αυτά με μία παραθυρική συνάρτηση ημιτόνου. Η παραθυρική συνάρτηση ημιτόνου περιέχει μόνο τη θετική φάση μίας περιόδου της συνάρτησης ημιτόνου.

Κατ' αντιστοιχία με τη γραμμική μετάβαση, όταν η ένταση του ενός ηχείου αυξάνεται, η ένταση του άλλου θα φθίνει. Για να επιτευχθεί αυτό, η παραθυρική συνάρτηση ημιτόνου διαβάζεται με διαφορά φάσης 180° ανάμεσα στα δύο κανάλια.

2.4.3 Η οκταφωνία

Οι διατάξεις πολυκάναλων συστημάτων με χρήση πολλαπλών ηχείων, έχουν αναπτυχθεί με σκοπό την δημιουργία σύνθετης ηχητικής κίνησης στο χώρο μέσω της διάχυσής του σήματος στο πολυκάναλο πεδίο (diffusion), καθώς και της συνύπαρξης πολλών σημάτων ταυτόχρονα (Wyatt 1999).

Έχουν αναπτυχθεί διάφορων τύπων διατάξεις οκταφωνίας. Μία διάταξη οκταφωνίας που χρησιμοποιείται συχνά, παρατίθεται στο επόμενο σχήμα:



Σχήμα 8 οκταφωνική διάταξη (Wyatt 1999)

Ας υποθέσουμε ότι θέλουμε ένα σήμα να εκτελέσει μία πλήρη περιστροφή στον οκταφωνικό χώρο.

Το σήμα θα εκτελέσει κίνηση rapping ανάμεσα στο κάθε γειτονικό ζεύγος ηχείων που σχηματίζεται στον οκταφωνικό χώρο, διαδοχικά. Η κίνηση rapping στο κάθε ζεύγος θα εκτελείται με τον ίδιο τρόπο με τον οποίο εκτελείται ανάμεσα σε δύο ηχεία στο στερεοφωνικό χώρο.

Για να επιτύχουμε ομαλή μετάβαση εφαρμόζουμε καμπύλες σταθερής ενέργειας καθώς η κίνηση που εκτελεί ένα σώμα που περιστρέφεται με κέντρο τον άξονα του κορμού μας είναι ίδια

μεταξύ δύο σημείων που το τόξο που καλύπτουν είναι μικρότερο από 90° .

Επειδή τα ηχεία είναι 8, η διαφορά φάσης με την οποία θα διαβάζεται η παραθυρική συνάρτηση ημιτόνου ανάμεσα σε δύο γειτονικά κανάλια, θα ισούται με $360^{\circ}/8 = 45^{\circ}$.

3. Θεωρητική ανάλυση των εφαρμογών

Έχοντας καλύψει το θεωρητικό υπόβαθρο των βασικών τεχνικών που χρησιμοποιούνται στις 4 εφαρμογές, σε αυτό το μέρος θα γίνει η παρουσίαση της λειτουργία τους.

3.1. spectral delay

Το spectral delay είναι μία μονάδα επεξεργασίας του ήχου, η οποία χωρίζει το σήμα σε φασματικές περιοχές κάθε μία από τις οποίες έχει διαφορετικό χρόνο καθυστέρησης.

Το σήμα κατά την είσοδό του στη μονάδα επεξεργασίας spectral delay αναλύεται σύμφωνα με την fft τεχνική σε παράθυρα που αντιστοιχούν σε $\text{window-size}/2$ μερίσματα. Η τιμή πλάτους και φάσης που αντιστοιχεί σε κάθε μέρισμα του κάθε παραθύρου αποθηκεύεται σε μία μνήμη η οποία έχει χωρητικότητα N παράθυρα δηλαδή $N \cdot \text{μέγεθος παραθύρου samples}$.

Μία μονάδα delay καθυστερεί τη πληροφορία πλάτους-φάσης του τρέχοντος μερίσματος κατά $f(N-n)$, όπου f το παράθυρο, N ο αύξων αριθμός του παραθύρου και n ο αριθμός των παραθύρων που θα ανατρέξει για να βρει τη πληροφορία πλάτους φάσης. Με αυτό το τρόπο οδηγείται στην έξοδο η πληροφορία που υπήρχε στο ίδιο μέρισμα κατά n παράθυρα πριν. Το κάθε μέρισμα μπορεί να έχει ξεχωριστή τιμή καθυστέρησης.

Για παράδειγμα, το N_1 παράθυρο θα μπορούσε να εμφανίζει κατά τη διαδικασία ανασύνθεσης (ifft):

το 1^ο μέρισμα με τη πληροφορία πλάτους-φάσης που υπήρχε στο 1^ο μέρισμα του N_1-200 παραθύρου,

το 2^ο μέρισμα με τη πληροφορία που υπήρχε στο 2^ο μέρισμα του N_1-150 παραθύρου,

το 3^ο μέρισμα με τη πληροφορία που υπήρχε στο 3^ο μέρισμα του N_1-300 παραθύρου

*·
·
·*

στο 512^ο μέρισμα τη πληροφορία που υπήρχε στο 512^ο μέρισμα του N_1-230 παραθύρου.

Ταυτόχρονα το κάθε μέρισμα ανατροφοδοτείται (feedback) ξεχωριστά. Δηλαδή το μέρισμα λαμβάνει στην είσοδό του το σήμα της εξόδου του πολλαπλασιασμένο με ένα συντελεστή.

Με αυτό το τρόπο έχουμε δημιουργήσει (μέγεθος παραθύρου/2) αριθμό φασματικών ζωνών όπου η κάθε φασματική ζώνη έχει το δικό της ανατροφοδοτούμενο χρόνο καθυστέρησης.

Ο χρήστης επιλεγεί από την γραφική διεπαφή (interface) της εφαρμογής ποια θέλει να είναι η καθυστέρηση που θα έχει κάθε μέρισμα ξεχωριστά καθώς και ο συντελεστή ανατροφοδότησής του.

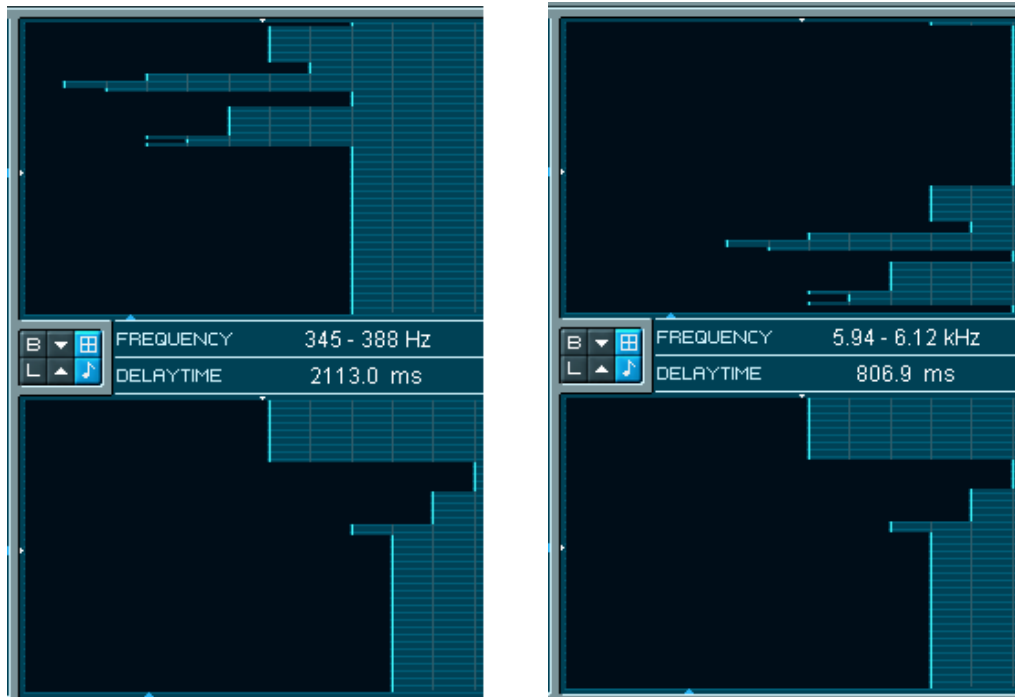
Τόσο στην εφαρμογή της native instruments όσο και στη δική μας η καταχώρηση των τιμών γίνεται μέσω ενός γραφικού περιβάλλοντος-αντικειμένου που ονομάζεται multislider και δίνει τη δυνατότητα στο χρήστη να μεταβάλλει το πλάτος εκατοντάδων sliders (μπάρες μεταβολής τιμών) ταυτόχρονα.

Ταυτόχρονα, έχουμε τη δυνατότητα να μεταβάλλουμε το χρόνο καθυστέρησης του κάθε μερίσματος ξεχωριστά με αυτοματοποιημένο τρόπο επιτυγχάνοντας με αυτό το τρόπο φαινόμενα καθυστέρησης μεταβλητού χρόνου.

Ο τρόπος με τον οποίο το επιτυγχάνουμε αυτό ποικίλει και αποτελεί την βασική διαφορά από την αντίστοιχη εφαρμογή της native instruments.

Ένα από τα βασικά χαρακτηριστικά του spectral delay της NI είναι η δυνατότητα οι χρόνοι αυτοί να είναι μεταβλητοί μέσω μίας μαζικής κίνησης όλων των παραμέτρων ταυτόχρονα, μετακινώντας την εικόνα τους προς μία διεύθυνση.

Στο παρακάτω σχήμα φαίνεται η κίνηση που έχουν οι sliders της τιμής των delay του ενός καναλιού καθώς μετατοπίζονται ταυτόχρονα ως «εικόνα» προς μία κατεύθυνση ενώ στο άλλο κανάλι παραμένουν σταθερά.



Σχήμα 9 ομαδοποιημένη κίνηση των sliders στο ένα κανάλι

Η κίνηση αυτή μπορεί να γίνει είτε χειροκίνητα είτε αυτοματοποιημένα μέσω μίας περιοδικής κίνησης η οποία επιτυγχάνεται μέσω μίας προκαθορισμένης κυματομορφής.

Στην δική μας εφαρμογή δοκιμάσαμε την ιδέα ο κάθε ξεχωριστός slider που ελέγχει τη καθυστέρηση του παραθύρου να κινείται μέσω ενός διαφορετικού random walk. Ένα multislider δύο περιοχών καθορίζει το πάνω και το κάτω όριο της κίνησής random walk του κάθε slider ξεχωριστά. Ταυτόχρονα υπάρχει η δυνατότητα ομαδοποίησης των sliders ώστε ανά συγκεκριμένο αριθμό, να εκτελούν ίδια κίνηση.

Με αυτό το τρόπο επιτυγχάνουμε η μεταβολή της ανακατανομής γειτονικών μερισμάτων, να είναι η ίδια.

Τέλος έχουμε δώσει τη δυνατότητα στο χρήστη να επιλέγει μέσω του multislider δύο περιοχών ποιες φασματικές περιοχές θα παραμείνουν σταθερές και ποιες όχι.

3.2. spectral gate

Η εφαρμογή spectral gate δίνει τη δυνατότητα στο χρήστη να ελέγξει την διέλευση της πληροφορίας του κάθε μερίσματος.

Αντίστοιχα με το spectral delay και σε αυτή την περίπτωση το σήμα κατά την είσοδό του στη μονάδα spectral gate αναλύεται σε παράθυρα που αντιστοιχούν σε $\text{window-size}/2$ μερίσματα.

Σε κάθε μερίσμα καταχωρείται και μία διαφορετική τιμή που αποτελεί τη τιμή του ορίου με βάση το οποίο θα λειτουργεί το gate κάθε μερίσματος.

Ο τρόπος λειτουργίας του gate είναι ο εξής: Ένας λογικός έλεγχος συντελείται στην έξοδο της τιμής πλάτους του τρέχοντος μερίσματος συγκρίνοντας το με το αντίστοιχο πλάτος που έχουμε καταχωρήσει εξωτερικά μέσω του multislider. Εάν το πλάτος ενός μερίσματος είναι μεγαλύτερο από τη τιμή του ορίου του gate, τότε δρομολογείται στην έξοδο με το πλάτος που είχε πριν τη σύγκριση.

Εάν το πλάτος ενός μερίσματος του κάθε frame είναι μεγαλύτερο από τη τιμή που του αντιστοιχείται τότε δρομολογείται στην έξοδο με μηδενικό πλάτος.

Με αυτό τον τρόπο το spectral gate μπορεί να χρησιμεύσει και ως ένας αποθορυβοποιητής που θα έχει τη δυνατότητα να λειτουργήσει ανάλογα με τα φασματικά χαρακτηριστικά του θορύβου βάθους. Προσαρμόζοντας ανάλογα τις τιμές gate των μερισμάτων ελέγχουμε το μέγεθος του θορύβου που θα αποκόψουμε ανάλογα με το μέγεθος του θορύβου που υπάρχει σε κάθε φασματική περιοχή. Λόγω του ότι έχουμε μεγάλο αριθμό φασματικών περιοχών (512 στη συγκεκριμένη εφαρμογή όσος και ο αριθμός των μερισμάτων), έχουμε τη δυνατότητα και μεγαλύτερης λεπτομέρειας στην αποκοπή θορύβου.

3.3. spectral waveshaping

Η τεχνική spectral waveshaping, μετασχηματίζει τις τιμές των δυναμικών του κάθε μερίσματος ξεχωριστά.

Η λογική με την οποία επιτυγχάνεται η τεχνική spectral waveshaping στο φάσμα είναι η ίδια με αυτή του μετασχηματισμού της χρονικής κυματομορφής. Επειδή όμως το σήμα έχει αναλυθεί σε φασματικές ζώνες-μερίσματα η τεχνική waveshaping εφαρμόζεται στο πλάτος της κάθε φασματικής ζώνης ξεχωριστά.

