ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Αλγοριθμική σύνθεση και αυτόματη ηχοποίηση/ μουσικοποίηση βίντεο με χρήση Η/Υ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χρήστος Π. Καραμήτσος

Επιβλέπων: Γιώργος Καμπουράκης
Καθηγητής ΕΜΠ

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Αλγοριθμική σύνθεση και αυτόματη ηχοποίηση/ μουσικοποίηση βίντεο με χρήση Η/Υ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χρήστος Π. Καραμήτσος

Επιβλέπων: Γιώργος Καμπουράκης
Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την
_____

_____	_____	_____
Γιώργος Καμπουράκης	Βασίλης Λούμος	Ελευθέριος Καγιάφας
Καθηγητής	Καθηγητής	Καθηγητής

Αθήνα, Μάιος 2014

.....

Χρήστος Π. Καραμήτσος
Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Χρήστος Καραμήτσος 2014

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Υπάρχει μια βαθιά εδραιωμένη αντίληψη από τον περισσότερο κόσμο, πως η διαδικασία της μουσικής σύνθεσης είναι μια αυθόρμητη και πλήρως καλλιτεχνική διαδικασία, στην οποία δε γίνεται χρήση τυποποιημένων μηχανισμών και διαδικασιών. Με μια προσεκτικότερη μελέτη όμως στο έργο των σπουδαιότερων δημιουργών, από την εποχή που η μουσική έπαψε να είναι αυτοσχεδιαστική, φαίνεται πως η αλήθεια απέχει αρκετά από αυτή την άποψη. Ειδικότερα, από την εποχή της ακμής για τη δυτική μουσική, περί τον 15^{ο} αιώνα και μετά, οι συνθέτες έχουν καταστρώσει κι εφαρμόσει αναρίθμητες διαφορετικές αλγοριθμικές διαδικασίες, με σκοπό την απόκτηση συστηματικότερης προσέγγισης, αλλά και την επίτευξη πρωτότυπων αποτελεσμάτων

Αυτές οι διαδικασίες, συχνά περιέχουν μαθηματικά στατιστικά μοντέλα, στοιχεία δηλαδή που προσθέτουν το στοιχείο της τυχαιότητας στη δημιουργία. Οι υπολογιστές φυσικά προσέφεραν τεράστια βοήθεια στους μουσικούς από τον 20^{ο} αιώνα και μετά, χάρη στους μεγάλους αποθηκευτικούς χώρους και την δυνατότητα για γρήγορους και ακριβείς υπολογισμούς. Σκοπός αυτής της εργασίας είναι η κατασκευή ενός υπολογιστικού προγράμματος, το οποίο κάνοντας χρήση της διαδικασίας του Video Sonification, δηλαδή ηχοποίησης βίντεο, ως μεθόδου αλγοριθμικής σύνθεσης, δημιουργεί μουσική σε πραγματικό χρόνο.

Λέξεις-κλειδιά

Αλγοριθμική σύνθεση, ηχοποίηση, πραγματικός χρόνος

Abstract

Although music composition has always been regarded as a purely spontaneous and artistic procedure, which uses no mechanisms or structures, a closer research of the compositional methods of artists through the years reveals that this is far from true. Throughout the years, especially since the rise of western music in the 15th centrury, composers have employed numerous algorithmic procedures in order to create music in a more consistent and systematical way. This procedures often included the use of mathematical statistical models, as well as elements that provided an extra chance factor. In that way, the result was not predetermined. Computers ofcourse proved to be of great aid to musicians of the 20th century, providing large databases and the ability to perform very fast calculations, that were impossible and also painful to perform by hand.

This thesis is exploring one of the possible and most impressive algorithmic composition tools computers are able to provide to a composer, through the use of Video Sonification. The program created for the purpose of this thesis creates a musical piece in real-time for any video chosen by the user.

Keywords

Video Sonification, Algorithmic composition, statisticam models

Ευχαριστώ θερμά τον επιβλέποντα καθηγητή κ. Γιώργο Καμπουράκη, καθώς και τον Κωνσταντίνο Μπακογιάννη, υποψήφιο διδάκτορα, για την πολύτιμη βοήθεια, καθοδήγηση και υποστήριξη που μου παρείχαν καθ' όλη τη διάρκεια εκπόνησης της παρούσας εργασίας, καθώς και για τις χρήσιμες ιδέες που προσέφεραν.

ΠΕΡΙΕΧΟΜΕΝΑ

1. Εισαγωγή……10
1.1 Αλγοριθμική μουσική……...10
1.2 Τι είναι μουσική……. ……..10
1.3 Σκοπός της Εργασίας…….…...11
1.4 Ιστορική Αναδρομή στην αλγοριθμική μουσική……..11
1.5 Επίδραση Η/Υ στην αλγοριθμική μουσική……...14
1.6 Μεταγενέστερες υπολογιστικές τεχνικές αλγοριθμικής σύνθεσης…...16
1.7 Φιλοσοφία, Αισθητική αλγοριθμικής σύνθεσης……17

2. Ήχος και τα χαρακτηριστικά του……...……..18
2.1 Ο τρόπος που αντιλαμβανόμαστε τον ήχο……18
2.2 Αρμονικοί ήχοι……...19
2.3 Μη αρμονικοί τόνοι-Θόρυβος……20
2.4 Μέθοδοι σύνθεσης ήχου……21
2.4.1 Προσθετική σύνθεση (Additive synthesis)……..21
2.4.2. Αφαιρετική σύνθεση (Subtractive Synthesis)……22
2.4.3 Σύνθεση με μικροδομές (Granular synthesis)……..23
2.4.4 Σύνθεση με διαμόρφωση πλάτους και (Amplitude-ring modulation synthesis)……23
2.4.5 Σύνθεση πίνακα κυματομορφής (Wavetable synthesis)……..23
2.4.6 Σύνθεση φυσικών μοντέλων (Physical Modelling Synthesis)…….24
2.5 Ρυθμικά στοιχεία ενός μουσικού κομματιού……..24
2.5.1 Μουσικές αξίες……24
2.5.2 Μέτρο-tempo……. …….25

3. Τί είναι Sonification(Ηχοποίηση)……..27
3.1 Ηχοποίηση εικόνας (Image Sonification) ……..27
3.2 Αλγοριθμική μουσική μέσω Sonification……...28
3.3 Video Sonification και αλγοριθμική μουσική……28
3.3.1 Video Sonification με οπτικό υλικό από το διάστημα…….28
3.3.2 Τρόπος χρήσης του Video Sonification στην παρούσα εργασία….29

4. Συνοπτική παρουσίαση του MAX/MSP……30
4.1 Βασικά στοιχεία και λειτουργίες του MAX/MSP……...30
4.2 Πρωτόκολλο MIDI (*Musical Instrument Digital Interface)* στο MAX/MSP……31
4.3 Δομημένος προγραμματισμός στο MAX/MSP……32

5. Χρώμα και μοντέλα προσομοίωσης του από Η/Υ…….35
5.1 Τι είναι χρώμα…….35
5.2 Πως αντιλαμβανόμαστε τα χρώματα;……36
5.3 Ραβδία και Κονία…….36
5.4 Χαρακτηριστικά αντίληψης των χρωμάτων……38
5.5Χρωματικά μοντέλα(C. models) και Χρωματικοί χώροι (c. spaces).....39
5.5.1 Χρωματικό μοντέλο και Χρωματικός χώρος RGB……..41
5.5.2 Χρωματικό μοντέλο και Χρωματικός χώρος HSV και HSL……41
5.6 Επεξεργασία βίντεο στο MAX/MSP……..43

6. Επεξήγηση αλγορίθμου……...45
6.1 Εξαγωγή δεδομένων από βίντεο……..47
6.1.1 Εύρεση του ρυθμού χρωματικής μεταβολής του βίντεο……48
6.1.2 Εντοπισμός σημαντικών εικονοστοιχείων (pixel)……...50
6.1.3 Υπολογισμός μέσης φωτοβολίας(luminosity) - επιλογή έντασης(volume)…….53
6.1.4 Υπολογισμός μέσων τιμών H, S, L για τα ενδιαφερόμενα εικονοστοιχεία…….54
6.2 Επεξεργασία δεδομένων και αναπαραγωγή ήχου……...55
6.2.1 Ορισμός Tempo……55
6.2.2 Ρυθμός μελωδίας……..57
6.2.3 Παραμετροποίηση HSL δεδομένων…….58
6.2.4 Κίνηση μπάσου……60
6.2.5 Κρουστά……62
6.2.6 Συγχορδίες……63
6.2.7 Πέμπτες……64
6.2.8 Επιλογή μουσικού οργάνου αναπαραγωγής……66
6.3 Μορφή παρουσίασης στο βασικό patcher……...67

7. Προτάσεις για εξέλιξη του αλγορίθμου……..68
7.1 Με χρήση αρμονικών ήχων-Τονικών κέντρων……68
7.2 Εισαγωγή Μη-αρμονικών τόνων, Θορύβου μέσω του MSP……69

8. Βιβλιογραφία/ Πηγές……71

1. Εισαγωγή

1.1 Αλγοριθμική Μουσική

Ο όρος αλγοριθμικη μουσική, εν έτη 2014, είναι πια άρρητα συνδεδεμένος με τον όρο υπολογιστική μουσική. Στην πραγματικότητα όμως, πρόκειται για έναν κλάδο, ο οποίος γνώρισε την πρώτη του μεγάλη ανάπτυξη στις αρχές του 20ου αιώνα, όπου οι Η/Υ δεν είχαν αναπτυχθεί αρκετά ώστε να φτάσουν να αποτελούν αναπόσπαστο κομμάτι πια αυτής της έννοιας . Ως αλγόριθμος, ορίζεται μια πεπερασμένη σειρά ενεργειών, αυστηρά καθορισμένων και εκτελέσιμων σε πεπερασμένο χρόνο, που στοχεύουν στην επίλυση ενός προβλήματος. Πιο απλά **αλγόριθμο** ονομάζουμε μία σειρά από εντολές που έχουν αρχή και τέλος, είναι σαφείς και εκτελέσιμες που σκοπό έχουν την επίλυση κάποιου προβλήματος.[1].

Στην περίπτωση της αλγοριθμικής μουσικής, λοιπόν, το "πρόβλημα" προς επίλυση είναι η σύνθεση ενός μουσικού κομματιού. Ο όρος μουσική, βέβαια, είναι μια άκρως παρεξηγήσιμη έννοια και για να γίνει αρχικά κατανοητή η έκταση της σύγχυσης, αρκεί να αναφερθεί πως έχουν υπάρξει αμέτρητες συνθέσεις στο παρελθόν, οι οποίες είχαν κατηγοριοποιηθεί, υποτιμητικά, ως "θόρυβος" την περίοδο δημιουργίας, αλλά και κυκλοφορίας τους, οι οποίες όμως κάποιες δεκαετίες μετά, χαρακτηρίστηκαν ως πρωτοποριακές και άρα αποδεκτές από τη μουσική κοινότητα (τουλάχιστον από κάποια σημαντική μερίδα της). Με το πέρασμα του χρόνου, τα όρια της έννοιας "μουσικό κομμάτι" έχουν διευρυνθεί σημαντικά.