Στην εφαρμογή μας spectral waveshaping, η συνάρτηση μετασχηματισμού χρησιμεύει για να μετατρέψει το τρέχων εισερχόμενο πλάτος του μερίσματος v με τιμή χ_v στη τιμή ψ_v .

Για να κατανοήσουμε τα ηχητικά φαινόμενα που εμφανίζονται από το μετασχηματισμό στα πλάτη των μερισμάτων μέσω της τεχνικής waveshaping, πρέπει να μελετήσουμε πώς κατανέμεται το πλάτος της κυματομορφής στα μερίσματα. Η τιμή του πλάτους κατά την ανάλυση με τη τεχνική fft διαμοιράζεται σε όλες τις φασματικές ζώνες.

Αυτό σημαίνει ότι, εάν για παράδειγμα, το μέσο πλάτος για ένα ηχητικό γεγονός που αναλύεται με μέγεθος παραθύρου 1024, έχει υπολογιστεί στα 60 dB RMS είναι ποιο πιθανό στατιστικά η τιμή του πλάτους που θα έχει το μερίσμα v τη χρονική στιγμή t να έχει τη τιμή 0.0575 dB RMS. Αυτό σημαίνει ότι οι τιμές των πλατών θα είναι μετατοπισμένες σε εξαιρετικά χαμηλές τιμές. Η μέγιστη δυναμική ενός σήματος ή ακόμα και μία μέση τιμή της είναι εξαιρετικά απίθανο να εμφανιστεί σε κάποιο μερίσμα σε αντίθεση με τη χρονική κυματομορφή του που είναι πολύ πιθανό.

Εάν εφαρμοστεί η τεχνική waveshaping στην αριθμητική τιμή του πλάτους του κάθε μερίσματος ξεχωριστά, μία πολύ μικρή αύξηση των ποιο χαμηλών τιμών πλάτους, θα οδηγήσει σε μία δραματική αύξηση του συνολικού πλάτους του σήματος.

Ένας επιπλέον λόγος εμφάνισης θορύβου από την αύξηση των χαμηλών δυναμικών, οφείλεται στο ότι στις χαμηλές χαμηλές δυναμικές κρύβονται συχνά θόρυβοι δυσδιάκριτοι ή ακόμα και μη ακουστοί λόγω της διαδικασίας ηχογράφησης ή/και της ψηφιοποίησης.

3.4. spectral spetialization

Η φασματική οκταφωνία είναι μία εφαρμογή που δίνει τη δυνατότητα στο χρήστη να τοποθετήσει τις τιμές των μερισμάτων της fft ανάλυσης/ανασύνθεσης και κατ' επέκταση διαφορετικές φασματικές ζώνες ενός αρχείου ήχου, στον οκταφωνικό χώρο. Ταυτόχρονα έχει τη δυνατότητα να ελέγχει τη κίνησή τους στον οκταφωνικό χώρο μέσω της κίνησης random walk.

Το σήμα χωρίζεται σε παράθυρα μέσω της ανάλυσης ανασύνθεσης μέσω fft και το κάθε μέρισμα αποκτάει διαφορετική θέση μέσα στην οκταφωνία. Με αυτό το τρόπο επιτυγχάνουμε μία κατανομή του φάσματος μέσα στο χώρο.

Ταυτόχρονα μέσω του random walk δίνουμε ξεχωριστή κίνηση σε κάθε μία από τις φασματικές περιοχές.

Ας υποθέσουμε ότι θέλουμε ένα συγκεκριμένο μέρισμα να εκτελέσει μία πλήρη περιστροφή στον οκταφωνικό χώρο.

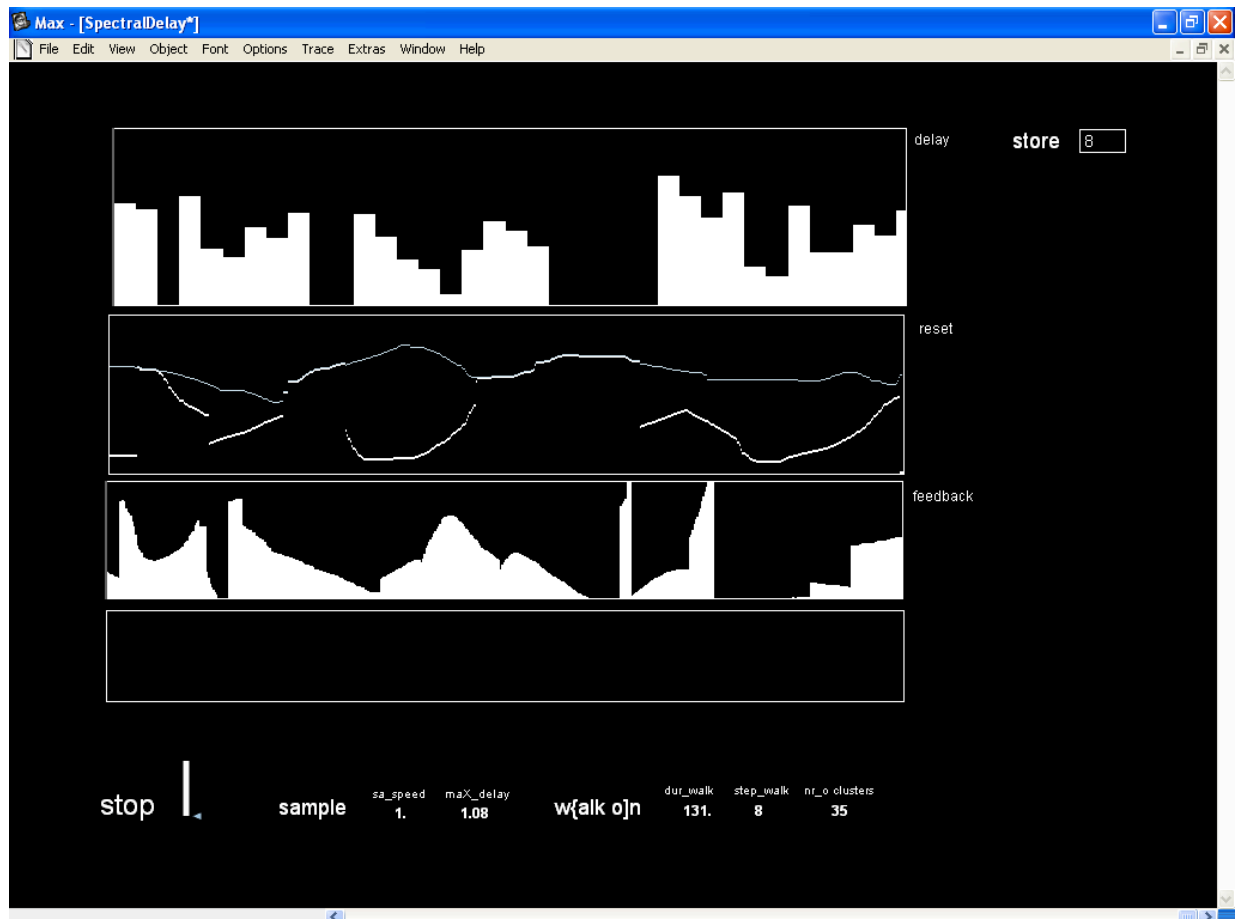
Από την ανάλυση/ανασύνθεση κάθε φορά που εμφανίζεται το επιθυμητό μέρισμα σε κάθε παράθυρο θα πολλαπλασιάζεται η τρέχουσα έντασή του με διακριτές τιμές που προέρχονται από τις ημιτονοειδείς μεταβάσεις ramping που θα συντελούνται μεταξύ των γειτονικών καναλιών με διαφορά φάσης 45° μέχρι να ολοκληρωθεί ο κύκλος και των 8 καναλιών διαδοχικά. Η διαδικασία αυτή αναλύθηκε στο προηγούμενο κεφάλαιο με τη περιστροφή ενός σήματος στον οκταφωνικό χώρο.

Η διαδικασία παραμένει η ίδια αλλά επειδή η κίνηση εξαρτάται από τις τιμές που δίνει το random walk στο κάθε μέρισμα θα έχει ως αποτέλεσμα το κάθε μέρισμα να εκτελεί μία random walk κίνηση στον οκταφωνικό χώρο.

4. Ανάλυση των παραμέτρων

Παρακάτω θα αναλύσουμε τις παραμέτρους που εμφανίζονται στην εικονική διεπαφή της κάθε εφαρμογής ξεχωριστά.

4.1 spectral delay



Σχήμα 10 η εικονική διεπαφή χρήστη της εφαρμογής spectral delay

Αρχικά θα επιχειρηθεί μία σύντομη σύνδεση με την προηγούμενη θεωρητική ανάλυση που αφορούσε τον αλγόριθμο spectral delay και την επεξήγηση των παραμέτρων που υπάρχουν στην εφαρμογή.

Όταν ανοίγει η εφαρμογή βλέπουμε τρία αντικείμενα multislidars.

Το τέταρτο πλαίσιο στο κάτω μέρος της διεπαφής είναι ένας φασματογράφος κατάλληλος για την επίβλεψη του φάσματος που παράγεται.

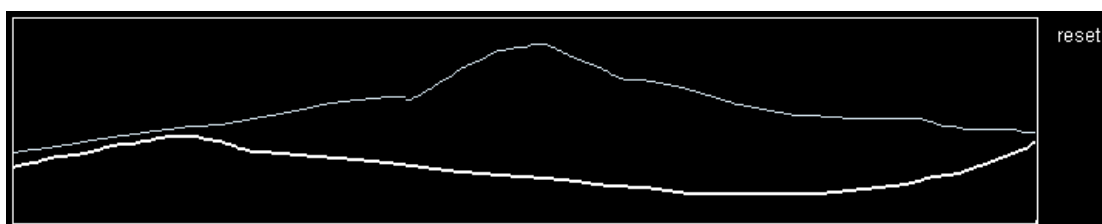
Κάθε ένας από τους τρεις multislidars περιέχει 512 κάθετους sliders των οποίων τη τιμή μπορούμε να επηρεάσουμε με μία απλή χειρονομία του ποντικιού με την προϋπόθεση ότι αυτή θα βρίσκεται εντός των ορίων του multislider που εμείς θέλουμε να επηρεάσουμε.

Οι 512 κάθετοι sliders είναι ισάριθμοι με των αριθμό των μερισμάτων που επηρεάζει ο κάθε slider.



Σχήμα 11 το αντικείμενο ελέγχου των τιμών καθυστέρησης των μερισμάτων

Κατά αυτό τον τρόπο το πρώτο από πάνω αντικείμενο multislider περιέχει 512 κάθετους sliders-κάθετες στήλες που επηρεάζουν την τιμή delay που θα έχει το κάθε μέρισμα ξεχωριστά.



Σχήμα 12 Οι δύο σειρές ορίων που ελέγχουν τη κίνηση random walk του κάθε μερίσματος ξεχωριστά

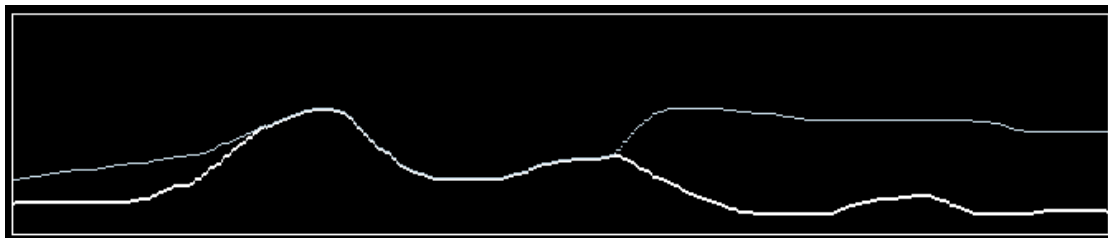
Το δεύτερο αντικείμενο multislider περιέχει δύο 2 σειρές που η κάθε μια αποτελείται από 512 sliders . Ο κάθε slider της πρώτης από πάνω σειράς (με χρώμα γαλάζιο), επηρεάζει το πάνω όριο του random walk του αντίστοιχου μερίσματος που ελέγχει. Ο κάθε slider της δεύτερης σειράς (με χρώμα λευκό) επηρεάζει το κάτω όριο.

Η συγκεκριμένες τιμές έχουν επίδραση μόνο εφόσον έχουμε ενεργοποιήσει το random walk από τον αντίστοιχο διακόπτη (walk on).

Όταν η άσπρη γραμμή πάει να ξεπεράσει τη γαλάζια τότε η τελευταία την ακολουθεί καθορίζοντας ουσιαστικά τα μέγιστα σημεία από τα οποία πέρασε.

Εάν θέλουμε να κατεβάσουμε τη γαλάζια γραμμή ποιο χαμηλά θα πρέπει να χρησιμοποιήσουμε τη παράμετρο `reset` δίπλα στο `multislider` ώστε η άσπρη γραμμή να γίνει ίση με τη γαλάζια.

Όταν το πάνω με το κάτω όριο slider είναι ίσα (επομένως τα μπλε με τα λευκά σημεία θα εφάπτονται) τότε οι sliders ακινητοποιούνται. Δημιουργώντας κατά τόπους περιοχές που τα όρια θα είναι άνισα ή ίσα, ελέγχουμε ποιες περιοχές θα έχουν κίνηση `random walk` και ποιες όχι αντίστοιχα.