1.2 Τι είναι μουσική;

Στην ουσία, πρόκειται για ένα θέμα όπου η υποκειμενική αισθητική, αλλά και η χρονική περίοδος, παίζει μεγάλο ρόλο. . Φυσικά και η έννοια της λεξης "θόρυβος" έχει αλλάξει σημαντικά σε σχέση με το παρελθόν και δεν έχει πια τόσο αρνητική χροιά. Κάτι τέτοιο γίνεται φανερό από το γεγονός ότι υπάρχουν πια μουσικά ρεύματα (noise music), των οποίων κύριο γνώρισμα αποτελεί η χρήση θορύβου με συνέπεια. Υπάρχουν και παραδείγματα όπου η ένταξη του γίνεται σε κατά τα άλλα μελωδικές καταστάσεις, όπως π.χ. ο Τζίμι Χέντριξ, ό οποίος έδωσε μια νέα διάσταση στον κιθαριστικό ήχο, παραμορφώνοντάς τον, ώστε να ισορροπεί ανάμεσα στον θόρυβο και στη μουσική.

Τόσο ο ορισμός της μουσικής, όσο και σχετικά με τη μουσική θέματα όπως η εκτέλεση, η σύνθεση και η σπουδαιότητά της, διαφέρουν από πολιτισμό σε πολιτισμό και ανάλογα με το κοινωνικό πλαίσιο. Ένας αυστηρός ορισμός θα ήταν ότι ως **μουσική** ορίζεται η τέχνη που βασίζεται στην οργάνωση ήχων με σκοπό τη σύνθεση, εκτέλεση και ακρόαση/λήψη ενός μουσικού έργου .[1] Η ερώτηση 'τί είναι μουσική;' έχει γίνει θέμα άπειρων συζητήσεων και φυσικά έχει δεχτεί πληθώρα απαντήσεων, όμως καμία απάντηση δεν έχει γίνει καθολικά αποδεκτή.

Ένας άλλος ορισμός, ο οποίος προέρχεται από το Βρετανικό λεξικό της Οξφόρδης, είναι ο εξής, πρόκειται για "μια από τις καλές τέχνες που ασχολείται με το συνδυασμό ήχων με σκοπό την ομορφιά ως προς τη φόρμα και την έκφραση των σκέψεων και συναισθημάτων." Αίσθηση έχει προκαλέσει και η δήλωση του σπουδαίου μουσικοσυνθέτη Έντγκαρ Βαρές (Edgar Varese), ο οποίος χαρακτηρίζει τη μουσική ως 'οργανωμένο ήχο',[1]. Πρόκειται για μια περιγραφή, η οποία χρησιμοποιείται αρκετά συχνά πια, ειδκά από τις αρχές του 20ου αιώνα και μετά, καθώς πολλές συνθέσεις δεν αποτελούνται πια αποκλειστικά και μόνο από νότες, αλλά κι από ήχους. Οι έννοιες αυτές θα αναλυθούν εις βάθος παρακάτω, έτσι ώστε να γίνει σαφέστερος ο διαχωρισμός τους.

1.3 Σκοπός της Εργασίας

Αυτή η εργασία πραγματεύεται την δημιουργία μουσικής μέσω της επεξεργασίας των παραμέτρων ενός οποιουδήποτε βίντεο. Για την επίτευηξη του σκοπού αυτού, γίνεται χρήση ενός προγαμματιστικού περιβάλλοντος, το MAX/MSP μαζί με το συμπληρωματικό του εργαλείο Jitter, το οποίο είναι υπεύθυνο για το χειρισμό και την επεξεργασία εικόνων και βίντεο. Το πρόγραμμα που δημιουργήθηκε κάνει χρήση των χρωματικών μεταβολών του βίντεο για κάθε καρέ σε πραγματικό χρόνο και δημιουργεί ένα μουσικό κομμάτι με χαλαρή δομή και ενδιαφέροντα μελωδικά και ρυθμικά μοτίβα, ανάλογα με τις παραλλαγές του εκάστοτε βίντεο. Αυτή η μορφή σύνθεσης, έχει το επιπρόσθετο πλεονέκτημα να δημιουργεί για τον ακροατή/θεατή ένα οπτικοακουστικό θέαμα. Επίσης, ο αλγόριθμος είναι μη-ντετερμινιστικός, κάνοντας χρήση στατιστικών μοντέλων, όπως αλυσίδες Markov. Το λογισμικό που χρησιμοποιήθηκε για την υλοποίηση της διαδικασίας είναι το MAX/MSP, ένα γραφικό προγραμματιστικό περιβάλλον ειδικά σχεδιασμένο για μουσική και πολυμέσα.

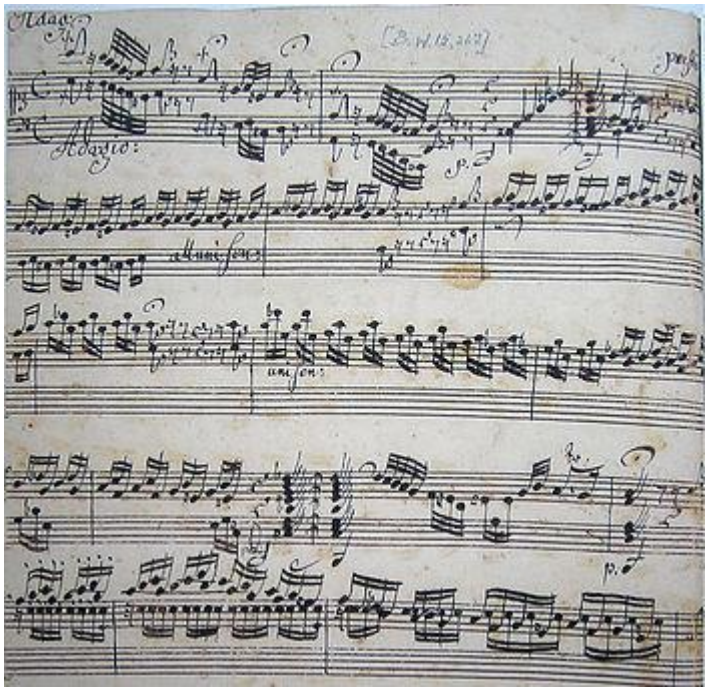
1.4 Ιστορική αναδρομή στην αλγοριθμική μουσική

Για να επανέλθουμε όμως στο αρχικό μας ζήτημα, η αλγοριθμική μουσική πραγματεύεται τη σύνθεση μουσικής μέσω μιας αυτοματοποιημένης διαδικασίας, καθορισμένης από το συνθέτη. Η αλήθεια είναι ότι η αλγοριθμική μουσική, ειδικά ύστερα από την εισαγωγή των υπολογιστών στη διαδικασία, έχει βρει κατά καιρούς αρκετούς πολέμιους. Πολλοί αμφισβητούν την καλλιτεχνική αξία των δημιουργημάτων αυτών, κυρίως διότι ο τρόπος γέννησης τους ενοχλεί και μοιάζει απόκοσμος και μηχανικός, σε σχέση με την επικρατούσα ιδέα για τον "αυθεντικό" καλλιτέχνη. Για έναν "αυθεντικό" καλλιτέχνη, σύμφωνα με την κοινή γνώμη, όλα τα δηιουργήματα του αποτελούν προϊόν στιγμιαίας έμπνευσης και καλλιτεχνικού οίστρου.

Στην πραγματικότητα όμως, κάτι τέτοιο δεν ισχύει, ακόμα και για τους σημαντικότερους, κατά γενική παραδοχή, συνθέτες του Δυτικού πολιτισμού. Εδώ και χίλια χρόνια χρησιμοποιούνται κανόνες, υπολογισμοί και ανέμπνευστες τυπικές διαδικασίες για τη σύνθεση όλων αυτών των κλασσικών εργών[2]. Το γεγονός αυτό όμως δε θα έπρεπε να

αντιλαμβάνεται ως κάποιο αρνητικό στοιχείο, καθώς αρκεί να αναλογιστούμε τις αντίστοιχες αναλογίες που επικρατούν για τη δημιουργία π.χ. μιας φαντασμαγορικής τοιχογραφίας, ή ενός εντυπωσιακού αρχιτεκτκονικού κτιρίου. Και στα 2 αυτά παραδείγματα είναι απαραίτητη κάποια τυπική προεργασία και σχεδιασμός. Ένας παρόμοιος σχεδιασμός υπάρχει στη μουσικη του δυτικού κόσμου, από την εποχή που έπαψε να είναι αυτοσχεδιαστική, πριν ακόμα εφευρεθεί η έννοια του αλγορίθμου ως λέξη.

Μεγαλοπρεπές κτίριο 18ου αιώνα στην Αυστρία

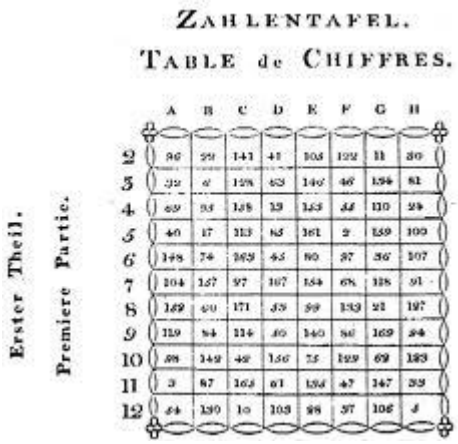


Απόσπασμα από τη ***Τοκάτα και Φούγκα σε Ρε Ελάσσονα*** *BWV 565, Γ.Σ. Μπαχ*

Οι απαρχες μιας τετοιας σκεψη, της σύνδεσης δηλαδή μιας τέχνης, με μια επιστημονική μέθοδο, βρισκονται στην εποχή της Αρχαίας Ελλάδας και συγκεκριμένα στον Πυθαγόρα και στους μαθητές του, ο οποίος πίστευε πως υπήρχε άμεση σχέση των αριθμών με τη μουσική και κατά προέκταση τη φύση. Όμως, όπως αναφέρθηκε προηγουμένως, ήταν απλά σκέψη, καθώς η ελληνική μουσική ήταν σχεδόν εξολοκλήρου αυτοσχέδια και άρα δεν είχε αναπτυχθεί κάποια συστηματική μέθοδο σύνθεσης.

Ένα από τα πρώτα παραδείγματα συστηματοποιημένης μεθόδου σύνθεσης, αποτελεί η σύνθεση με χρήση ενός κανόνα, κάτι το οποίο ξεκίνησε περί τα τέλη του 15ου αιώνα. Ο συνθέτης έγραφε τη μελωδία της βασικής φωνής, ύστερα επέλεγε έναν κανόνα, τον οποίο θα ακολουθούσε για τις υπόλοιπες φωνές. Αυτός μπορεί να ήταν π.χ. η δεύτερη φωνή να μιμηθεί ακριβώς τη μελωδία της πρώτης, ξεκινώντας όμως μερικά μέτρα αργότερα ή να αποτελεί κάποια αναστροφή της πρώτης φωνής κάνοντας χρήση συγκεκριμένων μουσικών διαστημάτων.

Κάτι που αξίζει να σημειωθεί, είναι ότι από το 18ο αιώνα ακόμα, υπήρχαν μέθοδοι σύνθεσης, οι οποίες περιλάμβαναν τυχαιότητα. Δηλαδή ενώ μιλάμε για αυτοματοποιήμενες διαδικασίες, το αποτέλεσμα δεν ήταν κάθε φορά ίδιο, καθώς ο αλγόριθμος ήταν μη ντετερμινιστικός. Ένα τέτοιο παράδειγμα ήταν το *Musikalisches Wurfelspie*l, του Βόλφγκανγκ Αμαντέους Μότσαρτ, το οποίο σημαίνει μουσικό παιχνίδι ζαριών. Εκεί, όπως υποδεικνύει και η ονομασία, έχοντας στη διάθεση του διάφορα αποσπάσματα μουσικών κομματιών, με τη χρήση ζαριών έφτιαχνε ένα ενιαίο σύνολο, το οποίο αποτελούνταν από ένα τυχαίο συνδυασμό των προαναφερθέντων υπομονάδων.



Πίνακας αντιστοίχισης για μουσικό παιχνίδι ζαριών του 18ου αιώνα

Στη συνέχεια, ο κλάδος γνώρισε πολύ μεγάλη εξέλιξη, στις αρχές του 20ου αιώνα, και συγκεκριμένα λίγο μετά τον 2ο Παγκόσμιο Πόλεμο, με τις μεθόδους του δωδεκαφθογγισμου και του σειριαλισμου να ξεχωρίζουν. Με τη χρήση τους εξαγόταν αποτέλεσμα για αμέτρητες παραμέτρους ενός κομματιού απο τις νότες , μέχρι το ρυθμό, αλλά ακόμα και τις δυναμικές. Οπότε από αυτή την εποχή και ύστερα μπορούμε να μιλάμε για μια πραγματικά συστηματοποιημένη διαδικασία δημιουργίας αλγορίθμικής μουσικής, με την έννοια που υπάρχει μέχρι σήμερα.[3]

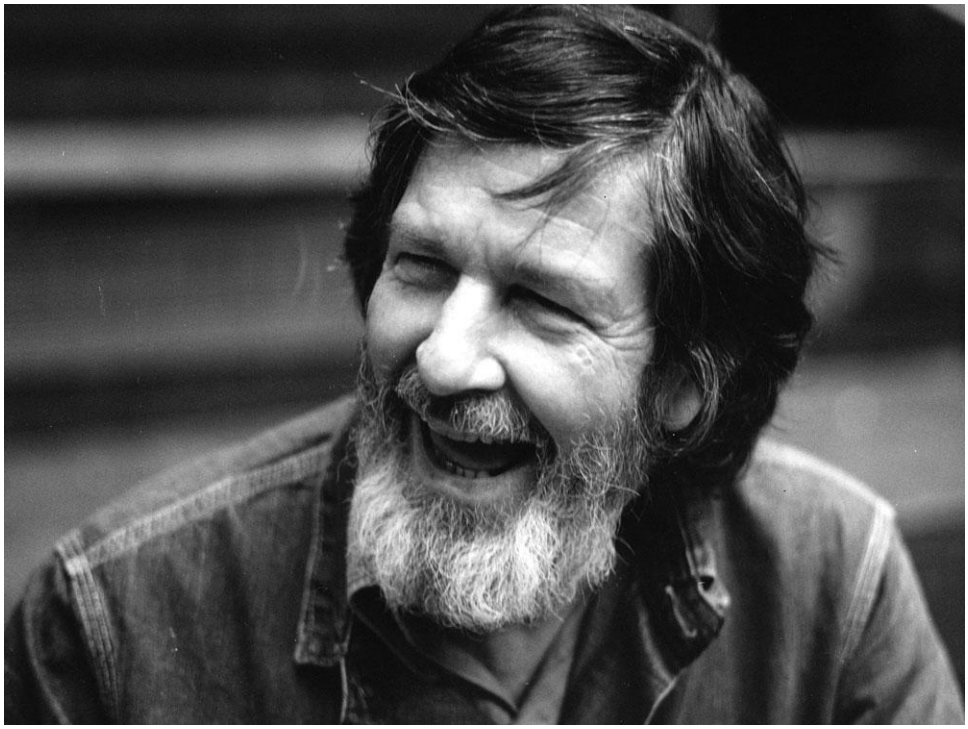
1.5 ΕΠΙΔΡΑΣΗ Η/Υ ΣΤΗΝ ΑΛΓΟΡΙΘΜΙΚΗ ΜΟΥΣΙΚΗ

Όπως θα φαντάζεται εύκολα κάποιος, η άνοδος των Η/Υ συνιστά ασύγκριτη βοήθεια σε αυτό το ρεύμα, καθώς προσέφεραν άπλετο αποθηκευτικό χώρο αλλά και απίστευτα ταχύτερη, ακριβέστερη και φυσικά ευκολότερη επεξεργασία των αλγορίθμων. Οι πρώτοι που πειραματίστηκαν κάνοντας χρήση Η/Υ για τη σύνθεση ενός μουσικού κομματιού ήταν οι **Lejaren Hiller** και **Leonard Isaacson,** 2 καθηγητές του πανεπιστημίου του Ιλλινόις το 1956. Το έργο που παρουσίασαν ονομάζονταν illiac suite, γραμμένο για κουαρτέτο εγχόρδων και το οποίο πήρε το όνομά του από τον υπολογιστή illiac, ο οποίος χρησιμοποιήθηκε για τη σύνθεση, αλλά και για την εκτέλεση του κομματιού [4]. Στην ουσία ο αλγόριθμος ξεκινούσε από την παραγωγή πολλών τυχαίων αριθμών, οι οποίοι αντιστοιχούσαν σε κάποια νότα μιας μουσικής σκάλας, μέσω μιας αντιστοίχισης που είχε γίνει προηγουμένως, αλλά αργότερα και σε ρυθμικά μοτίβα, δυναμικές και πολλές άλλες παραμέτρους. Στη συνέχεια αυτοί οι αριθμοί περνούσαν από διάφορους ελέγχους, χρησιμοποιώντας κανόνες μουσικής θεωρίας, στην ουσία φιλτράροντας τα τυχαία αρχικά αποτελέσματα [5].

Φωτογραφία του Lejaren Hiller με τον υπολογιστή Illiac

Τα χνάρια τους ακολούθησε κι ένας σπουδαίος Έλληνας συνθέτης, ο Ιωάννης Ξενάκης, παρουσιάζοντας τα έργα ST/4 και ST/10, το 1962. Τα αρχικά τους (ΣΤ) προέρχονται από το στοχαστικά, καθώς δημιουργήθηκαν κάνοντας χρήση στοχαστικών διαδικασιών. Αυτά είναι στατιστικά μοντέλα που ερμηνεύουν τη συμπεριφορά φαινομένων "μεγάλων αριθμών" [1]. Πρόκειται για ένα πρόγραμμα το οποίο εξήγαγε την παρτιτούρα του κομματιού, έχοντας στη διάθεση του μια λίστα από αρκετές νότες και κάποια πιθανοτικά βάρη, τα οποία εισήγαγε ο προγραμματιστής, . Κάποιες τελευταίες αποφάσεις που αφορούσαν τη δομή προέρχονταν από τυχαίους αριθμούς, οι οποίοι παρέχονταν πάλι από τον υπολογιστή. Αξίζει να σημειωθεί ότι ο ίδιος ο συνθέτης επεξηγεί πλήρως τη διαδικασία στο βιβλίο που εξέδωσε, με τίτλο, *Formalized Music*. Σε αντίθεση βέβαια με το *illiac suite*, το έργο εκτελούνταν από πραγματικούς οργανοπαίχτες, με χρήση παραδοσιακών μουσικών οργάνων[3]. Αξίζει να γίνει ιδιαίτερη μνεία στον John Cage, ο οποίος ήταν ένας από τους πρώτους που κατάλαβαν την αξία της βοήθειας που προσέφεραν οι υπολογιστές. Συνεργάστηκε μάλιστα με τον Lejaren Hiller, έναν από τους δημιουργούς του illiac suite, για να φτιάξουν μια σύνθεση ονόματι *HPSCHD*, η οποία ουσιαστικά αποτελούσε κολλάζ διάφορων αποσπασμάτων από Μότσαρτ και άλλους καταξιωμένουςν παλαιότερους συνθέτες. Οι αποφάσεις για τη σειρά εκτέλεσης των αποσπασμάτων παίρνονταν από τον υπολογιστή με παραγωγή τυχαίων αριθμών.[6]

Ιωάννης Ξενάκης



John Cage

1.6 Μεταγενέστερες υπολογιστικές τεχνικές αλγοριθμικής σύνθεσης

Μία προσέγγιση που έγινε εφικτή μέσω των τεράστιων δυνατοτήτων των Η/Υ είναι τα συστήματα τεχνητής νοημοσύνης. Αυτά μοιάζουν με συστήματα που έχουν κάποιους προκαθορισμένους κανόνες, δηλαδή μια "γραμματική", με τη διαφορά πως μπορούν να διαμορφώσουν τους κανόνες αυτόνομα, η αλλιώς να προσαρμοστούν "μάθαίνοντας". Ένα τέτοιο παράδειγμα αποτελεί το Experiments in Musical Intelligence (EMI), του David Cope. Αυτό το σύστημα διαθέτει μια ευρεία βάση δεδομένων από κανόνες, περιγραφές μουσικών ειδών και στρατηγικές σύνθεσης, όμως μπορεί επιπρόσθετα, με χρήση παρτιτούρες ενός συγκεκριμένου δημιουργού ως είσοδο, να φτιάξει δικούς του κανόνες και γραμματική. Έχει γίνει επιτυχημένη χρήση του για παραγωγή μουσικής που να προσομοιώνει συνθέσεις των Bach, Brahms, Mozart, Bartók. [3]

Γενετικοί αλγόριθμοι

Ένα παρακλάδι αλγοριθμικής σύνθεσης με χρήση συστημάτων τεχνητής νοημοσύνης είναι οι γενετικοί αλγόριθμοι, που είναι μία πολύ πρόσφατη τεχνική που ανήκει στον αυτόματο προγραμματισμό. Γενικά, οι γενετικοί αλγόριθμοι χρησιμοποιούνται για την εύρεση βέλτιστων λύσεων σε πολύ-παραμετρικά προβλήματα. Αν και συνήθως η διαδικασία της επιλογής βασίζεται σε προκαθορισμένα κριτήρια καταλληλότητας, υπάρχουν εφαρμογές κατά τις οποίες ο χρήστης καλείται να αλληλεπιδρά με το σύστημα κατά τη διάρκεια της επιλογής.[7]. Στην περίπτωση ενός προγράμματος μουσικής σύνθεσης, ο δημιουργός πρέπει να προγραμματίσει μια "κριτική" συνάρτηση, οποία επιβλέπει πολλά αποτελέσματα του

βασικού προγράμματος και αποφασίζει ποια είναι καταλληλότερα για το τελικό αποτέλεσμα. Στη συνέχεια παράγονται νέες συνθέσεις οι οποίες βασίζονται σε όσες πέρασαν τον αρχικό έλεγχο. Η διαδικασία επαναλαμβάνεται ως ότου καταλήξει σε μια βέλτιστη λύση ή μέχρι να γίνει ένας συγκεκριμένος αριθμός επαναλήψεων. Σε αυτή την περίπτωση λοιπόν, ο χρήστης δίνει στον υπολογιστή διάφορα εργαλεία, τα οποία επιλέγει να χρησιμοποιήσει έτσι ώστε να πληροί τις προϋποθέσεις ενός επιθυμητού αποτελέσματος, το οποίο είναι ορισμένο από το συνθέτη.

1.7 Φιλοσοφία, Αισθητική αλγοριθμικής σύνθεσης

Η αισθητική, είναι σίγουρα μια έννοια υποκειμενική, αλλά παρόλα αυτά έχουν διατυπωθεί αρκετές θεωρίες για το πώς θα πρέπει να κρίνεται η αξία ενός καλλιτεχνικού έργου. Στην περίπτωση μας, είναι εμφανές ότι έχει μεγάλη αξία, ο ίδιος ο αλγόριθμος που καταστρώνεται, συγκρίσιμη ακόμα και με το τελικό ηχητικό αποτέλεσμα. Επομένως θα ήταν εύλογη σκέψη, ο εκάστοτε αλγόριθμος να κρίνεται από κάποια κοινώς αποδεκτά κριτήρια. Τέτοια κριτήρια έχουν διατυπωθεί από τον Donald Knuth, διεθνώς αναγνωρισμένο επιστήμονα της πληροφορικής και είναι τα εξής. Ο αλγόριθμος να εμπεριέχει απλότητα, οικονομία, κομψότητα και να χειρίζεται σχετικά εύκολα[8]. Όμως, οι περισσότεροι από τους καταγγεγραμμένους αλγορίθμους δεν ακολουθούν αυτά τα κριτήρια.