Σχήμα 13 οι δύο σειρές ορίων ελέγχουν την διέλευση της κίνησης

Δηλαδή στο παραπάνω παράδειγμα, όλοι οι sliders που ελέγχουν τη καθυστέρηση των μερισμάτων της μεσαίας περιοχής θα είναι ακινητοποιημένα.



Σχήμα 14 οι τιμές ανατροφοδότησης των μερισμάτων

Το τρίτο από πάνω πλαίσιο περιέχει 512 sliders που επηρεάζουν τη τιμή `feedback` του κάθε μερισματος ξεχωριστά.

Οι υπόλοιπες παράμετροι σχετίζονται με τη λειτουργία των αντικειμένων `multislider` καθώς και με την αναπαραγωγή του ηχητικού δείγματος.



Σχήμα 15 οι παράμετροι ελέγχου

Με την επιλογή Start/stop θέτουμε σε λειτουργία/σταματάμε την εφαρμογή.

Ο slider επηρεάζει την ένταση του ήχου.

Με την επιλογή «sample» διαλέγουμε το δείγμα ήχου.

Με το «dry/wet» ελέγχεται η αναλογία του πρωτότυπου (dry) και του επεξεργασμένου (wet) σήματος. Σε τιμή 0 ακούμε μόνο το πρωτότυπο σήμα ενώ σε τιμή 1 μόνο το επεξεργασμένο.

Με την επιλογή «max delay» ορίζεται η μέγιστη καθυστέρηση που μπορεί να έχει ένα μέρισμα.

Δηλαδή εάν στο πρώτο multislider έχουμε επιλέξει ένα κάθετο slider να έχει τη μέγιστη τιμή του(οπότε θα υπάρχει μία κάθετη λευκή γραμμή που θα εφάπτεται και στις δύο οριζόντιες πλευρές του πρώτου multislider)

Δίπλα βρίσκονται οι παράμετροι του random walk



Σχήμα 16 οι παράμετροι ελέγχου του random walk

Με την επιλογή walk-off walk-on απενεργοποιούμε και ενεργοποιούμε το random walk αντίστοιχα.

Αυτό σημαίνει ότι έχοντας ενεργοποιημένη την παράμετρο random walk (walk on) τα sliders που ελέγχουν τη παράμετρο delay των 512 μερισμάτων (πρώτο multislider) θα κινούνται αυτόματα, το καθένα με τη δική του κίνηση random walk. Όλα όμως μαζί τα sliders θα ελέγχονται ομαδικά από τρεις παραμέτρους.

Η πρώτη παράμετρος (dur_walk) καθορίζει το πόσο γρήγορα θα κινούνται οι sliders σε μονάδα msec. Εάν όπως στο παράδειγμα είναι ρυθμισμένη στη τιμή 20 τότε ανά 20 msec θα δίνεται μία νέα τιμή random walk σε όλους τους sliders «ταυτόχρονα».

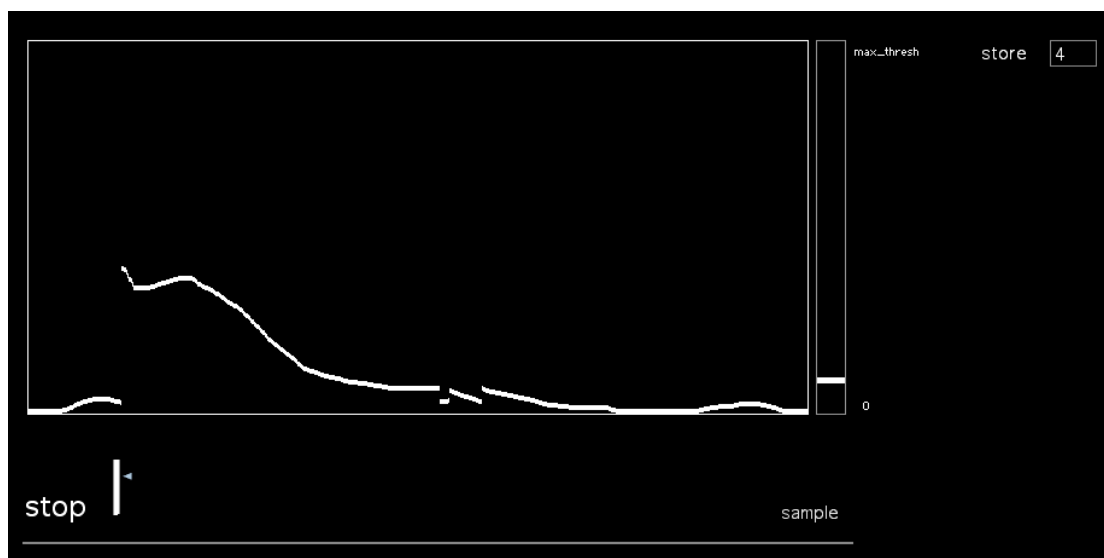
Η δεύτερη παράμετρος step_walk καθορίζει πόσο μεγάλο θα είναι το βήμα του random walk που θα εκτελεί ο κάθε slider ξεχωριστά.

Το slider μπορεί να πάρει 127 τιμές ξεχωριστά, άρα εάν έχουμε ρυθμίσει το step walk στη τιμή 8, μπορεί ο κάθε slider εάν ο μηχανισμός random walk του δώσει την ίδια κατεύθυνση για 16 φορές, να φτάσει από τη κατώτατη τιμή (το 0) στην ανώτατη (το 127).

Η τρίτη παράμετρος nr_clusters καθορίζει πόσες ομάδες (clusters) θα σχηματίσουν οι 512 sliders.

Δηλαδή αν επιλέξουμε 32 ομάδες τότε ανά 16 διαδοχικά μερίσματα οι 512 sliders θα σχηματίσουν 32 ομάδες. Αυτό σημαίνει ότι οι sliders 1-16 θα εκτελούν ταυτόχρονα την ίδια κίνηση random walk το ίδιο και οι sliders 177-192.

4.2 Spectral gate



Σχήμα 17 Η εικονική διεπαφή χρήστη της εφαρμογής spectral gate

Η εφαρμογή spectral gate ελέγχεται ουσιαστικά από δύο αντικείμενα, ένα multislider και ένα slider, όπου και τα δύο ρυθμίζουν τη τιμή gate του κάθε μερίσματος.

Το αντικείμενο multislider στέλνει 512 ξεχωριστές τιμές που αντιστοιχούν στις 512 τιμές ορίου του gate που θα αντιστοιχηθούν στα 512 μερίσματα που περιέχει κάθε παράθυρο.

Οι πρώτοι επομένως sliders ελέγχουν τη τιμή ορίου του gate που αντιστοιχεί στα πρώτα μερίσματα (χαμηλές συχνότητες) και οι τελευταίοι sliders στα τελευταία μερίσματα (ψηλές συχνότητες).

Η χαμηλότερη τιμή των sliders είναι το 0. ενώ η μέγιστη ρυθμίζεται από το slider δίπλα στο multislider, ο οποίος έχει ως κατώτατη τιμή το 0. και ως ανώτατη το μέγιστο όριο (max thresh) της δυναμικής του μερίσματος.

Επομένως η μέγιστη τιμή που μπορεί να δεχτεί ο κάθε slider που ανήκει στο multislider και κατ' επέκταση η παράμετρος gate του κάθε μερίσματος ξεχωριστά, είναι εξαρτώμενη από το slider που ελέγχει τη τιμή του μέγιστου ορίου του [gate].

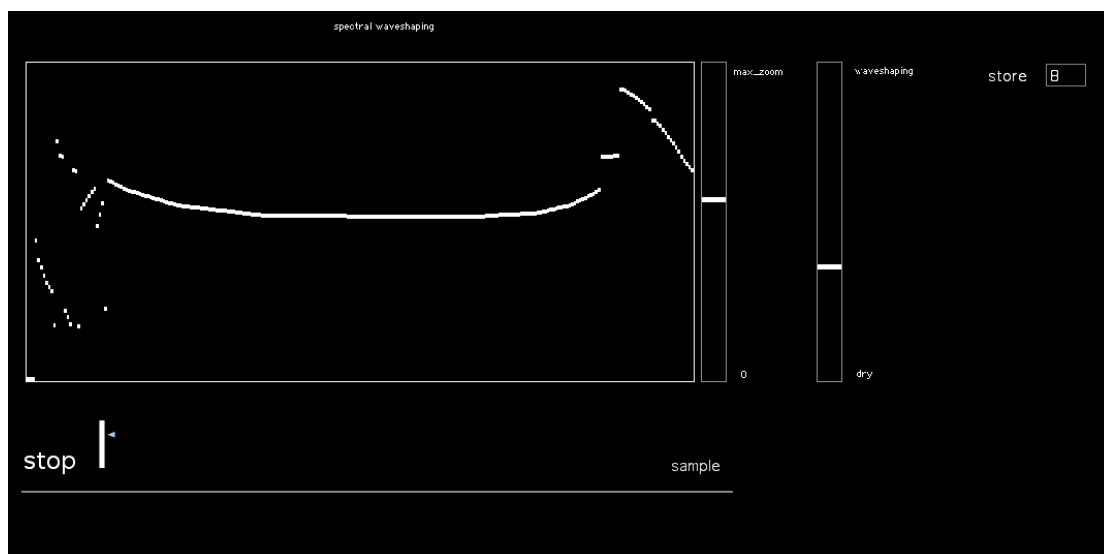
Όταν η τιμή του slider είναι 0., τότε δεν αποκόπτεται καμία δυναμική από το σήμα ενώ όταν πάρει τη μέγιστη τιμή (max

thresh) ο multislider έχει τη δυνατότητα να αποκόψει από το σήμα οποιαδήποτε δυναμική μέσω του gate.

Η συγκεκριμένη επιλογή, της προσθήκης ελέγχου του μέγιστου ορίου που θα αποκόπτεται από το σήμα, κρίθηκε επιβεβλημένη καθώς ύστερα από ανάλυση των δυναμικών των μερισμάτων, καταλήξαμε στο συμπέρασμα ότι όλη η ενέργεια των δυναμικών και κατ' επέκταση ο κύριος όγκος του σήματος είναι συγκεντρωμένος πολύ χαμηλές τιμές. Κατόπιν δοκιμών διαπιστώσαμε ότι μία τιμή του μονού slider της τάξης $1/20$ έως $1/10$ της τιμής της μέγιστης δυναμικής του μερίσματος, ενδέχεται να δώσει πολύ καλύτερο έλεγχο σε σχέση με αυτόν που θα είχαμε εάν ελέγχαμε όλο το δυναμικό εύρος του σήματος.

Η μέγιστη τιμή που θα δώσουμε εξαρτάται πάντα και από τα φασματικά χαρακτηριστικά του ίδιου του σήματος.

4.3 spectral waveshaping



Σχήμα 18 Η εικονική διεπαφή χρήστη της εφαρμογής spectral waveshaping

Το περιβάλλον ελέγχου της εφαρμογής spectral waveshaping είναι παρόμοιο με το περιβάλλον ελέγχου της εφαρμογής spectral gate.

Ένα αντικείμενο multislider δημιουργεί γραφικά τη συνάρτηση μετασχηματισμού καταχωρώντας τις τιμές σε μία μνήμη 250 θέσεων, που έχει τη δυνατότητα υπολογισμού κάθε ενδιάμεσης θέσης, με τη μέθοδο του σταθμισμένου μέσου (interpolation).

Ένας μονός slider μας επιτρέπει να διαχωρίσουμε τη δυναμική περιοχή του κάθε μερίσματος σε δύο περιοχές μέσω ενός ορίου (thresh).

Σε τιμές από 0. έως thresh, οι δυναμικές του μερίσματος θα υπόκεινται το μετασχηματισμό πλάτους μέσω της συνάρτησης μετασχηματισμού.

Από thresh μέχρι το μέγιστο πλάτος που δίνει η ανάλυση μέσω fft, το σήμα θα βγει στην έξοδο χωρίς μετασχηματισμό.

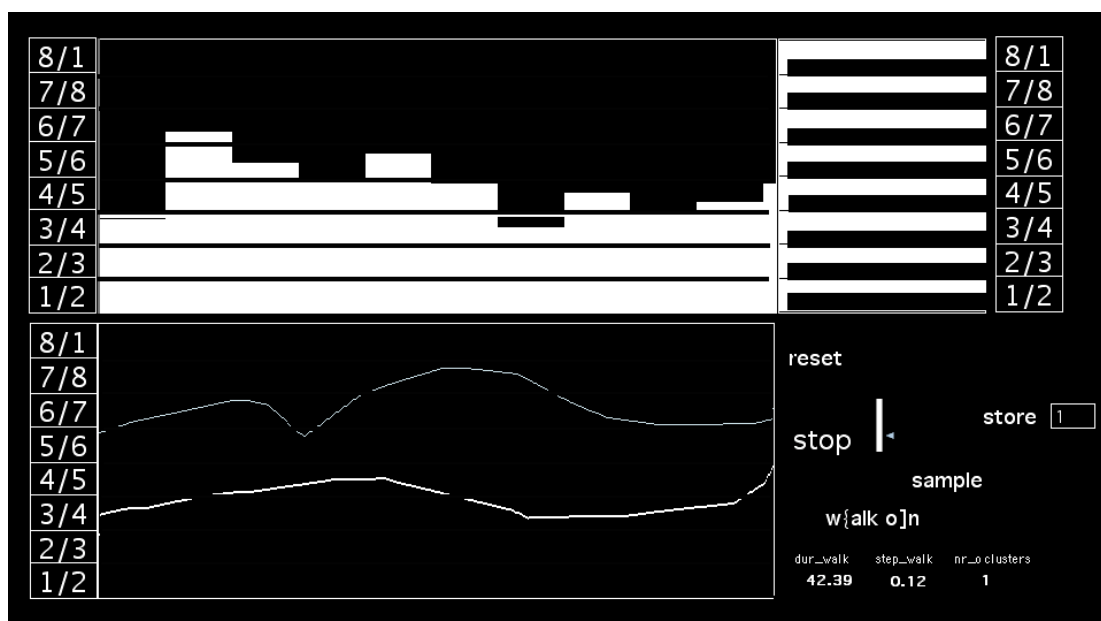
Η συγκεκριμένη παράμετρος κρίθηκε απαραίτητη λόγω της ανομοιόμορφης κατανομής των δυναμικών στα μερίσματα όπως ήδη αναφέραμε στην εφαρμογή spectral gate. Κατ' αυτό το τρόπο ο μετασχηματισμός των πολύ χαμηλών τιμών δυναμικών, επιφέρει μία δραματική αύξηση της συνολικής έντασης του σήματος. Ένας επιπλέον λόγος που ενισχύεται το φαινόμενο, είναι ότι στις πολύ χαμηλές δυναμικές κρύβονται συχνά ανεπιθύμητοι θόρυβοι (πχ θόρυβος βάθους), οι οποίοι μετά το μετασχηματισμό της τιμής

τους, αναδύονται στην επιφάνεια του προσλαμβανομένου σήματος.

Για τον ίδιο λόγο η τιμή thresh αυξάνει εκθετικά ώστε να υπάρχει μεγαλύτερη ανάλυση στις χαμηλότερες τιμές.

Τέλος ο δεύτερος μονός slider αλλάζει το ποσοστό έντασης μεταξύ των δύο περιοχών. Όταν η μία περιοχή από τις δύο έχει ένταση [$\alpha * amp$], η δεύτερη θα έχει ένταση [$(1-\alpha)*amp$], όπου α η τιμή που δίνει ο slider στη κλίμακα 0 έως 1, και amp η τρέχουσα δυναμική του σήματος.

4.4 spectral spatialization



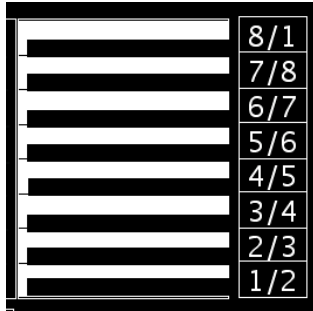
Σχήμα 19 Η εικονική διεπαφή χρήστη της εφαρμογής spectral spatialization

Ανοίγοντας την εφαρμογή spectral spatialization, αντικρίζουμε τις ίδιες παραμέτρους που συναντήσαμε και στην εφαρμογή spectral delay και οι οποίες αφορούν το random walk. Βλέπουμε ένα αντικείμενο multislider στο οποίο εκτελείται το random walk όπου τα όρια του κάθε μερίσματος ξεχωριστά ελέγχονται από ένα multislider δύο περιοχών. Επίσης συναντάμε ξανά τις παραμέτρους random walk on/off, dur walk, step walk, nr of clusters όπως ακριβώς έχουμε δει και αναλύσει στην εφαρμογή spectral delay. Το random walk όμως στη προκειμένη περίπτωση και γενικότερα ο έλεγχος του multislider μίας περιοχής, εφαρμόζεται στην τοποθέτηση των μερισμάτων στον οκταφωνικό χώρο. Οι αριθμοί δίπλα στους multislider και οι διαχωριστικές γραμμές εξυπηρετούν αυτόν ακριβώς το σκοπό, το να κατατοπίζουν το χρήστη σε ποια περιοχή του οκταφωνικού χώρου βρίσκεται. Όταν για παράδειγμα βρίσκεται μέσα στις διαχωριστικές γραμμές που οριοθετούνται από τους αριθμούς 5/6, σημαίνει ότι η κάτω γραμμή της περιοχής αντιπροσωπεύει το 5⁰ κανάλι και η πάνω γραμμή το 6⁰.



Σχήμα 20 περιοχή ελέγχου panning

Ταυτόχρονα δίπλα στο multislider μίας περιοχής αντικρίζουμε 8 άσπρες γραμμές που ρυθμίζουν τις εντάσεις των αντίστοιχων καναλιών.



Σχήμα 21 οι εντάσεις των 8 καναλιών

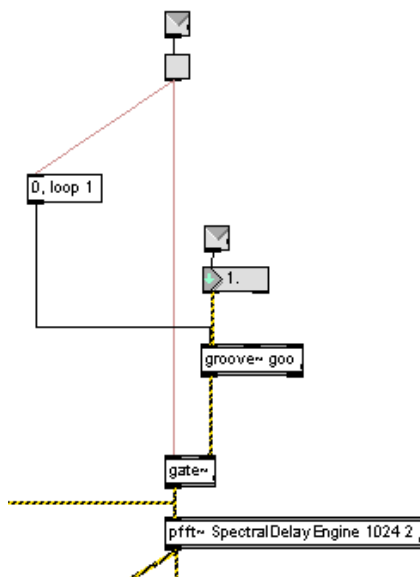
5. Ανάλυση των αλγορίθμων

Και οι 4 εφαρμογές κατασκευάστηκαν σε περιβάλλον max/msp. Το μέγεθος παραθύρου που χρησιμοποιείται έχει μέγεθος 1024. Το παράθυρο είναι τύπου `hanning`. Παρακάτω θα αναλύσουμε το τρόπο κατασκευής των αλγορίθμων των εφαρμογών.

Το αντικείμενο που επιλέχθηκε για να αναπαράγει το δείγμα του ήχου προς διαμόρφωση είναι το `[groove]` το οποίο μέσω ενός αντικειμένου `[gate]` υπόκειται σε έλεγχο εάν θα αναπαράγει τον ήχο ή όχι. Ως βασικά ορίσματα δέχεται:

A) ένα αριθμό ο οποίος επιλέγει σε msec από ποιο χρονικό σημείο θα ξεκινήσει να αναπαράγεται το ηχητικό δείγμα και στον οποίο έχουμε θέσει τη τιμή 0 (δηλαδή από την αρχή του),

B) τη ταχύτητα αναπαραγωγής του δείγματος η οποία έχει καθοριστεί μέσω του αντικειμένου `[sig~ 1]`, να είναι ίση με το πρωτότυπο δείγμα.

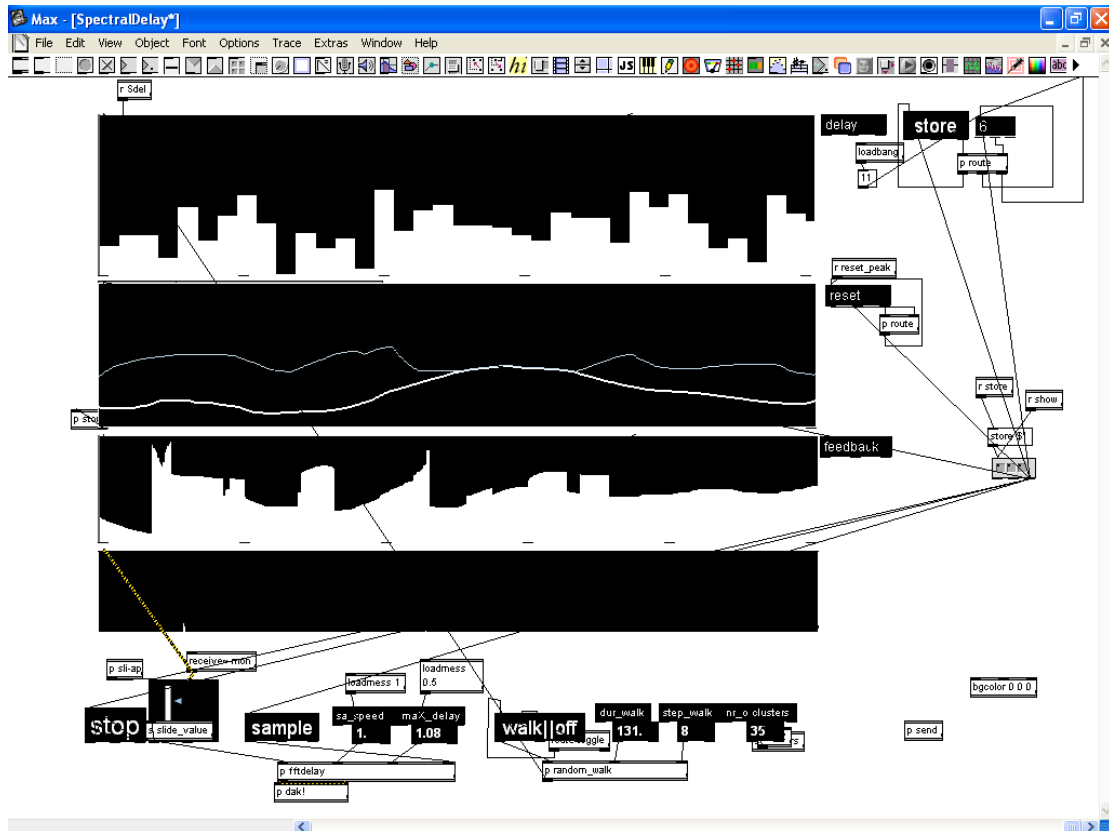


Σχήμα 22 το sampler

Μέσω ενός αντικειμένου `[gate~]` το οποίο ελέγχεται εξωτερικά από την επιλογή `start/stor`, επιλέγουμε την διέλευση ή όχι του σήματος.

5.1. spectral delay

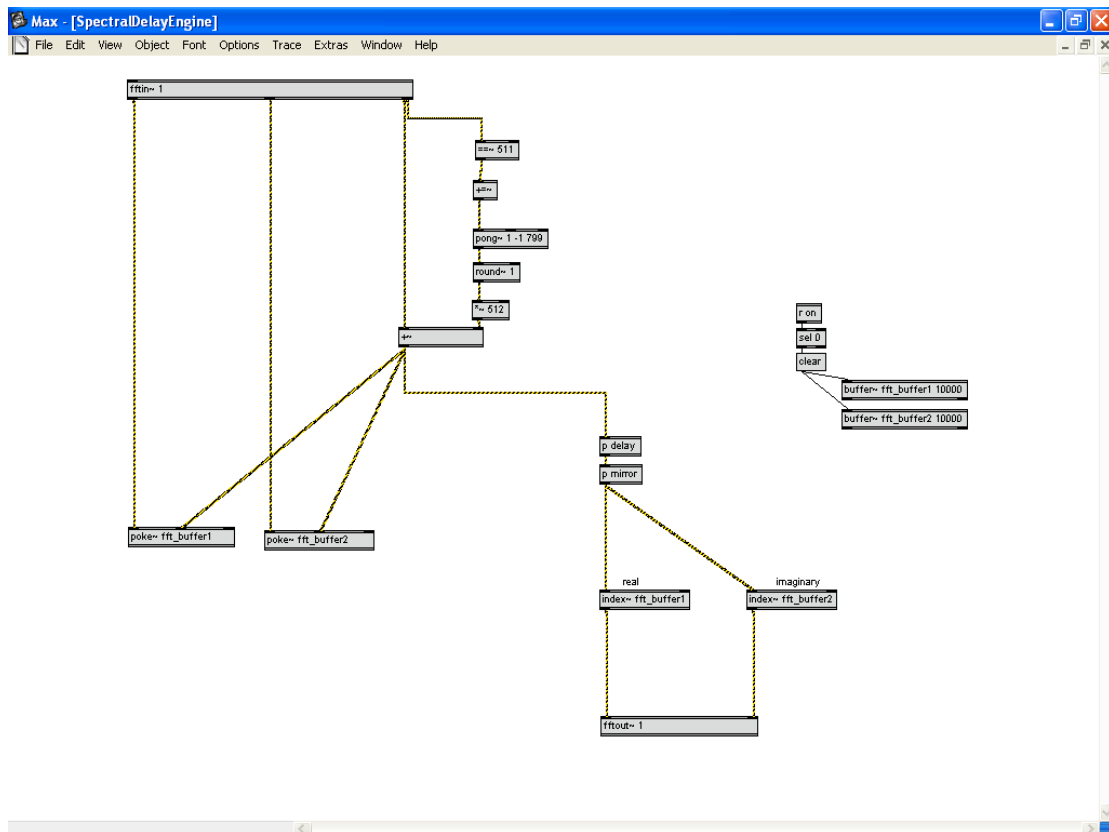
5.1.1 Η δημιουργία καθυστέρησης στα μερίσματα



Σχήμα 23 το κεντρικό patch της εφαρμογής spectral delay «ξεκλειδωμένο»

Ο αλγόριθμος βρίσκεται μέσα στο patch που διαβάζει το [rfft] με την ονομασία spectral delay engine το οποίο είναι ξεχωριστό.

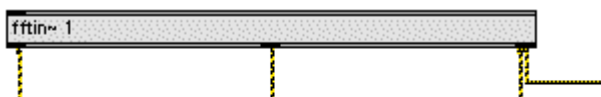
Αρχικά θέλουμε να επιτύχουμε την αποθήκευση των πραγματικών και των φανταστικών μερών του κάθε μερίσματος, ενός πεπερασμένου αριθμού παραθύρων, σε μία μνήμη (Boyle 2004)



Σχήμα 24 ο αλγόριθμος δημιουργίας καθυστέρησης στα μερίσματα

Το αντικείμενο [rfft] αναλύει το σήμα με μέγεθος παραθύρου τον αριθμό του πρώτου ορίσματος (1024) και overlap τον αριθμό του δεύτερου ορίσματος (2).

Το σήμα αναλύεται σε 512 φασματικές περιοχές και το αντικείμενο [fftin] μας δίνει τα αποτελέσματα της ανάλυσης σε πραγματικό χρόνο.