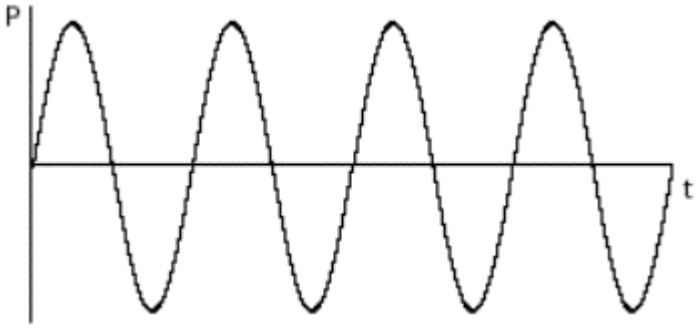
Γι'αυτό μια ομάδα οπτικών αλλά και ακουστικών καλλιτεχνών διατύπωσαν τα δικά τους κριτήρια, τα οποία στην ουσία προσμετρούν τόσο την πρωτοτυπία και την αξία της σύλληψης, όσο και το ηχητικό αποτέλεσμα αυτής, δηλαδή την εκτέλεση. Το πρόβλημα σε αυτή την ιδέα έγκειται στο ότι πολλές φορές, ο απλός ακροατής δεν γνωρίζει τη σκέψη που οδήγησε στη σύνθεση, επομένως δεν μπορεί να κρίνει συνολικά το έργο. Άρα τελικά, ο δημιουργός οφείλει να καταστρώσει έναν πρωτότυπο αλγόριθμο που θα πετύχει τον επιθυμητό στόχο και φυσικά να κάνει τις απαραίτητες δοκιμές έτσι ώστε να σιγουρευτεί πως το ηχητικό αποτέλεσμα είναι ικανοποιητικό, γιατί εν τέλει αυτό είναι που θα μείνει στην ιστορία.

2. Ήχος και τα χαρακτηριστικά του

Όλες οι συνθέσεις μέχρι και τον 20ο αιώνα, αποτελούνταν από ένα σύνολο νοτών γραμμένων σε παρτιτούρες. Από τον 20ο αιώνα και μετά όμως παρατηρούντα έργα στα οποία ο ακροατής συναντάει "ήχους", σε ένα αποτέλεσμα μη μελωδικό, τουλάχιστον με τα συμβατικά κριτήρια. Είναι, λοιπόν, επιτακτική μία σύντομη επεξήγηση των 2 αυτών εννοιών έτσι ώστε να μπορέσουν να αποσαφηνιστούν πλήρως.

2.1 Ο τρόπος που αντιλαμβανόμαστε τον ήχο

Ότι ακούμε, οφείλεται στις διαταραχές της πίεσης στον αέρα, του εκάστοτε περιβάλλοντος. Αυτές οι διαταραχές αποτελούν στην πραγματικότητα, μικροσκοπικές διακυμάνσεις της φυσιολογικής πίεσης, η οποίες προκαλούνται από αντικείμενα που δονούνται. Οι δονήσεις είναι στην ουσία ταλαντώσεις, οι οποίες ως γνωστόν, σε μη ιδανικές συνθήκες, είναι φθίνουσες. Δηλαδή αρχικά, ο ταλαντωτής κινείται από μια μέγιστη ακραία θέση στην ακριβώς αντίθετη μερια και εναλλάξ για ένα πολύ σύντομο χρονικό διάστημα και με την πάροδο του χρόνου, το πλάτος της ταλάντωσης φθίνει μέχρι να γίνει μηδενικό και το αντικείμενο να ακινητοποιηθεί. Ένα μικρόφωνο, τοποθετημένο σε κοντινή απόσταση από το αντικείμενο, θα είχε την εξής γραφική παράσταση πίεσης σε συνάρτηση με το χρόνο στα πρώτα ms.

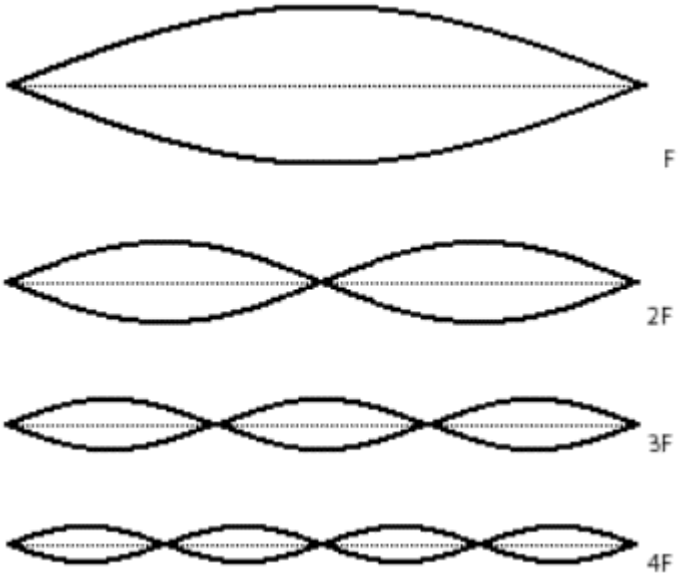


Η ημιτονοειδής ταλάντωση της πίεσης συνιστά το ηχητικό μας κύμα, με πλάτος το μέγιστο εύρος ταλάντωσης της πίεσης. Η παλμική δόνηση αυτή του αντικειμένου λέγεται και απλή αρμονική κίνηση, επειδή παρόλο που η ταχύτητα του δεν είναι σταθερή καθόλη τη διάρκεια, ο χρόνος στον οποίο πραγματοποιεί έναν επαναλαμβανόμενο κύκλο είναι σταθερός. Ο αριθμός των παλμικών δονήσεων που πραγματοποιούνται στη διάρκεια ενός δευτερολέπτου ορίζεται ως η συχνότητα της δόνησης.

Για να μπορέσει το αυτί μας όμως να αντιήφθεί αυτό τον ήχο θα πρέπει να να πληρεί κάποιες προϋποθέσεις ως προς το πλάτος της πίεσης, αλλά και της συχνότητας ταλάντωσης. Το πλάτος της πίεσης του αέρα θα πρέπει να κυμαίνεται από 10^{-9} ως και 10^{-3} φορές επί την ατμοσφαιρική πίεση. Μικρότερες τιμές δεν γίνονται αντιληπτές, ενώ μεγαλύτερες μπορεί να προκαλέσουν βλάβες στο τύμπανο. Για τη συχνότητα το εύρος είναι 20 μέχρι 20.000 κύκλους ανά δευτερόλεπτο (η μονάδα μέτρησης είναι το Hz).

2.2 Αρμονικοί ήχοι

Όσα προαναφέρθηκαν εξηγούν την απλή αρμονική κίνηση, όμως στην πραγματικότητα, τα περισσότερα αντικείμενα, όπως η χορδή της κιθάρας, ταλαντώνονται με παραπάνω από έναν τρόπο ταλάντωσης ταυτόχρονα και συνεπώς περιέχουν πολλές επί μέρους παλμικές δονήσεις σε περισσότερες από μια συχνότητες. Οι ήχοι που παράγονται με αυτόν τον τρόπο,ονομάζονται σύνθετοι τόνοι. Ο συνδυασμός των διαφορετικών συχνοτήτων και των αντίστοιχων πλατών αυτών των ήχων συνιστούν το φάσμα τους και η κάθε συχνότητα ξεχωριστά ορίζεται ως συνιστώσα. Όταν όλες οι συνιστώσες ενός ήχου αποτελούν ακέραια πολλαπλάσια μιας θεμελιώδους συχνότητας (F), τότε ο τόνος χαρακτηρίζεται ως αρμονικός και οι συνιστώσες του αρμονικές συνιστώσες.

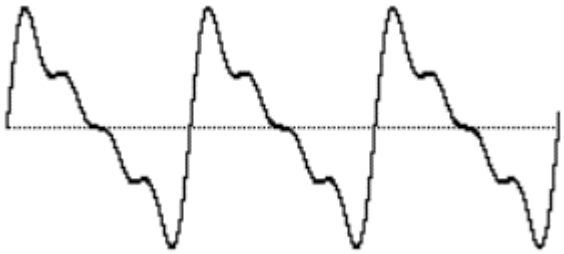


διάφοροι τρόποι που μπορεί να ταλαντωθεί μια χορδή

Το άθροισμα των συνιστωσών αυτών παραμένει ένα περιοδικό κύμα, με συχνότητα τη θεμελιώδη. Αυτό είναι ένα φαινόμενο που δεν ισχύει μόνο για τα ηχητικά κύματα, καθώς από το θεώρημα του Fourrier είναι γνωστό ότι κάθε περιοδικό κύμα μπορεί να εκφραστεί ως άθροισμα ημίτονοειδών συναρτήσεων, με συχνότητες ακέραια πολλαπλάσια μιας θεμελιώδους και διαφορετικά πλάτη και φάσεις. Η μαθηματική αναπαράσταση του μοντέλου είναι η εξής. [9]

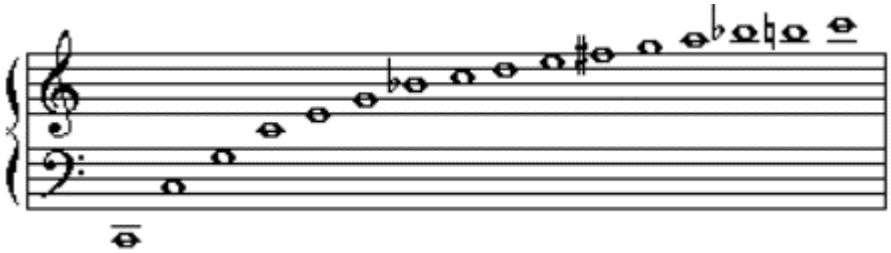
$$\boxed{g(t) = C + \sum_{k=1}^{\infty} d_k \sin(2\pi F k t + \beta_k)}$$

Οπου t είναι ο χρόνος, F η θεμελιώδης συχνότητα, dk τα διάφορα πλάτη των συνιστωσών και βκ οι αντίστοιχες φάσεις και ο C είναι πραγματικός αριθμός, που ισούται με το μέσο όρο της g(t) (Συνήθως είναι 0 για ηχητικά κύματα).



Περιοδικό ηχητικό σήμα με πολλές συνιστώσες

Γίνεται φανερό λοιπόν ότι το κύμα είναι άθροισμα συνιστωσών με συχνότητες F,2F,3F... κλπ. Μία απεικόνιση των αρμονικών συνιστωσών ενός τέτοιου τόνου με μουσικούς όρους είναι η παρακάτω.



Οι 15 πρώτες αρμονικές ενός θεμέλιου φθόγγου σε παρτιτούρα

Οι παραπάνω όροι ξεκινώντας από το χαμηλότερο στον υψηλότερο αναπαριστούν την F,2F,3F κ.ο.κ. συνιστώσα. .Οι ήχοι λοιπόν που χρησιμοποιούνται στη συμβατική μουσική ονομάζονται *μουσικοί φθόγγοι* και είναι περίπου 100. Οι μουσικοί φθόγγοι έχουν τρία κύρια χαρακτηριστικά τα οποία αντιλαμβάνεται το αυτί, την ένταση, την οξύτητα και την χροιά. Τα χαρακτηριστικά αυτά είναι σχετικά με το πλάτος, τη συχνότητα και σύνθεση των επιμέρους παλμικών δονήσεων που τους προκαλούν. Για μια συγκεκριμένη π.χ. θεμελιώδη συχνότητα το αυτί αντιλαμβάνεται το όργανο προέλευσης αυτού του αρμονικού ήχου μέσω των όρων dk, δηλαδή τα πλάτη των συνιστωσών αυτού του ήχου. Αυτός είναι ο λόγος που κάθε όργανο έχει το δικό του χαρακτηριστικό ηχόχρωμα. [10]

2.3 Μη αρμονικοί τόνοι-Θόρυβος

Αντίθετα όταν οι παλμικές δονήσεις είναι ασύμμετρες και ακανόνιστες, όπως π.χ. σε μια καμπάνα, παράγουν ήχους που δεν έχουν ορισμένο ύψος και ονομάζονται μη αρμονικοί τόνοι. Εκεί οι συχνότητες των συνιστώσων δεν αποτελούν ακέραιο πολλαπλάσιο ενός

θεμέλιου φθόγγου και ως συνέπεια, το ηχητικό κύμα δεν επαναλμβάνεται περιοδικά. Σε αυτή την περίπτωση ο ήχος δεν ακούγεται σαν ένα ενιαίο σύνολο και ειδικά όταν οι αρμονικές είναι πολλές και τελείως μη αρμονικές δεν αντιλαμβανόμαστε μια θεμελιώδη συχνότητα, σε αντίθεση με τους μουσικούς φθόγγους. Όταν ο ήχος είναι τόσο σύνθετος, που περιέχει πολλές διαφορετικές συχνότητες χωρίς κάποια προφανή μαθηματική σχέση, αντιλαμβανόμαστε τον ήχο ως *θόρυβο*.

Επομένως όλοι οι ήχοι μπορούν να είναι από οτιδήποτε από καθαροί και προβλέψιμοι τόνοι (ημιτονοειδές κύμα) μέχρι και ολοκληρωτικά τυχαίοι (λευκός θόρυβος). Οι περισσότεροι στην πραγματικότητα βρίσκονται κάπου ενδιάμεσα των 2 ακραιών αυτών καταστάσεων. Για παράδειγμα μια νότα, ή αλλιώς μουσικός φθόγγος από τρομπέτα ή πιάνο είναι αρμονικός τόνος και πλησιάζει περισσότερο καθαρό και προβλέψιμο τόνο, ενώ τα κύμβαλα των κρουστών τείνουν προς το θόρυβο.

2.4 Μέθοδοι σύνθεσης ήχου

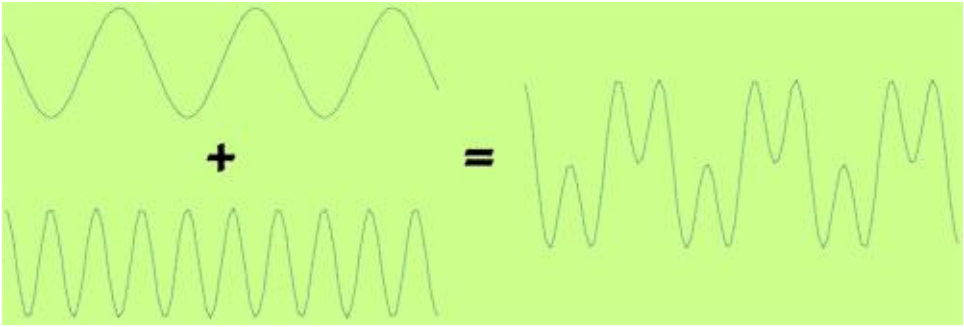
Η παρούσα εργασία αφορά την αλγοριθμική σύνθεση ηλεκτρονικών ήχων μέσω video sonification με τη χρήση φυσικά ηλεκτρονικού υπολογιστή. Με τον όρο ηλεκτρονικός ήχος περιγράφεται ο ήχος που προκύπτει από τη σύνθεση ψηφιακών ήχων με ηλεκτρονικά μέσα. Εφόσον αναλύθηκαν οι βασικότερες δομικές μονάδες ενός μουσικού κομματιού, δηλαδή οι μουσικοί φθόγγοι ή νότες και οι μη αρμονικοί τόνοι αλλά και ο θόρυβος, θα προχωρήσουμε στην απαρίθμηση των βασικότερων μεθόδων σύνθεσης ηλεκτρονικών ήχων.

2.4.1 Προσθετική σύνθεση (Additive synthesis)

Αποτελεί μία από τις πρώτες μεθόδους που χρησιμοποιήθηκαν και στην ουσία βασίζεται στο θεώρημα του Fourier το οποίο αναφέρθηκε και προηγουμένως. Δηλαδή ότι ενας σύνθετος τόνος μπορεί να δημιουργηθεί από το άθροισμα πολλών απλών τόνων, οι οποίοι είναι ημιτονοειδείς συναρτήσεις διαφόρων πλατών και φάσεων.

$$s(t) = \sum_i a_i(t) sin(2\pi f_i t + \phi_i)$$

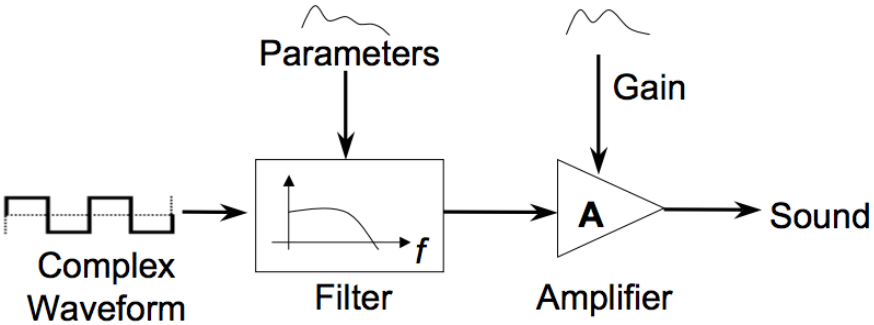
Με χρήση ψηφιακών ταλαντωτών λοιπόν ο υπολογιστής μπορεί να παράγει οποιοδήποτε σύνθετο τόνο και η ταχύτητα και ισχύς του αποτελούν το μόνο περιορισμό για το βαθμό πολυπλοκότητας της τελικής κυματομορφής. Είναι παρόλα αυτά μια τεχνική που δε χρησιμοποιείται πια καθώς θεωρείται ξεπερασμένη.



2 ηχητικά περιοδικά κύματα προσθέτονται δημιουργώντας ένα σύνθετο τόνο

2.4.2 Αφαιρετικη σύνθεση (Subtractive Synthesis)

Η τεχνική αυτή είναι στην ουσία ακριβώς αντίθετη από την προσθετική σύνθεση. Ως αρχική κυματομορφή χρησιμοποιείται ένα ηχητικό σήμα πλούσιο σε αρμονικές συνιστώσες, όπως ο λευκός θόρυβος, ή ένας τετραγωνικός παλμός, και στη συνέχεια αυτή φιλτράρεται έτσι ώστε να παραμείνουν οι επιθυμητές συνιστώσες. Η απλότητα της τεχνικής την κατέστησε ως ιδανική επιλογή για την κατασκεύη των αναλογικών synthesizers Moog, στη δεκαετία του 60 και 70.

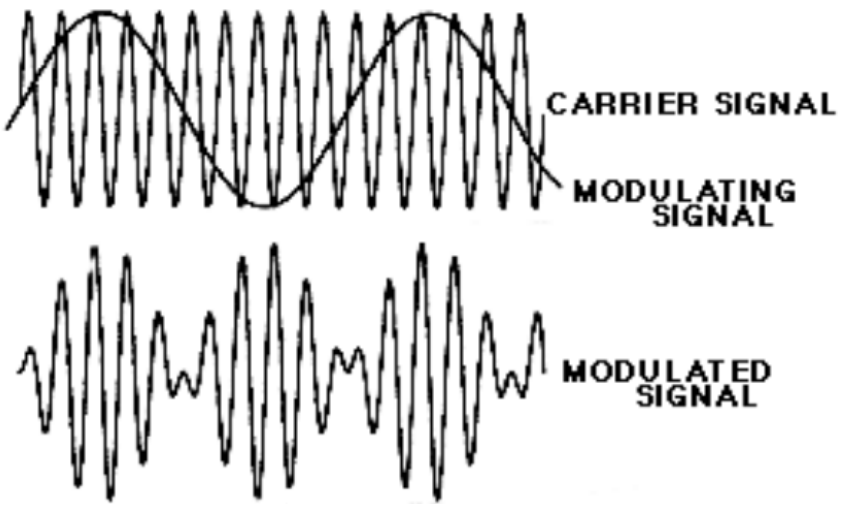


2.4.3 Σύνθεση με μικροδομές (Granular σύνθεσεις)

Όπως γίνεται φανερό από τον τίτλο, εδώ γίνεται χρήση μικροδομών, δηλαδή ηχητικά σήματα διάρκεις από 10-100 ms συνήθως. Αυτά τα σήματα αποτελούνται από μια περιβάλλουσα και μια κυματομορφή. Το τελικό σήμα διαμορφώνεται από το συνδυασμό πολλών τέτοιων μικροδομών στη σειρά, με διάφορες περιβάλλουσες και δίαρκειες κάθε φορά, δημιουργόντας έτσι διαφορετικά ηχοχρώματα αλλά και θεμελιώδη συχνότητα. Η μεγαλύτερη της χρησιμότητα έγκειται λοιπόν στην επεξεργασία ψηφιακού ήχου, όταν θέλουμε να διαμορφώσουμε κατάλληλα τη διάρκεια του ή τη βασική συχνότητα ξεχωριστά

2.4.4 Σύνθεση με διαμόρφωση πλάτους και (Aplitude-ring modulation synthesis)

Εδώ σε αντίθεση με τις προηγούμενες τεχνικές, προκειται για μη-γραμμική μεθόδο σύνθεσης καθώς δημιουργούνται συχνότητες στο τελικό σήμα οι οποίες δεν παράχθηκαν από τους ταλαντωτές . Το ηχητικό σήμα της πήγης, ο φορέας (carrier signal), πολλαπλασιάζεται με ένα άλλο σήμα, το σήμα διαμόρφωσης (modulation signal). Στη διαμόρφωση πλάτους το σήμα διαμόρφωσης έχει μη αρνητικές τιμές (πλάτος σήματος από 0 εώς 1) ενώ στο ring modulation έχει και αρνητικές (-1 εώς 1). Το σήμα διαμόρφωσης δηλαδή χρησιμοποιείται ως περιβάλλουσα για το φορέα.



2.4.5 Σύνθεση πίνακα κυματομορφής (Wavetable synthesis)

Στην ουσία πρόκειται για μια περιορισμένη μορφή προσθετικής σύνθεσης σε πραγματικό χρόνο (real time), η οποία χρησιμοποιείται σε αρκετά synthesizers από τα τέλη της δεκαετίς του 70 και μετά (π.χ. Korg). Η βασική λειτουργία έγκειται στη αποθηκευμένων περιοδικών κυματομορφων οι οποίες έχουν με διάρκεια ενα κύκλο. Αυτές αναπαράγονται ύστερα αφού δεχτούν κάποια διαμόρφωση στο πλάτος αλλά και στην ταχύτητα. Με αυτόν τον τρόπο μπορεί να αναπαραχθεί οποιοσδήποτε θεμέλιος φθόγγος. Η κύρια διαφορά με την προσθετική σύνθεση, είναι ότι δε γίνοντια υπολογισμοί πρόσθεσης ταλαντωτών, αλλά υπάρχει πίνακας κυματομορφών που αποτελούν "τμήματα" από όλα τα ηχοχρώματα.