Σχήμα 25 το αντικείμενο fftin~

Από τη πρώτη αριστερά έξοδο του αντικειμένου βγαίνει το πραγματικό μέρος και από δεύτερη το φανταστικό.

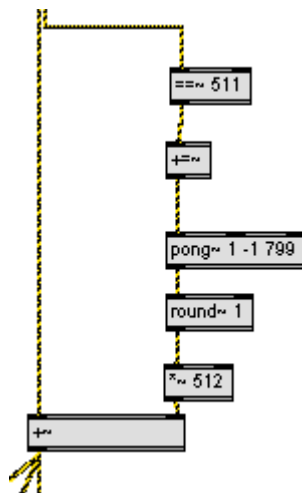
Από τη τρίτη βγαίνει ο αριθμός του μερίσματος που αντιστοιχεί στη κάθε χρονική στιγμή στο μιγαδικό.

Δεδομένου ότι έχουμε ανάλυση με μέγεθος παραθύρου 1024 κατα συνέπεια η συχνότητα του κάθε μερίσματος θα καθορίζεται από τη τιμή:

$44100/1024$ * αύξοντα αριθμό μερίσματος = 43.066 *αύξοντα αριθμό μέρισμα. πχ το 57 μέρισμα θα έχει συχνότητα $43.066*57=2454.762\text{HZ}$.

Όπου 44100 είναι η συχνότητα δειγματοληψίας.

Άρα από τη δεξιά έξοδο έχουμε ουσιαστικά τη συχνότητα.



Σχήμα 26 η δημιουργία δείκτη ανάγνωσης παραθύρων

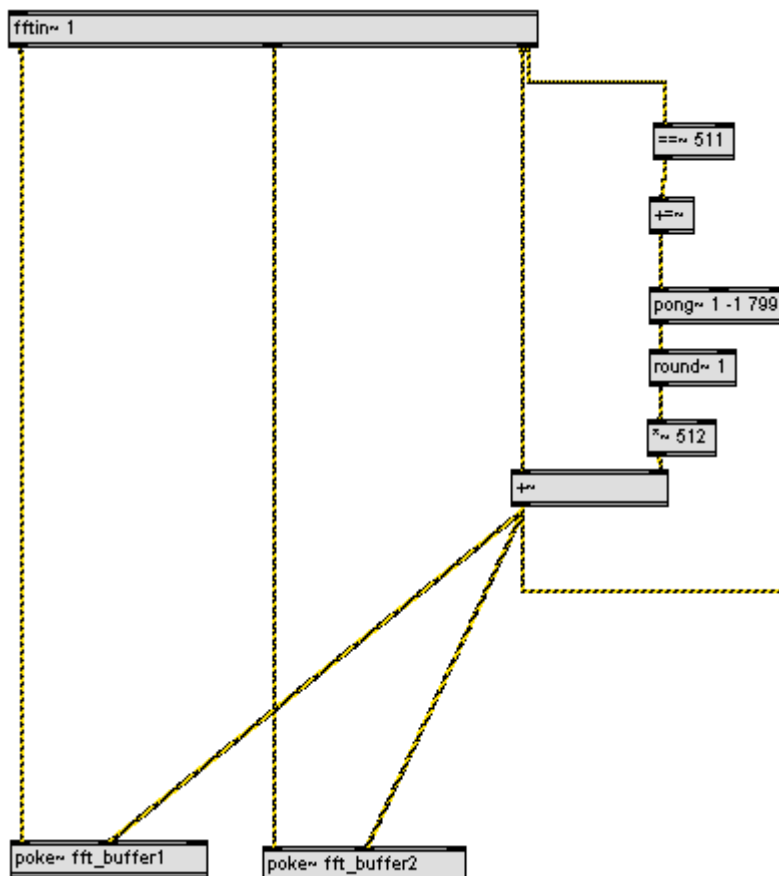
Όταν η τιμή του δείκτη ανάγνωσης ισούται με 511, μία λογική συνθήκη εκτελείται μέσω του αντικειμένου `[==~ 511]`, η οποία στέλνει τη τιμή 1 όταν έρχεται στην είσοδο του συγκριτή τιμή ίση με το 511 και 0 όταν έρχεται οποιαδήποτε άλλη.

Αρχικά θέλαμε η συνθήκη να δίνει 1 όποτε έρχεται καινούριο παράθυρο δηλαδή όποτε στέλνεται από τη δεξιά έξοδο του `[fftin~]` η τιμή 0 ως αύξων αριθμός του τρέχοντος μερίσματος. Επειδή όμως παρουσιάστηκαν μη αναμενόμενες ηχητικές αλλοιώσεις αντικαταστήσαμε τη τιμή 0 με τη τιμή 511 κατόπιν δοκιμών.

Ένας accumulator κάθε φορά που αλλάζει το παράθυρο αθροίζει το σήμα εισόδου (που έχει τιμή 1) λειτουργώντας ως αύξοντας δείκτης του τρέχοντος παραθύρου.

Το αντικείμενο [pong~] οριοθετεί τη μέγιστη τιμή του δείκτη του παραθύρου. Όταν ο δείκτης ξεπερνάει τη τιμή 799, επιστρέφει αυτόματα στη τιμή 0.

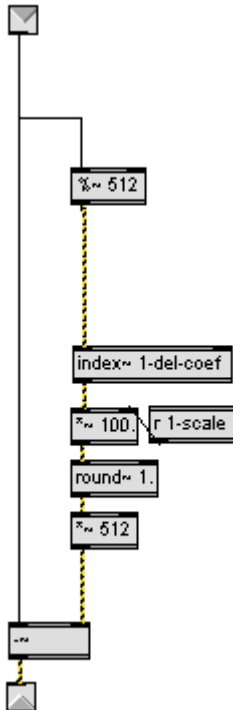
Το αντικείμενο [round~] χρησιμοποιείται στην έξοδο του αντικείμενου [pong~] καθώς διαπιστώσαμε ότι σε μεγάλες τιμές προστίθεται ένα μη αναμενόμενο δεκαδικό μέρος στις ακέραιες τιμές του δείκτη ανάγνωσης, που οδηγεί σε ανεπιθύμητες ηχητικές αλλοιώσεις του σήματος.



Σχήμα 27 καταχώρηση των πραγματικών και των φανταστικών τιμών των μερισμάτων σε μνήμες

Το πραγματικό μέρος αποθηκεύεται μέσω του αντικειμένου [poke~] στο fft_buffer1 και το φανταστικό στο fft_buffer2.

Το αντικείμενο [poke~] αποθηκεύει τη τιμή του πραγματικού ή φανταστικού μέρους στη πρώτη αριστερά είσοδο και σε θέση που υποδεικνύει το σήμα που έρχεται στην μεσαία είσοδο.



Σχήμα 27 η δημιουργία του αναδρομικού δείκτη ανάγνωσης

Μέσα στο sub-patcher με το όνομα delay βρίσκεται ο αλγόριθμος που επιλέγει ποιες προγενέστερα αποθηκευμένες τιμές θα αναπαραχθούν στην έξοδο.

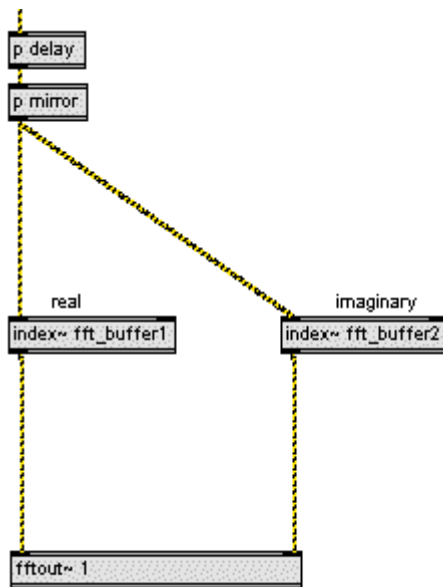
Το index μετατρέπεται πάλι σε αύξοντα αριθμό μερίσματος μέσω της συνάρτησης [%~].

Κατόπιν επιλέγει τον αύξοντα αριθμό του παραθύρου που θα διαβαστεί σε καθυστέρηση και το οποίο έχει καθορισθεί δυναμικά από τον χρήστη στο πρώτο πλαίσιο και έχει αποθηκευτεί στο [buffer~ 1-del-coef].

Κατόπιν η τιμή στρογγυλοποιείται και πολλαπλασιάζεται με το 512 ώστε να δώσει ένα ακέραιο αριθμό που θα αντιστοιχεί στον αριθμό των παραθύρων που θα αφαιρεθούν από τη τρέχουσα τιμή του index για να βρεθεί ποιο προγενέστερο μέρισμα θα χρησιμοποιηθεί κατά την ανασύνθεση του τρέχοντος παραθύρου (δημιουργείται μία ομάδα από 512 καινούρια μερίσματα).

Η τιμή αυτή του index, που αντιστοιχεί στον αύξοντα αριθμό του μερίσματος, θα περάσει από το subpatcher με το όνομα mirror. Σε αυτό εάν ξεπεραστεί ο αριθμός 409600 που αποτελεί τη μέγιστη τιμή sample που ο buffer καταχωρεί το πραγματικό και το

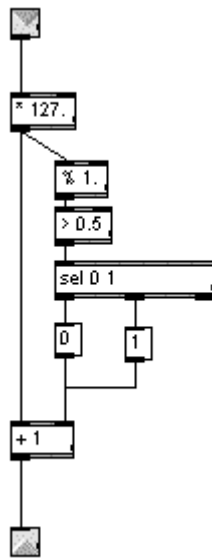
φανταστικό μέρος (800 παράθυρα * 1024/2 μέγεθος παραθύρου/2 = 409600) ξανααγυρνάει πάλι τη μέτρηση από το 0.



Σχήμα 28 Η ανασύνθεση των τιμών

Ο τελικός δείκτης ανάγνωσης επιλέγει την πραγματική και τη φανταστική τιμή από τα αντικείμενα [index~fft_buffer1] και [index~fft_buffer2] και κατόπιν οδηγείται στο αντικείμενο [fftout~ 1] μαζί με τις υπόλοιπες πραγματικές/φανταστικές τιμές των μερισμάτων που έχουν προκύψει με τον ίδιο τρόπο, για να ανασυνθέσουν το νέο παράθυρο.

Οι υπόλοιποι αλγόριθμοι που υπάρχουν μέσα στην εφαρμογή υποστηρίζουν τον κεντρικό αλγόριθμο που αναλύθηκε παραπάνω.



Το sub-patcher round μετατρέπει το εύρος τιμών 0 -1 σε εύρος τιμών 0 - 127

και είναι επιφορτισμένο με την εργασία να στρογγυλοποιεί τις δεκαδικές τιμές

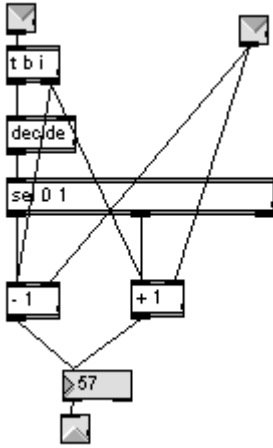
προς το πλησιέστερο μεγαλύτερο ακέραιο αριθμό σε περίπτωση που το δεκαδικό μέρος είναι μεγαλύτερο από 0.5

και προς το μικρότερο εάν είναι μικρότερο από 0.5.

Σχήμα 31 η στρογγυλοποίηση των τιμών

Μετάπειτα δημιουργείται με βάση το δείκτη ανάγνωσης λίστα δύο στοιχείων

[index, τιμή καθυστέρησης] μέσω το αντικειμένου [pack]

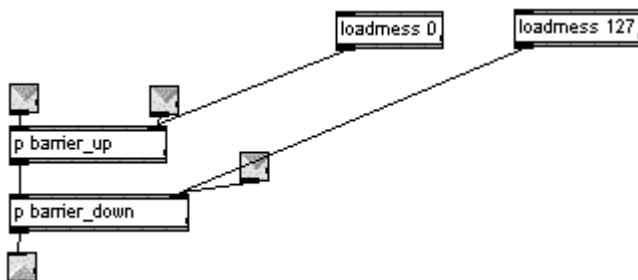


Σχήμα 33 Η επιλογή του τυχαίου βήματος

Κάθε φορά που περνάει στο subpatcher random walk, ένας αριθμός επιλέγει από ένα άλλο τυχαίο αριθμό μεταξύ 0 ή 1. Εάν είναι 0 τότε ο αριθμός που ενεργοποίησε τη διαδικασία επιλογής προχωράει στη πρόσθεση με το -1 βήμα random walk. Εάν είναι 1 τότε προστίθεται με το 1 βήμα random walk.

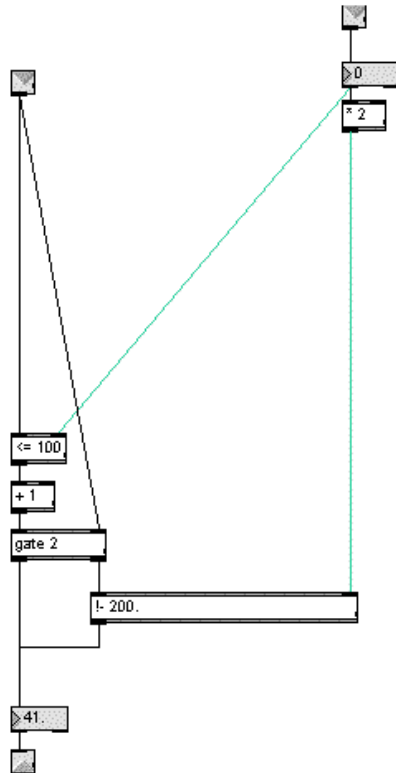
Επομένως μέσω αυτής της διαδικασίας επιλέγεται η επόμενη τιμή καθυστέρησης του μερίσματος που αντιστοιχεί στο συγκεκριμένο δείκτη ανάγνωσης.