2.4.6 Σύνθεση φυσικών μοντέλων(Physical Modeling Synthesis)

Είναι τεχνικές που μιμούνται τον ήχο από διάφορα μουσικά όργανα, περιγράφοντας με αριθμητικούς τύπους και διαφορικές εξισώσεις τον τρόπο με τον οποίο είναι κατασκευασμένο ένα μουσικό όργανο ή προσομοιώνοντας συστήματα που συμπεριφέρονται σαν ένα αληθινό μουσικό όργανο [7]. Πρόκειται για μια από τις πολυπλοκότερες αλλά και πιο ενστικτώδεις προσεγγίσεις καθώς η κύρια ιδέα είναι να προσομοιωθεί η απόκριση ενός συστήματος, που είναι το μουσικό όργανο, σε ένα ερέθισμα που μπορεί να προέρχεται από π.χ. δόνηση, αέρα ή χορδή.

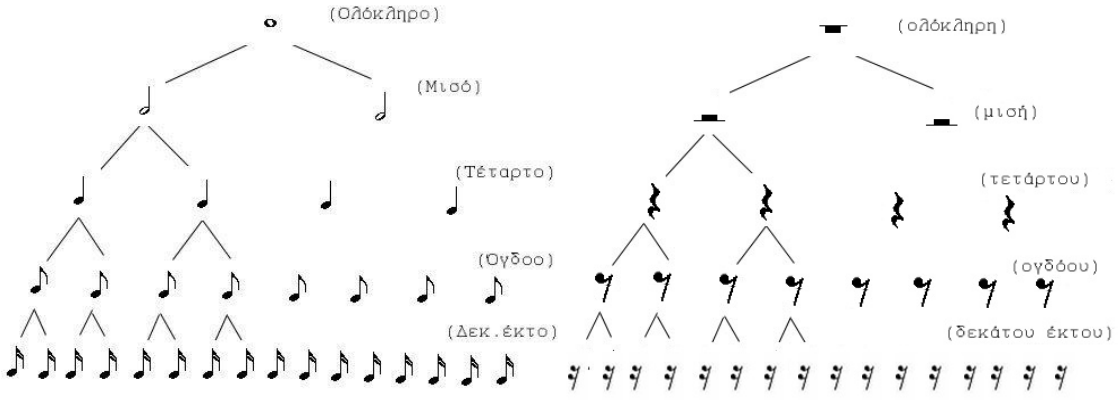
2.5 Ρυθμικά στοιχεία ενός μουσικού κομματιού

2.5.1 Μουσικές αξίες

Για να δημιουργηθεί ένα μουσικό κομμάτι χρειάζονται 2 βασικά στοιχεία: η μελωδία και ο ρυθμός. Η μελωδία αποτελείται από τους μουσικούς φθόγγους, στους οποίους αναφερθήκαμε σε προηγούμενη ενότητα. Ο ρυθμός ταυτίζεται με την οργάνωση του χρόνου. Η οργάνωση αυτή βρίσκεται σε απόλυτη εξάρτηση με τους τονισμούς και τη διάρκεια των ήχων. Η διάρκεια των μουσικών φθόγγων αποδίδεται με διάφορα σύμβολα τα οποία ονομάζονται αξίες διαρκείας των φθόγγων, ή πιο απλά αξίες των φθόγγων. Σε κάθε μία από τις αξίες των φθόγγων αντιστοιχεί και μια ισοδύναμη αξία παύσης. Όμως, οι αξίες αυτές δεν είναι απόλυτες, αλλά σχετικές με επιπλέον στοιχεία που δίνονται, τις υποδείξεις ρυθμικής αγωγής. Οι κυριότερες αξίες, αρχίζοντας από τις μεγαλύτερες και καταλήγοντας στις μικρότερες, είναι οι εξής.[10]

Φθογγόσημο	Ονομασία	Διάρκεια	Παύση
	Ολόκληρο	4 χρόνους	
	Μισό	2 χρόνους	
	Τέταρτο	1 χρόνο	
	Όγδοο	1/2 χρόνο	
	Δέκατο Έκτο	1/4 χρόνου	

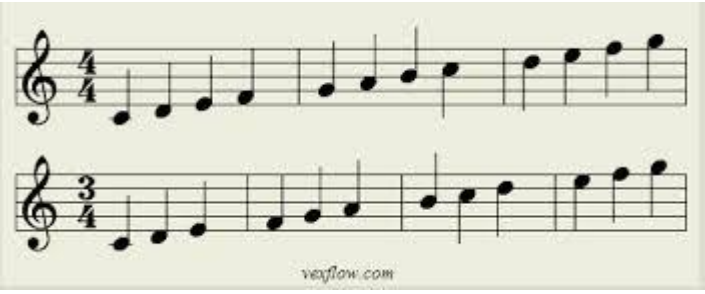
Κάθε μία από τις αξίες των φθόγγων ή των παύσεων με 2 από τις αμέσως μικρότερες της.



2.5.2 Μέτρο-tempo

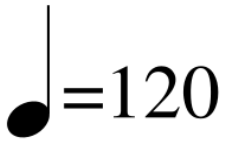
Για να μπορέσει να απεικονιστεί η μουσική αξία των φθόγγων χρησιμοποιείται το μέτρο, το οποίο είναι μια απόλυτα ισόποση διαίρεση του χρόνου. Κάτι αντίστοιχο με αυτό που συμβαίνει στην καθημερινή ζωή, όπου ο χρόνος διαιρείται σε ώρες, λεπτά, δευτερόλεπτα κλπ. Στη μουσική η διαίρεση αυτή γίνεται με μια υποθετική ροή ισόχρονων κτύπων, που ονομάζονται παλμοί. Η έννοια του παλμού είναι αυτό που θέλουμε να εκφράσουμε όταν κτυπάμε ρυθμικά το πόδι μας στο άκουσμα κάποιας μουσικής. Οι παλμοί μπορεί να είναι γρήγοροι ή αργοί, αλλά σίγουρα είναι ισόχρονοι, έτσι ώστε ο χρόνος να διαιρείται σε ισόποσα διαστήματα.

Για τη γραφική αναπαράσταση των παλμών του μέτρου χρησιμοποιείται κάποια από τις αξίες φθόγγων που αναφέρθηκαν προηγουμένως. Οι αξίες αυτές αντιπροσωπεύονται από αριθμούς. Το ολόκληρο με το αριθμό 1, το μισό με το 2, το τέταρτο με το 4 κ.ο.κ.. Ο αριθμός των παλμών που περιέχεται μέσα σε κάθε μέτρο και ο αριθμός που αντιπροσωπεύει το είδος της αξίας που διαλέχτηκε για να αναπαραστήσει το κάθε παλμό τοποθετούνται μαζί, με μορφή κλάσματος, στην αρχή του μουσικού κομματιού.



Το κλάσμα αυτό, που έχει σαν αριθμητή τον αριθμό των παλμών και σαν παρονομαστή την αξία των παλμών, ονομάζεται χρόνος του μουσικού κομματιού. Το πάνω παράδειγμα έχει 4 παλμούς ανά μέτρο και ο χρόνος του είναι 4 τέταρτα. Στο δεύτερο παράδειγμα έχουμε 3 παλμούς ανά μέτρο και χρόνο 3 τέταρτα.

Η ταχύτητα με την οποία γίνεται η επανάληψη των παλμών ενός μέτρου ονομάζεται ρυθμική αγωγή(tempo). Το όργανο που είναι υπεύθυνο για για τον ορισμό αυτής της ταχύτητας είναι ο μετρονόμος, ένα όργανο το οποίο με τους απόλυτα ισόχρονους κτύπους του είναι κατάλληλο για να δείχνει τον αριθμό των παλμών σε κάθε λεπτό(bpm, beats per minute). π.χ.



Σε αυτή την περίπτωση κάθε χτύπος του μετρονόμου αντιστοιχεί σε ένα τέταρτο, και σε κάθε λεπτό υπάρχουν συνολικά 120 τέταρτα.

Ένας μετρονόμος. Κάθε χτύπος αντιστοιχεί στη μουσική αξία που ορίζουμε

3. Τι είναι Sonification;

Ο ελληνικός αντίστοιχος όρος είναι ηχοποίηση, αλλά δεν είναι δόκιμος, οπότε στη συνέχεια της παρούσας εργασίας θα χρησιμοποιείται ο αγγλικός όρος, δηλαδή Sonification. Πρόκειται για μια διαδικασία μετάδοσης πληροφορίας μέσω ήχων, που δεν περιέχουν ομιλία, χρησιμοποιώντας την ίδια αρχή με την οπτικοποίηση (visualization), όπου εκεί τα δεδομένα αντιστοιχίζονται με παραμέτρους που αντιλαμβάνεται το μάτι, όπως σχήμα, χρώμα κλπ.. Η οπτικοποίηση χρησιμοποιείται εδώ και αρκετό καιρό ευρέως. Ένα απλό παράδειγμα είναι τα γραφήματα, όπου μπορεί κάποιος εύκολα να αντιληφθεί τη διακύμανση πολλών τιμών σε συνάρτηση με κάποια άλλη παράμετρο (π.χ. χρόνο) ευκολότερα και γρηγορότερα. Το Sonification από την άλλη κάνει χρήση παραμέτρων αντιληπτών απο το αυτί, όπως συχνότητα, ένταση.

Παρόλο που ως έννοια αναπτύχθηκε τα τελευταία χρόνια, στην ουσία η βασική αρχή χρησιμοποιούνταν ευρέως από τον επιστημονικό κόσμο εν άγνοια του. Ένα από τα παλαιότερα παραδείγματα είναι ο μετρητής Geiger, ο οποίος εφευρέθηκε το 1908, και χρησιμοποιείται ακόμα και σήμερα, πληροφορώντας το χρήστη για τα επίπεδα ραδιενέργειας σε ένα περιβάλλον με τη χρήση ήχων. Είναι μια σημαντικότατη επιστήμη που χρησιμοποιείται για την ανίχνευση μοτίβων σε μεγάλο όγκο δεδομένων, πράγμα που θα ήταν δυσκολότερα αντιληπτό από οπτική παρατήρηση.

3.1 Ηχοποίηση εικόνας (Image Sonification)

Στην περίπτωση του Image Sonification, τα δεδομένα μας είναι εικόνες και είναι ένας από τους πιο ενδιαφέροντες και πρακτικούς τομείς του κλάδου, καθώς συνιστά εργαλείο για ανθρώπους με προβλήματα στην όραση. Με κατάλληλη αντιστοίχιση των παραμέτρων της εικόνας με τις παραμέτρους του ήχου, αυτά τα άτομα μπορούν να αντιληφθούν μια εικόνα. Για να γίνει εφικτό όμως κάτι τέτοιο θα πρέπει τα άτομα αυτά να έχουν μια σχετική εκπαίδευση, ως προς το τι σημαίνει ο κάθε ήχος.

Γενικότερα, Image Sonification χρησιμοποιείται και σε άλλες πρακτικές περιπτώσεις, όπως ο παλμογράφος στα νοσοκομεία. Εκεί ο χτύπος της καρδιάς καταγράφεται πρώτα από όλα οπτικά και ύστερα, ακούμε ένα υψίσυχνο ήχο για κάθε χτύπο της καρδιάς του ασθενή. Με αυτόν τον τρόπο γίνεται εφικτή η απομακρυσμένη παρακολούθηση της καταστασής του. Όλη η διαδικασία μπορεί να προσομοιωθεί από ένα σύστημα με είσοδο και έξοδο. Η αντιστοίχιση των παραμέτρων της εικόνας, όπως σχήμα, ύψος, χρώμα με τις παραμέτρους του ηχητικού αποτελέσματος όπως συχνότητα, ένταση, ρυθμός είναι το σύστημα μας. Η είσοδος του είναι η εκάστοτε εικόνα, και η έξοδος το ηχητικό αποτέλεσμα.