Όπως έχουμε πει ο αριθμός αυτός περιορίζεται από δύο barriers. Στη περίπτωση που ξεπεράσει το όριο ανακλάται δηλαδή εκτελεί το βήμα με την αντίθετη διεύθυνση από αυτή που του είχε δώσει το random walk. Η διαδικασία αυτή πραγματοποιείται μέσα στο subpatcher barriers.



Σχήμα 34 τα δυο όρια τιμών του random walk

Όπως παρατηρούμε διέρχεται από δύο barriers που αντιστοιχούν στο πάνω και το κάτω όριο που έχει το random walk σε κάθε ξεχωριστό μέρισμα.

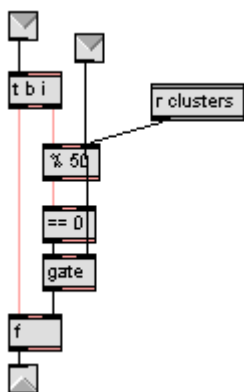


Σχήμα 35 η δημιουργία του ορίου

Από την δεξιά είσοδο έρχεται ο αριθμός που είχε επιλεγεί από το barrier_up ή από το barrier_down αντίστοιχα.

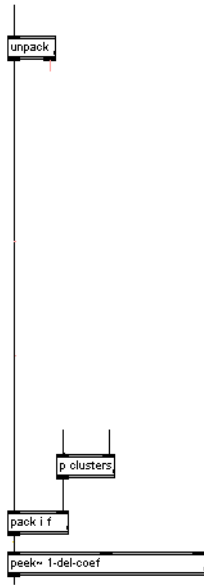
Εάν έχει ξεπεράσει τον αριθμό (μικρότερο στο barrier_down και μεγαλύτερο στο barrier_up) τότε θα περάσει από το συμπλήρωμα του διπλάσιου ορίου. Διαφορετικά θα περάσει ως έχει.

Το τελικό βήμα random walk διέρχεται από το sub-patcher clusters από το οποίο ομαδοποιείται κατά την εκτέλεση του βήματος με άλλα μερίσματα-sliders.



Ο αριθμός του τυχαίου βήματος μπορεί να διέρχεται μέσα από το gate μόνο όταν ο λογικός έλεγχος είναι αληθής δηλαδή μόνο όταν το modulo δίνει 0 όταν διέρχεται από αυτό. Στην αντίθετη περίπτωση εκτελεί το ίδιο βήμα με το slider που έδωσε αποτέλεσμα 0.

Σχήμα 36 η δημιουργία των clusters

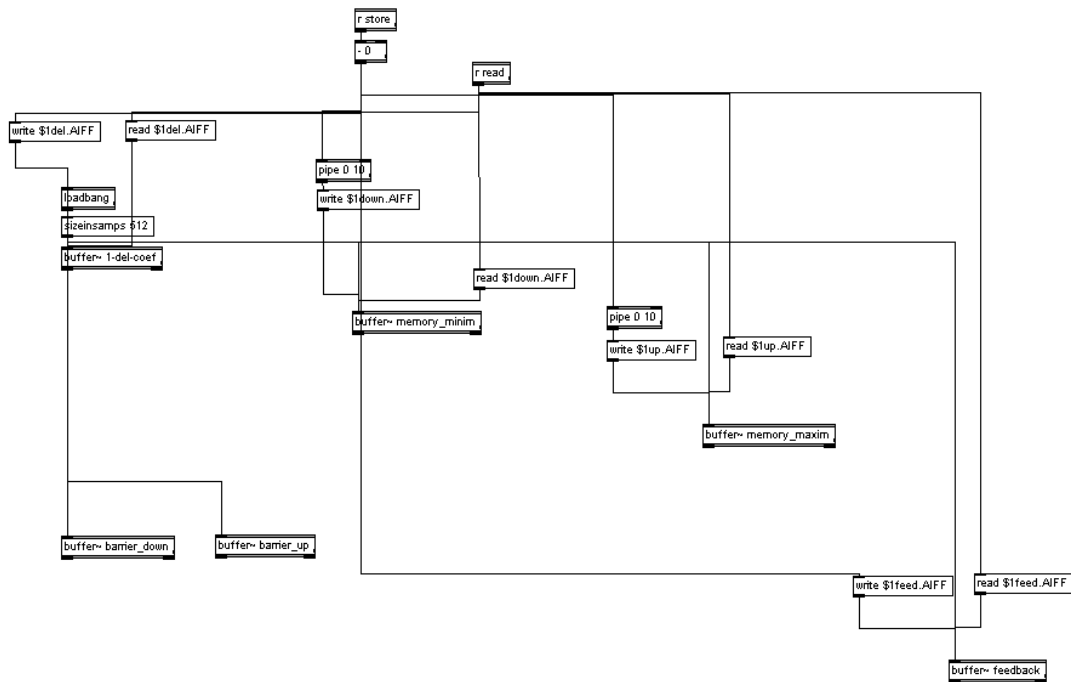


Η νέα τιμή καθυστέρησης
 θα δημιουργήσει μία λίστα δύο στοιχείων μαζί με
 τον αντίστοιχο δείκτη ανάγνωσης
 και εισερχόμενη
 στην αριστερή είσοδο
 της συνάρτησης [peek~]
 θα καταχωρηθεί
 στη θέση με αριθμό
 αυτόν του δείκτη ανάγνωσης και τιμή αυτήν της
 νέας τιμής καθυστέρησης

Σχήμα 36 η καταχώρηση της νέας τιμής random walk

Το σύστημα αποθήκευσης

Οι λίστες που έχουν πάνω από 253 στοιχεία δε μπορούν να αποθηκευτούν στο αντικείμενο preset γι' αυτό μία λύση που βρήκαμε είναι να καταχωρούνται ως aiff αρχεία και να ανακαλούνται από το χρήστη μέσω ενός κεντρικού preset.



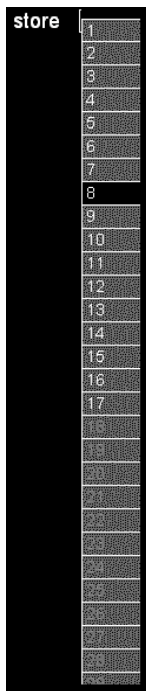
Σχήμα 37 το σύστημα αποθήκευσης

Οι εντολές write και read είναι αυτές που γράφουν σε αρχείο και διαβάζουν από αρχείο.

To preset

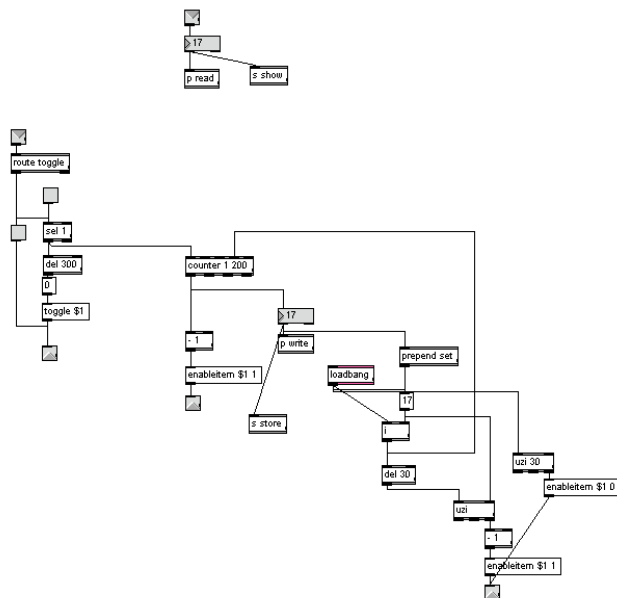
Ο αλγόριθμος έχει κατασκευαστεί με τα εξής χαρακτηριστικά

store 8

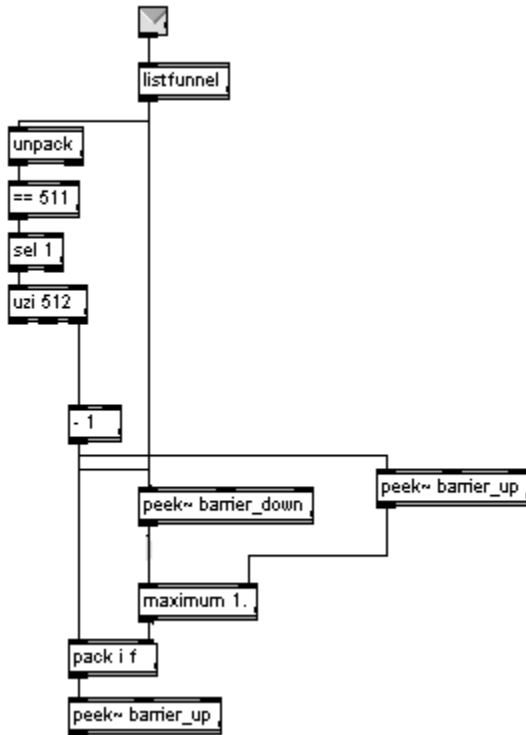


ο χρήστης επιλέγει το επιθυμητό preset από το λευκό ορθογώνιο και με βάση τους αριθμούς που είναι ενεργοποιημένοι.

Όταν πατήσει store θα ενεργοποιηθεί και θα αποθηκευτούν όλοι οι παράμετροι της εφαρμογής στον αμέσως επόμενο αριθμό (πχ στο παράδειγμα το 18)



Σχήμα 38 η εικονική διεπαφή και ο αλγόριθμος δημιουργίας του preset



Σχήμα 40 η καταχώρηση της μέγιστης τιμής του μερίσματος

Όταν θα μεταβάλλουμε τη διπλή σειρά των sliders μία φορά τότε ο barrier_up μέσω του συνδυασμού των αντικειμένων [listfunnel], [peek~] γεμίζει με τις τιμές της κάτω σειράς.

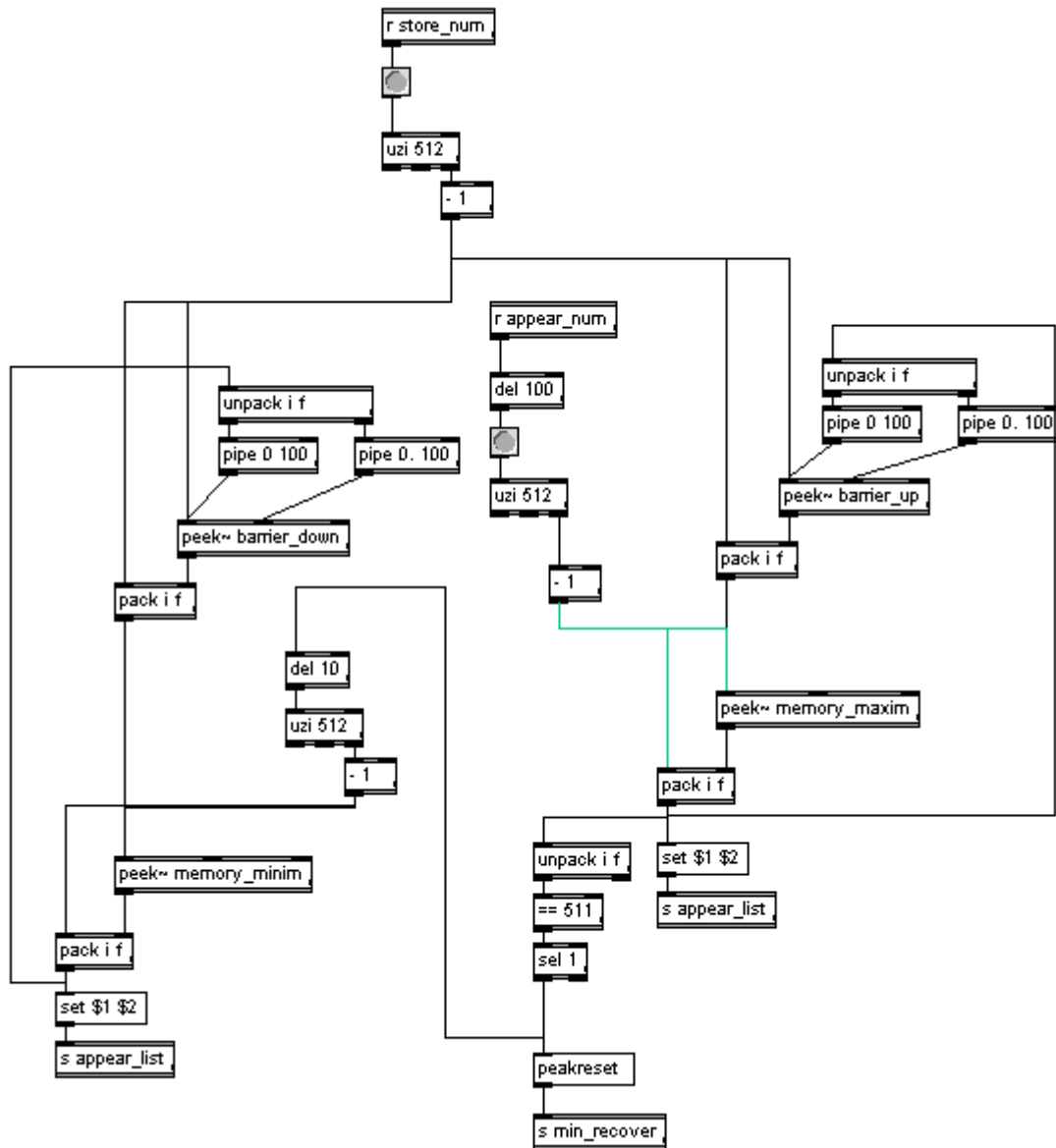
Κατόπιν η συνάρτηση uzi ενεργοποιείται για να βγουν κατά αύξοντα αριθμό οι αριθμοί που θα χρησιμεύσουν ως οι δείκτες ανάγνωσης της συνάρτησης [peek~ barrier_down] εξαναγκάζοντας τις τιμές να εξέλθουν από αυτή και να συγκριθούν με αυτές που έχουν τον ίδιο δείκτη (θέση στη μνήμη).

Το μέγιστο των δύο τιμών καταγράφεται στη μνήμη barrier_up μέσω του δείκτη αφού πρώτα δημιουργείται λίστα δύο στοιχείων.

Επιπλέον αυτές οι τιμές πρέπει να έχουν τη δυνατότητα να ανακαλούνται όποτε ο χρήστης επιθυμεί μέσω αντικειμένων που αποθηκεύουν τις τιμές τους όπως το preset.

Το αντικείμενο preset όμως δε μπορεί να δει μέσα σε μία μνήμη όπως και δε μπορεί να αποθηκεύσει λίστα μεγαλύτερη από 253 στοιχεία (ενώ η εφαρμογή δουλεύει με 512). Επιλέξαμε ως λύση να αποθηκεύσουμε τις τιμές σε aiff αρχεία όπως έχουμε πράξει και με τις υπόλοιπες λίστες.

Όταν ο χρήστης διαλέγει την επιλογή store τότε ενεργοποιείται το αντικείμενο [uzi] το οποίο στέλνει τιμές από 0 έως 511. Οι τιμές αυτές θα χρησιμεύσουν ως δείκτες ανάγνωσης που θα διαβάσουν τις τιμές των barrier_up και barrier_down. Κατόπιν ο δείκτης ανάγνωσης και η τιμή που του αντιστοιχεί στοιχειοθετούνται σε λίστες των δύο στοιχείων (δείκτης ανάγνωσης, τιμή) ώστε να αποθηκευτούν εκ νέου στις μνήμες memory_minim και memory_maxim.



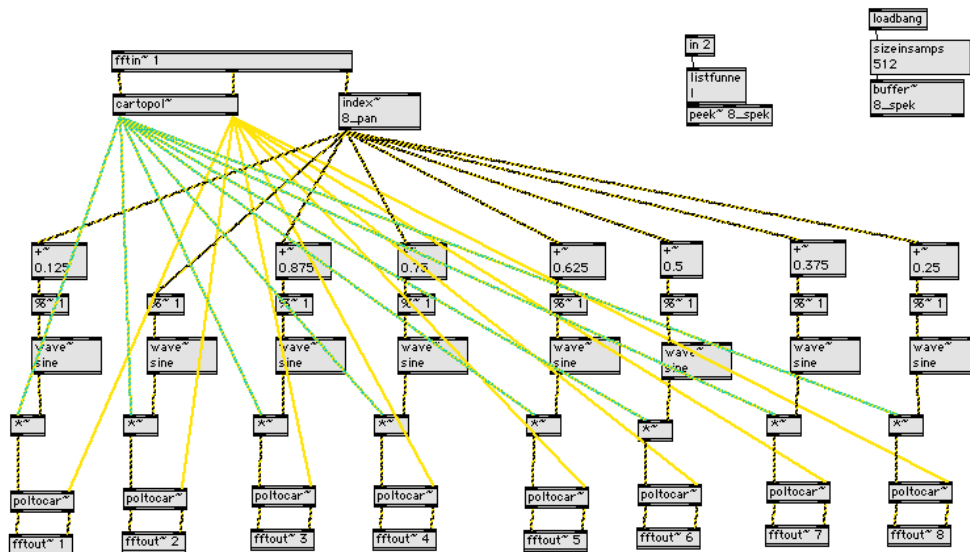
Σχήμα 41 ο αλγόριθμος ανάκλησης των τιμών

Εάν είναι το πλάτος του μερίσματος είναι μεγαλύτερο από το πλάτος που καταχωρούμε από το πρώτο μονό slider

(amp_μερίσματος ϵ (threh, 650)) τότε θα φύγει το πλάτος του μέρισμα της δεύτερης εισόδου. Εάν είναι μικρότερο (amp_μερίσματος ϵ (0, thresh)) τότε θα φύγει το πλάτος του μερίσματος της πρώτης εισόδου.

Πρώτου βγει στην έξοδο του [selector~] θα πολλαπλασιαστεί με ένα συντελεστή που έρχεται από το δεύτερο slider ώστε να καθοριστεί η αναλογία του πλάτους ανάμεσα στις δύο περιοχές για να οδηγηθεί στην έξοδο.

5.4 spectral spatialization



Σχήμα 44 η καταχώρηση της μέγιστης τιμής του μερισματος
to subpatcher pfft~ pan8_engine~

Έχοντας αναλύσει τον αλγόριθμο του random walk στην εφαρμογή spectral delay, το ενδιαφέρον μας επικεντρώνεται στην εξήγηση του αλγόριθμου τοποθέτησης των μερισμάτων στην οκταφωνία.

Όπως έχουμε αναφέρει στο κεφάλαιο 3.4.3, η οκταφωνία επιτυγχάνεται όταν το σήμα πολλαπλασιάζεται με μία παραθυρική συνάρτηση ημίτονου η οποία θα διαβάζεται με διαφορά φάσης 45° ανάμεσα σε δύο γειτονικά ηχεία.

Η φάση δίνεται από 0 έως 1 αντί από 0 έως 360° . Έως εκ τούτου διαφορά φάσης 45° αντιστοιχεί σε 0.125 διαφορά φάσης για να διαβαστεί από το αντικείμενο [wave~]. Επειδή όμως θέλουμε το σήμα να βρίσκεται μόνο στο πρώτο κανάλι στη χαμηλότερη τιμή του multislider μονής περιοχής, βάζουμε ως αρχική φάση στο κανάλι 1 τη τιμή 0.125. Σε κάθε επόμενο κανάλι θα προστίθεται η τιμή 0.125. Ένα modulo 1 διασφαλίζει ότι εάν η τιμή ξεπεράσει το 1 θα προστεθεί το υπόλοιπο που ξεπερνάει το 1 στη τιμή 0.

Η μνήμη έχει καταχωρημένο το ημίτονο μόνο στο πρώτο τέταρτο της έκτασης της ενώ στο υπόλοιπο μέρος είναι αποθηκευμένες μόνο μηδενικές τιμές. Με αυτό το τρόπο επιτυγχάνουμε ότι όταν το σήμα δεν πολλαπλασιάζεται με το ημίτονο, να στέλνεται με μηδενικό πλάτος στην έξοδο.

Από τη δεξιά έξοδο του αντικειμένου [fftin~] εξέρχεται ο αύξων αριθμός των μερισμάτων του τρέχοντος παραθύρου. Η κάθε τιμή χρησιμεύει ως δείκτης ανάγνωσης του αντικειμένου [index~ 8-spek], που στέλνει στην έξοδο τις τιμές που έχουμε αποθηκεύσει στη μνήμη 8-spek επιλέγοντας τη τιμή που αντιστοιχεί στη θέση που φέρει τον ίδιο αριθμό με το δείκτη ανάγνωσης.

Η τιμή αυτή προστίθεται στην αρχική φάση του κάθε καναλιού.

Ο τελικός συντελεστής που προκύπτει πολλαπλασιάζεται με την ένταση του τρέχοντος μερίσματος η τιμή της οποίας προκύπτει από τη συνάρτηση [roltoCAR~]. και κατόπιν μεταβαίνει στην έξοδο για να υπολογιστεί μαζί με τα υπόλοιπα μερίσματα του παραθύρου, σε δύο το πολύ εξόδους (αφού η θέση του μερίσματος προκύπτει από δύο κανάλια το μέγιστο) από τις 8 υπάρχουσες (οι συναρτήσεις fftout1,2,3,...,8 πραγματοποιούν την inverse fft στο αντίστοιχο κανάλι) .

6. Παραγόμενα ηχητικά φαινόμενα

Για να γίνει εφικτός ο έλεγχος από τον μουσικό ή τον σχεδιαστή ήχου των ηχητικών φαινομένων που εμφανίζονται κατά τη διάρκεια χρήσης μίας μονάδας επεξεργασίας, απαιτείται να υπάρξει μία κατάταξη τους σε βασικές ομάδες-τεχνικές.

6.1 spectral delay

Θα αναφερθούμε στα ηχητικά φαινόμενα που παρουσιάζονται με βάση τη διαμόρφωση των περιοχών στο multislider που ελέγχει την εκάστοτε παράμετρο.

6.1.1. Η παράμετρος delay

- Όταν οι τιμές καθυστέρησης στα μερίσματα έχουν ομαδοποιηθεί (clusters) με τιμές πχ από 20Hz έως 600Hz 50msec, από 600Hz έως 6000Hz 100msec κοκ, τότε εμφανίζονται φαινόμενα τύπου step-filtering.

Όσο πιο μικρός είναι ο αριθμός των ομάδων (clusters) που εμφανίζονται και όσο πιο μεγάλος είναι ο μέγιστος χρόνος καθυστέρησης (παράμετρος max_delay) τόσο τα φαινόμενα step-filtering γίνονται πιο έντονα (ηχητικό παράδειγμα sample 1).



Σχήμα 45 σταδιακή διαβάθμιση των τιμών

- Όταν μία περιοχή από γειτονικά μερίσματα παρουσιάζει μία συνεχή σταδιακή διαβάθμιση των τιμών καθυστέρησής τους, τότε παράγεται η πιο χαρακτηριστική διαμόρφωση που προκαλεί το spectral delay στον ήχο προσδίδοντας του μία «ρευστή» χροιά.

Το ηχητικό φαινόμενο μπορεί να μελετηθεί εάν λάβουμε μία μόνο περίοδο της σταδιακής διαβάθμισης των τιμών που αποτελείται από μία συνεχή άνοδο και από μία συνεχή πτώση (ηχητικό παράδειγμα sample 2).



Σχήμα 46 συνεχής διαβάθμιση των τιμών

Στη συγκεκριμένη φασματική περιοχή θα παράγεται μία χαρακτηριστική περιοδική «ρευστοποίηση» του αρχικού ηχοχρώματος.

Εάν έχουμε δύο διαφορετικές φασματικές περιοχές που συντελείται μία σταδιακή άνοδος και πτώση των τιμών θα ακούγονται ταυτόχρονα και οι δύο διαμορφώσεις που προκαλούνται στον ήχο, εάν έχουμε τρεις θα ακούγονται τρεις χαρακτηριστικές διαμορφώσεις(ηχητικό παράδειγμα sample 3) ΚΟΚ.



Σχήμα 46 τρεις περιοχές συνεχούς διαβάθμισης

Όσο πιο μεγάλος είναι ο χρόνος καθυστέρησης τόσο πιο έντονα ακούγεται το συγκεκριμένο φαινόμενο. Εάν έχουμε μία ενδιάμεση περίπτωση σε σχέση με τις δύο προηγούμενες περιπτώσεις, δηλαδή η σταδιακή διαβάθμιση των τιμών να επιτυγχάνεται με τη κάθε τιμή που αλλάζει να παραμένει ίδια για n αριθμό sliders, τότε θα ακούμε ταυτόχρονα φαινόμενα «ρευστοποίησης» του ήχου και φαινόμενα step filtering(ηχητικό παράδειγμα sample 4).

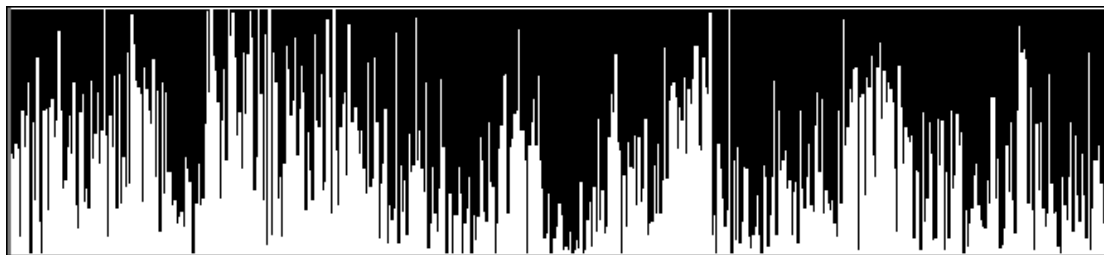


Σχήμα 46 συνεχής και σταδιακή διαβάθμιση των τιμών

- Στις δύο προηγούμενες περιπτώσεις, τα μερίσματα ανήκαν σε παράθυρα τα οποία σχετίζονταν με κάποιο τρόπο μεταξύ τους.

Στη πρώτη περίπτωση όπου εμφανίζονταν φαινόμενα step-filtering, ομάδες γειτονικών μερισμάτων είχαν ίδια τιμή, στη δεύτερη περίπτωση, τα γειτονικά μερίσματα προέρχονταν από γειτονικά παράθυρα. Όταν έχουμε τυχαίες τις επιλογές των τιμών

delay στη κάθε μπάντα τότε τείνουμε να ακούμε ένα είδος αλλοίωσης του σήματος που τείνει προς τη παραμόρφωση. Τα μερίσματα σε αυτή τη περίπτωση ανήκουν σε παράθυρα τα οποία δε σχετίζονται μεταξύ τους γι' αυτό το λόγο το σήμα χάνει τα ποιοτικά του χαρακτηριστικά αποκτώντας στοιχεία θορύβου που οφείλονται στη τυχαιότητα των τιμών (ηχητικό παράδειγμα sample 5).