3.2 Αλγοριθμική μουσική μέσω Sonification

Εκτός από πρακτικές εφαρμογές όμως, Sonification μπορεί να χρησιμοποιηθεί και για καλλιτεχνικούς σκοπούς. Αυτό είναι το θέμα που πραγματεύεται η παρούσα εργασία, σύνθεση μουσικής μέσω **video sonification,** δηλαδή η δημιουργία ενός μουσικού κομματιού κάνοντας χρήση ενός οποιουδήποτε βίντεο ως πηγή δεδομένων. Ο κλάδος αυτός του sonification βρίσκεται ακόμα σε πρώιμο στάδιο, με πολλούς μουσικούς, αλλά και μηχανικούς να πειραματίζονται πάνω σε αυτόν.

3.3 Video Sonification και αλγοριθμική μουσική

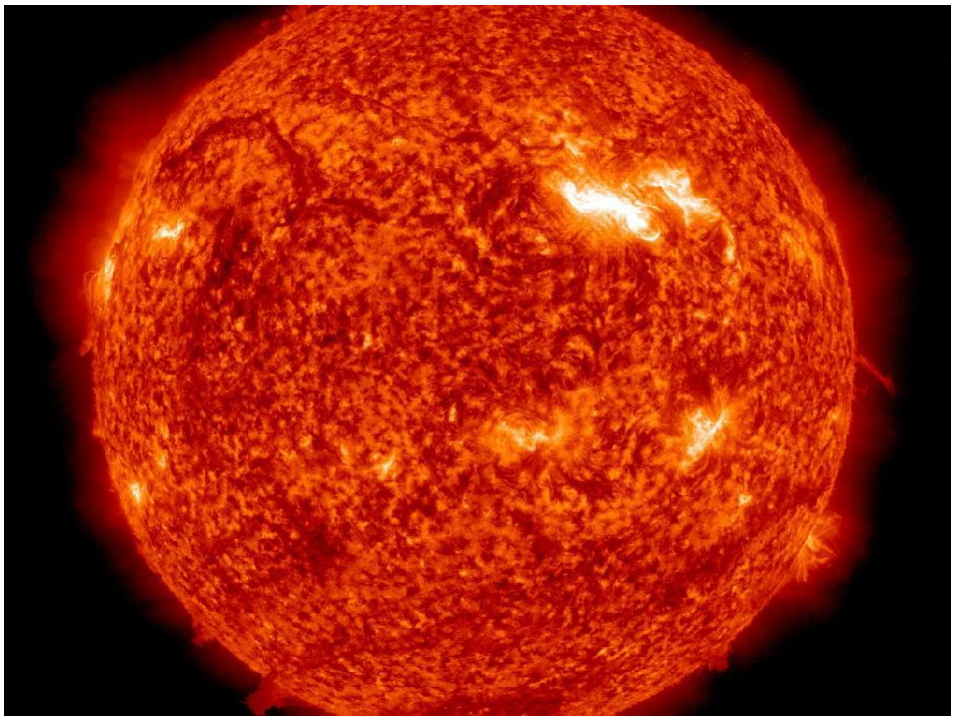
Το video sonification εχει διτή χρησιμότητα στην σύνθεση αλγοριθμικής μουσικής

- Χρησιμοποιώντας τις κατάλληλες παραμέτρους από το βίντεο, εξάγουμε δεδομένα τα οποία μπορούν να χρησιμοποιηθούν στον αλγόριθμο μας. Τα δεδομένα αυτά έχουν μια συσχέτιση μεταξύ τους, καθώς είναι τιμές που προκύπτουν από ένα συγκεκριμένο βίντεο. Με αυτόν τον τρόπο, θεωρητικά, το μουσικό αποτέλεσμα ακολουθεί τη συνοχή.
- Το οπτικοακουστικό θέαμα που μπορεί να προσφέρει στο θεατή/ακροατή, συνδυάζοντας ένα όμορφο αισθητικά οπτικό θέαμα, με μια εξίσου αξιόλογη μουσική, η οποία βρίσκεται σε αρμονία με το βίντεο. Φυσικά τα soundtrack των ταινιών, σε συνδυασμό με τη σκηνή αποτελούν ένα από τα χαρακτηριστικότερα παραδείγματα ενός τέτοιου θεάματος. Η διαφορά έγκειται, στο ότι σε αυτήν την περίπτωση η μουσική παράγεται με πιο αφαιρετικό τρόπο, χωρίς να υπάρχει κάποια άμεση σύνδεση με το οπτικό υλικό.

3.3.1 Video Sonification με οπτικό υλικό από το διάστημα

Ένας από τους πρωτοπόρους και σπουδαιότερους επιστήμονες/ συνθέτες οι οποίοι ασολούνται με το ζήτημα είναι ο Robert Alexander, διδάκτορας του πανεπιστήμιου University of Michigan, με θέμα της διατριβής του, το "Desing Science". Αυτός ασχολήθηκε με οπτικό υλικό της NASA από το διάστημα, παράγοντας αξιοθαύμαστα αποτελέσματα, τα οποία μάλιστα εκτός από καλλιτεχνική αξία είχαν και πρακτική χρησιμότητα. Συγκεκριμένα, η NASA, έχει καταγράψει οπτικό υλικό από τον ήλιο, όπου ακόμα και μη ορατές ακτινοβολίες, όπως υπεριώδη και υπέρυθρη έχουν αναπαρασταθεί με "ψεδο-χρώματα". Με χρήση αυτών των δεδομένων, τα οποία έχουν καταγραφεί σε διάστημα 43 χρόνων, ο Robert Alexander ηχοποίησε με συγκεκριμένο τρόπο την περιστροφή του ηλίου αλλά και άλλες συγκεκριμένες δραστηριότητες που συμβαίνουν περιοδικά με διακριτούς ήχους. Με αυτόν τον τρόπο οι επιστήμονες μπορούν να κατανοήσουν καλύτερα ορισμένα σημαντικά φαινόμενα όπως ηλιακές καταιγίδες.

Αυτή είναι μια αρκετά ιδιαίτερη περίπτωση Sonification, καθώς το αποτέλεσμα έχει επιστημονικό, αλλά και καλλιτεχνικό ενδιαφέρον. Το έργο του έχει λάβει εξίσου εξαιρετικές κριτικές κι από τους 2 τομείς, δηλαδή τον επιστημονικό, αλλά και τον καλλιτεχνικό.



Robert Alexander

3.3.2 Τρόπος χρήσης του Video Sonification στην παρούσα εργασία

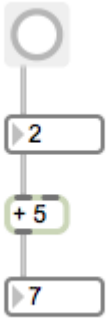
- κάνοντας χρήση ενός υψηλού επιπέδου προγραμματιστικού περιβάλλοντος, το MAX/MSP.
- Με αντιστοίχιση διάφορων παραμέτρων του βίντεο με τις επιθυμητές ηχητικές παραμέτρους.
- με χρήση μουσικών κανόνων έγινε εφικτό ένα είδος στοιχειώδους αλγοριθμικής σύνθεσης, σε αντιστοίχιση με το βίντεο προς επεξεργασία.

4. Συνοπτική παρουσίαση του MAX/MSP

Για την πλήρη κατανόηση της διαδικασίας που ακολουθήθηκε για τον αλγόριθμου, τον οποίο πραγματεύεται η παρούσα εργασία, κρίνεται σκόπιμο να γίνει παρουσίαση του τρόπου λειτουργίας του MAX/MSP. Για την παρουσίαση αυτή χρησιμοποιούνται αρκετά παραδείγματα μέσα από τα οποία φαίνονται οι δυνατότητες, αλλά και οι περιορισμοί του προγραμματιστικού αυτού εργαλείου.

4.1 Βασικά στοιχεία και λειτουργίες του MAX/MSP

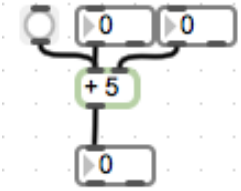
Το MAX/MSP είναι μια γραφική προγραμματιστική γλώσσα, ή μάλλον ορθότερα προγραμματιστικό περιβάλλον, του οποίου ο χειρισμός διαφέρει σημαντικά από τα περισσότερα δημοφιλή περιβάλλοντα, καθώς δεν περιλαμβάνει πολλές γραμμές κώδικα. Είναι αντικείμενοστρεφής και η λειτουργία του βασίζεται στη λογική ενός διαγράμματος ροής και είναι αρκετα διαδραστική ως προς το χρήστη. Χρησιμοποιείται κυρίως από συνθέτες, μουσικούς εκτελεστές αλλά και σχεδιαστές λογισμικού και ερευνητές. Κάθε πρόγραμμα που δημιουργείται στο MAX για να επιτελέσει μια συγκεκριμένη λειτουργία,ονομάζεται patcher. Για να κατανοηθεί καλύτερα η λειτουργία του ακολουθεί παράδειγμα ενός patcher με πολύ απλή λειτουργία



Το patcher του παραδείγματος, πραγματοποιεί πρόσθεση 2 αριθμών με χρήση 4 συνολικά αντικειμένων. Όπως φαίνεται από το σχήμα, παρατηρείται η δομή διαγράμματος ροής, όπου η πληροφορία μεταδίδεται από πάνω προς τα κάτω, με το αντικείμενο πιο ψηλά στην ιεραρχία να είναι το "κουμπί" (button), το οποίο μόλις πατηθεί πραγματοποιείται η επιθυμητή λειτουργία στέλνοντας "bang"(χτύπημα) στο επόμενο στη σειρά αντικείμενο. Τα αντικείμενα που περιέχουν μόνο αριθμούς ονομάζονται number boxes και το περιεχόμενο τους μπορεί εύκολα να μεταβληθεί με τη χρήση του ποντικιού. Το αντικείμενο που πραγματοποιεί την πράξη είναι ένα object και ανάλογα με την ονομασία που πληκτρολογούμε, μπορεί να εκτελέσει αναρίθμητες διαφορετικές λειτουργίες. Σε αυτή την περίπτωση προσθέτει τον αριθμό 5 σε οποιαδήποτε αριθμό δεχτεί στην αριστερή του είσοδο και εξάγει το αποτέλεσμα από τη μοναδική έξοδο που διαθέτει.

Καταλαβαίνουμε λοιπόν τη σημασία της ιεραρχίας στο MAX/MSP όπως και επίσης τη δυνατότητα που παρέχει στο χρήστη να παράγει το αποτέλεσμα με ελεγχόμενο τρόπο,

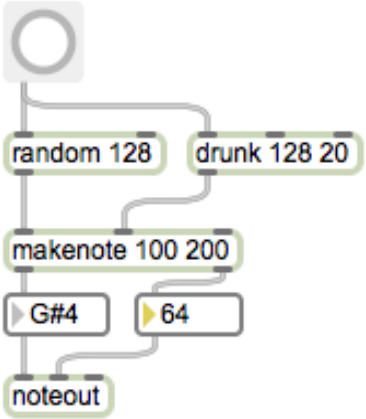
δίνοντας του μεγάλη ελευθερία και άνεση. Σημαντικό πλεονέκτημα αποτελεί επίσης το γεγονός ότι για κάθε αντικείμενο, ο χρήστης μπορεί να ανατρέξει στο σχετικό εγχειρίδιο το οποίο εξηγεί τη λειτουργία του, αλλά και επιπλέον προσφέρει παραδείγματα με εφαρμογές που κάνουν χρήση τα συγκεκριμένα αντικείμενα. Τα περισσότερα objects έχουν 2 εισόδους, την αριστερή ως κύρια, η οποία μόλις δεχτεί είσοδο δίνει σήμα ώστε να εξαχθεί αποτέλεσμα από μία ή παραπάνω εξόδους του object. Οι υπόλοιπες είσοδοι, συνήθως χρησιμοποιούνται για αποθήκευση δεδομένων που θα χρησιμοποιηθούν αργότερα από το object για την εξαγωγή του αποτελέσματος, γι' αυτό και θεωρούνται ανενεργές.