Σχήμα 47 τυχαίες τιμές

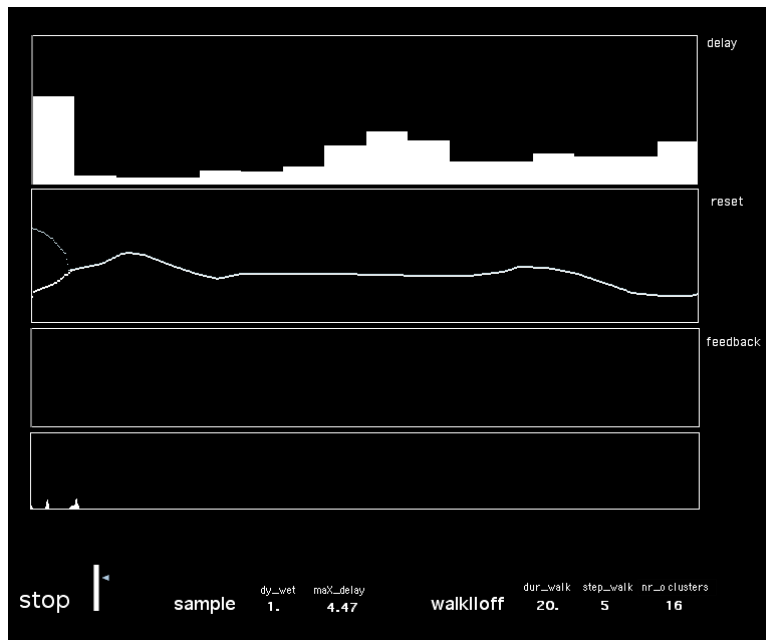
Σε γενικές γραμμές θα μπορούσαμε να ορίσουμε το ηχόχρωμα του διαμορφωτή όταν είναι απενεργοποιημένη η επιλογή random walk, ως ένα συνδυασμό από αυτές τις τρεις διαφορετικές ομαδοποιήσεις των τιμών του delay.

6.1.2. Η παράμετρος delay όταν μεταβάλλεται από το random walk

Για να γίνει ποιο κατανοητά τα ηχητικά φαινόμενα που αφορούν τον αυτοματοποιημένο έλεγχο των τιμών του delay μέσω του random walk, πρέπει να εξετάσουμε τι συμβαίνει στο σήμα όταν μόνο μία ομάδα γειτονικών sliders εκτελεί την κίνηση random walk.

Αυτή η τεχνική επιτυγχάνεται έχοντας ενεργοποιημένο το random walk (walk on). Οι δύο σειρές sliders στο δεύτερο multislidder που ελέγχουν τα όρια της κίνησής του, πρέπει να είναι ίσες σε όλη την έκταση του φάσματος με εξαίρεση μία περιοχή στην οποία θα είναι άνισες ώστε να επιτρέπεται η κίνηση μίας μόνο ομάδας sliders.

Στο σχήμα που ακολουθεί μόνο η πρώτη ομάδα sliders εκτελεί κίνηση random walk.



Σχήμα 48 μεταβολή των τιμών μέσω random walk

Σε αυτή τη περίπτωση οι sliders που θα εκτελούν τη κίνηση random walk ως ομάδα, θα διαλέγουν σε κάθε αλλαγή της τιμής τους, τα μερίσματα που ελέγχουν, από το ίδιο παράθυρο. Επομένως, μία περιοχή του φάσματος θα διαβάζεται κάθε φορά που δίνεται μία νέα τιμή random walk, από διαφορετικό σημείο της κυματομορφής.

Κατ' αυτό το τρόπο θα μπορούσαμε να πούμε ότι γίνεται μία μικροδομική σύνθεση στη συγκεκριμένη περιοχή του φάσματος και κατ' επέκταση σε όσες περιοχές του φάσματος εκτελείται η κίνηση random walk(ηχητικό παράδειγμα sample 6).

Όσο πιο μεγάλη είναι η περίοδος του random walk, καθώς και όσο πιο μικρός είναι ο αριθμός των φασματικών περιοχών τόσο πιο ευδιάκριτα θα είναι τα φαινόμενα μικροδομικής σύνθεσης στη κάθε περιοχή ξεχωριστά.

6.1.2. Η παράμετρος feedback

Ο τρίτος multislider ελέγχει τις τιμές feedback των μερισμάτων.

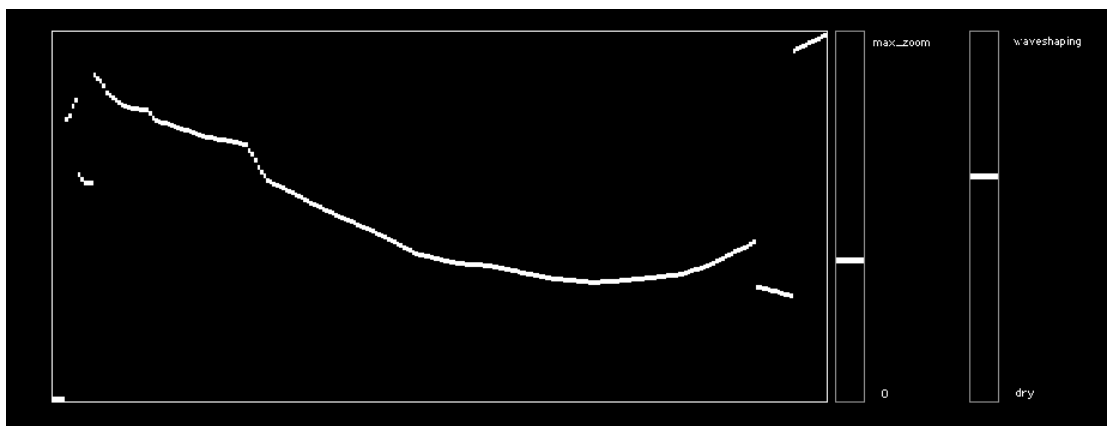
Όσο πιο μεγάλη τιμή feedback έχει ένα μερίσμα, τόσο ο ήχος του μερισματος θα τείνει να χάσει τα χαρακτηριστικά των δυναμικών του αποκτώντας μία σταθερή ένταση ενώ το φασματικό του περιεχόμενο θα ακούγεται σε συνεχή επανάληψη(ηχητικό παράδειγμα sample 7).

Με αυτό το τρόπο ο χρήστης έχει τη δυνατότητα να προσδώσει χαρακτηριστικά δρόνου στο τελικό ηχητικό αποτέλεσμα.

6.2. spectral gate

Η εφαρμογή spectral gate πέραν από τη δυνατότητα αποθρομβοποίησης που παρέχει στο χρήστη, τείνει ακουστικά να έχει χαρακτηριστικά φίλτρου, καθώς ο χρήστης επιλέγει ανάλογα με τη δυναμική που έχει το τρέχων μέρισμα, ποιες συχνότητες θα περάσουν στην έξοδο και ποιες όχι(ηχητικό παράδειγμα sample 8).

6.3. spectral waveshaping



Σχήμα 49 ενίσχυση των χαμηλών δυναμικών

Η ενίσχυση των χαμηλών δυναμικών ενισχύει τα transients του σήματος. Ο θόρυβος που εισέρχεται στο σήμα χαρακτηρίζεται από λεπτομέρεια ως προς τη κίνηση των δυναμικών αλλά και του φάσματος καθώς προέρχεται από λεπτομέρειες του φάσματος οι οποίες μέχρι ποιο πριν δε γινόντουσαν εύκολα αντιληπτές(ηχητικό παράδειγμα sample 9).

Η ενίσχυση των ποιο ψηλών δυναμικών προσδίδουν στο φάσμα χαρακτήρα διέλευσης του σήματος από φίλτρο(ηχητικό παράδειγμα sample 10).

6.4. spectral spatialization

Τα ηχητικά φαινόμενα spectral panning είναι δυσδιάκριτα καθώς η ανθρώπινη αντίληψη αδυνατεί να παρακολουθήσει τη κίνηση 512 φασματικών ζωνών στο χώρο. Για να γίνουν ποιο ευδιάκριτα πρέπει ο αριθμός των ομάδων των μερισμάτων που εκτελούν τη κίνηση random walk να είναι όσο το δυνατόν ποιο μικρός.

7. Συμπεράσματα

Έχοντας ολοκληρώσει και τις 4 εφαρμογές θα θέλαμε να επισημάνουμε κάποιες παρατηρήσεις για την ανάπτυξη των εφαρμογών.

Το περιβάλλον της max/msp έχει σχεδιαστεί για να ελέγχει τη πληροφορία που φέρουν τα σήματα μέσω του scheduler όπως αναπτύχθηκε στο περιβάλλον της max.

Το περιβάλλον δεν είναι τόσο ευέλικτο στον έλεγχο των σημάτων από το audio-rate όπως απαιτείται από το χρήστη όταν δουλεύει την fft ανάλυση/ανασύνθεση στο περιβάλλον της max/msp.

Κάποιες φαινομενικά απλές διαδικασίες μπορεί να αποδειχτούν εξαιρετικά δύσκολες λόγω του ότι ο χρήστης είναι υποχρεωμένος να δουλεύει σε audio-rate.

Όσον αφορά τις εφαρμογές, επιτύχαμε με τον έλεγχο της παραμέτρου delay των μερισμάτων μέσω του random walk να δημιουργήσουμε φαινόμενα μικροδομικού τύπου σε διαφορετικές φασματικές περιοχές. Σε αντίθεση με την εφαρμογή spectral delay της native instruments που αντιμετωπίζει το φάσμα σαν μία ενιαία περιοχή πάνω στην οποία εφαρμόζεται η επεξεργασία μετακινώντας όλη την εικόνα των παραμέτρων delay ταυτόχρονα, στη δική μας εφαρμογή διαιρέσαμε το φάσμα σε διαφορετικές ζώνες κάθε μία από τις οποίες τη διαχειριστήκαμε διαφορετικά.

Ταυτόχρονα όμως χάθηκαν τα έντονα ηχητικά φαινόμενα που παρατηρήσαμε όταν οι τιμές των γειτονικών μερίσματα ανήκουν σε παραπλήσια παράθυρα και τα οποία η εφαρμογή της NI τα εκμεταλλεύεται μετακινώντας τη συνολική εικόνα των παραμέτρων delay παλινδρομικά προς μία διεύθυνση.

Θα μπορούσε να επιτευχθεί ενδιαφέρων συνδυασμός μεταξύ των δύο τεχνικών εάν μετακινούνταν ένας μικρός αριθμός sliders με τη τεχνική του random walk, ενώ οι ενδιάμεσες τιμές ανάμεσα σε κάθε ζευγάρι γειτονικών sliders που εκτελούν τη κίνηση random walk, υπολογίζονται από τη μέθοδο σταθμισμένου μέσου.

Σε κάθε περίπτωση η παρούσα εργασία προσφέρει λύσεις σε όποιον θέλει να ασχοληθεί με την φασματική επεξεργασία του ήχου μέσω του προγραμματιστικού περιβάλλοντος max/msp, ανοίγοντας ένα δρόμο για μεταγενέστερες εφαρμογές.

Βιβλιογραφία

Boyle D 2004, 'Spectral delays with frequency domain processing', *Proc. Of the 7th Int. Conference of Digital Audio Effects*, pp. 42-44.

Διαμαντόπουλος,Τ 2004, *Προγραμματισμός και Σύνθεση Ήχου*, Εκδόσεις Ελλην, Αθήνα

Γαγάλης,Ν ,Γλαμπεδάκης,Μ , Κωτσάκης,Β, Κομισόπουλος,Β, Λευκαδίτης,Χ, Τραχάλιος,Μ 1992, *Μαθηματικός Λογισμός*, Μακεδονικές Εκδόσεις, Αθήνα

Miranda E 1998, *Computer Sound Design*, Elsevier Ltd, Great Britain

Native Instruments 2001, Spectral delay ver. 2.1,
http://www.computersandmusic.com/sitemap/p/Native_Instruments_Spektral_Delay_388.htm

πρόσβαση 18 Φεβρουαρίου 2008

Roads, C 1999, *The Computer Music Tutorial*, MIT Press, Massachusetts

Roads, C, Pope, S.T, Piccialli, A & Depdi, G 1997, *Musical Signal Processing*, Swets & Zeitlinger, Netherlands

Serra,Μ 1993,'Stochastic composition and stochastic timbre', *Perspectives of new music*, Vol 31, No 1, pp. 238-57.

VirSyn 2004, Cube ver. 2,
http://www.virsyn.de/en/E_Products/E_CUBE/e_cube.html

πρόσβαση 18 Φεβρουαρίου 2008

Wikipedia the free encyclopedia, *Random Walk*,

http://en.wikipedia.org/wiki/Random_walk,
http://en.wikipedia.org/wiki/Window_function,

πρόσβαση 6 Ιανουαρίου 2008

Wyatt,S 1999, *investigative studies on sound diffusion/projection*, e Contact,

<http://cec.concordia.ca/econtact/Diffusion/Investigative.htm>.

πρόσβαση 15 Ιουνίου 2008

Θα ήθελα να εκφράσω ένα θερμό ευχαριστώ στον επιβλέποντα
καθηγητή μου
Νικόλα βαλσαμάκη, στη Κατερίνα Τζεδάκη, στον Παναγιώτη
Κόκορα, για όλες τις υπέροχες συζητήσεις, παραστάσεις, εικόνες
που μου προσέφεραν.