Σε αυτό το patcher, οποιαδήποτε αλλαγή του δεξιού number box, το αποτέλεσμα παραμένει 0. Με αλλαγή του αριστερού number box, ή με "bang" από το button, τα οποία και τα 2 συνδέονται με την αριστερή είσοδο του object που επιτελεί την πρόσθεση, θα εμφανιστεί αυτόματα το αποτέλεσμα της πρόσθεσης των 2 αριθμών, στο number box που είναι συνδεδεμένο με την έξοδο του. Ο αριθμός 5 που έχει οριστεί σαν πρωταρχικός προσθετέος, αντικαθίσταται αυτόματα από δεξί number box, μόλις μεταβληθεί.

4.2 Πρωτόκολλο MIDI (*Musical Instrument Digital Interface)*

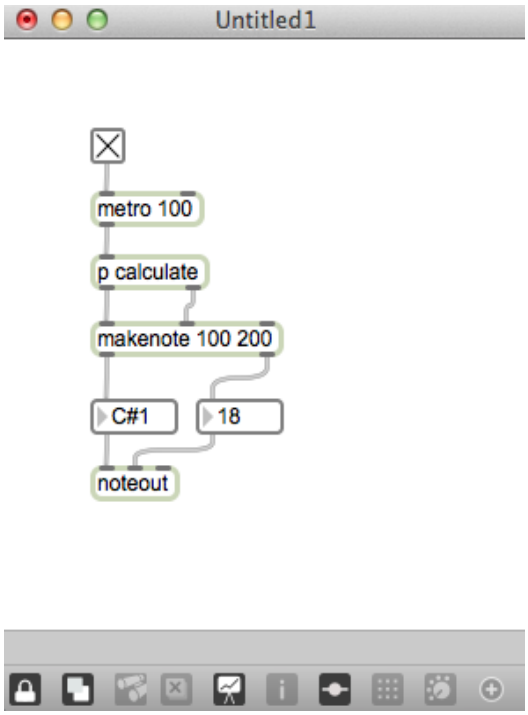
Στις εφαρμογές που έχουν να κάνουν με παραγωγή μουσικών ήχων, το MAX κάνει χρήση του πρωτόκολλου επικοινωνίας MIDI. Το MIDI (αποσκοπεί στην επικοινωνία και τον συγχρονισμό μεταξύ ηλεκτρονικών μουσικών οργάνων υπολογιστών και άλλων ηλεκτρονικών συσκευών. Δε μεταδίδει ηχητικό σήμα, αλλά μηνύματα που περιέχουν πληροφορίες σχετικά με το τονικό ύψος και την ένταση μιας νότας. Τα MIDI μηνύματα, όταν πρόκειται για τονικό ύψος (pitch) δέχονται τιμές από 0-127 και κάθε ένας από τους 128 αριθμούς αντιστοιχεί σε ένα μουσικό φθόγγο ξεκινώντας από τη νότα C-1 για το 0 κ.ο.κ.. Για κάθε ένα από τους 12 μουσικούς φθόγγους χρησιμοποιούνται τα εξής σύμβολα C, C#, D, D#, E, F, F#, G, G#, A, A#, B και δίπλα από τους φθόγγους αντιστοιχεί το νούμερο της οκτάβας στην οποία βρίσκεται ο φθόγγος. Ακολουθεί ένα ακόμα παράδειγμα, ενός πιο εκλεπτυσμένου patcher του MAX, όπου γίνεται χρήση MIDI μηνυμάτων, αλλά και τυχαιότητας με σχετικά απλό τρόπο.



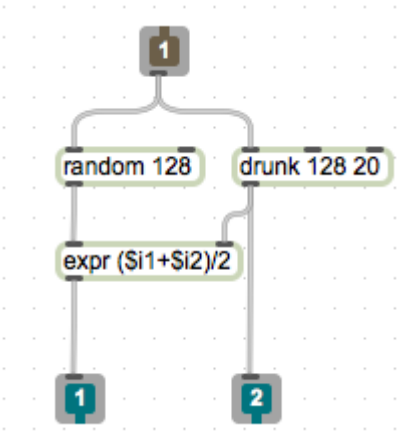
Κάνοντας "κλικ" στο button το αντικείμενο "random 128" υπολογίζει έναν τυχαίο αριθμό από το 0-127 και τον στέλνει στην έξοδο του. Το αντικείμενο "drunk 128 20" υπολογίζει κι αυτό αρχικά έναν τυχαίο αριθμό στο ίδιο εύρος με το random, αλλά η διαφορά τους έγκειται ότι στο ότι στα επόμενα "κλικ", οι επόμενοι τυχαίοι αριθμοί θα έχουν μεταξύ τους διακύμανση +/-20. Οι έξοδοι των **random** και **drunk** χρησιμεύουν για τον υπολογισμό του τονικού ύψους και της έντασης της νότας που θα παιχτεί με χρήση του επόμενου αντικειμένου, του **makenote**. Αυτό χρησιμεύει για να στέλνει τα 2 αυτά μηνύματα ομαδοποιημένα στο **noteout**, στέλνοντας ύστερα μηνύματα μηδενικής έντασης. Το **noteout**, το τελευταίο αντικείμενο του διαγράμματος χρησιμεύει τέλος για την παραγωγή των midi ήχων με αποτέλεσμα να ακούγεται η συγκεκριμένη νότα από το πρόγραμμα οδήγησης ήχου (audio driver) που χρησιμοποιεί ο υπολογιστής.

4.3 Δομημένος προγραμματισμός στο MAX/MSP

Όπως και με τις περισσότερες γλώσσες προγραμματισμού υπάρχει η δυνατότητα για δομημένο προγραμματισμό, έχοντας υποπρογράμματα, μέσα στο πρόγραμμα μας, τα οποία ονομάζονται subpatchers. Με αυτόν τον τρόπο μπορούμε να δημιουργήσουμε ένα patcher με διάφορες σύνθετες λειτουργίες, το οποίο θα έχει καθαρή και απλούστερη δομή. Τα subpatchers έχουν τη μορφή απλών object στα οποία ο χρήστης επιλέγει μια επιθυμητή ονομασία, που να έχει κάποιο νόημα ως προς την λειτουργία που επιτελεί. Με διπλό κλικ στο αντικείμενο αυτό ανοίγει το subpatcher.



Στο άνωθεν patcher, παρατηρούμε ότι το τελευταίο σκέλος είναι ίδιο με του προηγούμενου. Το υψηλότερο στην ιεραρχία αντικείμενο λέγεται toggle (διακόπτης) και στέλνει τιμή 0 ή 1 στο **metro 100**. Το **metro 100** λειτουργεί στην ουσία όπως ένας μετρονόμος και βρίσκεται εν ενεργεία, όταν ο διακόπτης είναι κλειστός, στέλνοντας "bang" ανά 100 ms στο "p calculate". Με διπλό κλικ στο "p calculate" μπορούμε να εξετάσουμε τη λειτουργία του.



Γνωρίζουμε τα random, drunk από το προηγούμενο παράδειγμα. Στη συνέχεια πραγματοποιείται μία πράξη, το αποτέλεσμα των 2 αριθμών προστίθεται και διαιρείται με το 2. Με αυτόν τον τρόπο παίρνουμε το μέσο όρο τους. Το καφέ αντικείμενο με τον αριθμό 1 παραμετροποιεί την είσοδο του subpatcher, και αντίστοιχα τα 1,2 με μπλε χρώμα τις εξόδους. Σε αυτό το σημείο προσθέτουμε ότι στο MAX αντικείμενα με παραπάνω από μια εξόδους, έχουν ιεραρχία από δεξιά προς τα αριστερά. Επομένως το makenote πρώτα δέχεται την πληροφορία για την ένταση (velocity) και μετά για το φθόγγο(pitch).

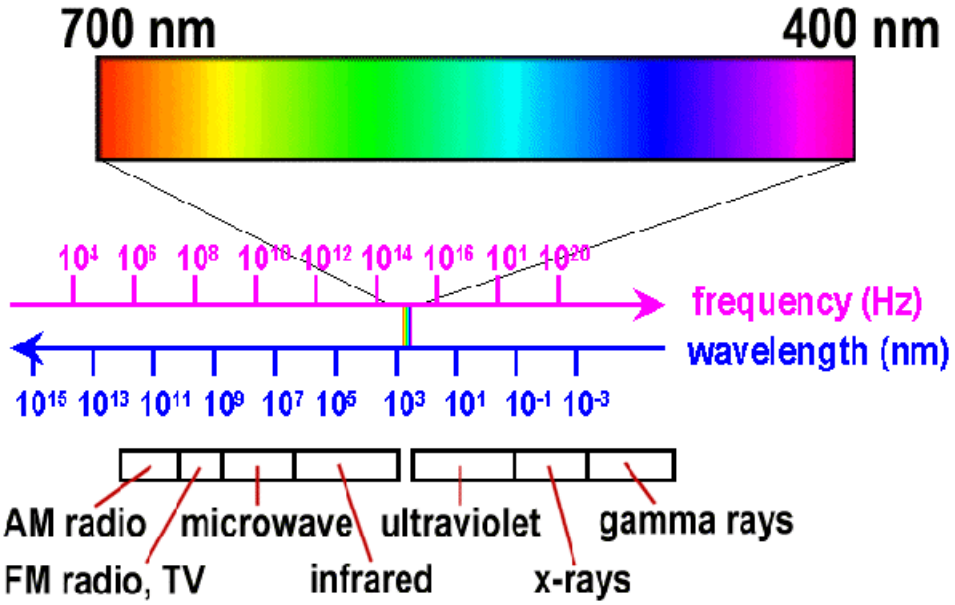
Θα μπορούσαμε να κατασκευάσουμε ένα patcher που να επιτελεί την ίδια λειτουργία χωρίς τη χρήση του **p calculate**, αλλά με αυτόν τον τρόπο έχουμε ένα αποτέλεσμα με πιο συμπαγή και κατανοητή δομή και φυσικά φιλικότερο ως προς το χρήστη. Υπάρχουν 2 είδη **subpatcher**, αυτά που αποθηκεύονται στη βιβλιοθήκη και μπορούν να επαναχρησιμοποιηθούν σε διαφορετικά patchers με απλή ανάκληση του ονόματος τους και αυτά που δημιουργούνται αποκλειστικά για ένα patcher. Η διαφορά τους έγκειται στο ότι τα πρώτα δημιουργούνται σε ξεχωριστό αρχείο και αποθηκεύονται ενώ τα τελευταία μέσα στο κύριο patcher και η ονομασία τους έχει το πρόθεμα "p " (p subpatcher_name).

5. Χρώμα και μοντέλα προσομοίωσης του από Η/Υ

Ο αλγόριθμος που χρησιμοποιήθηκε στο πρόγραμμα το οποίο πραγματεύεται η εργασία βασίζεται σε αρκετά μεγάλο βαθμό στην επεξεργασία του χρώματος ενός βίντεο. Επομένως κρίνεται σκόπιμο να γίνει αναφορά στις ιδιότητες, αλλά και τον τρόπο με τον οποίο το αντιλαμβανόμαστε.

5.1 Τι είναι Χρώμα

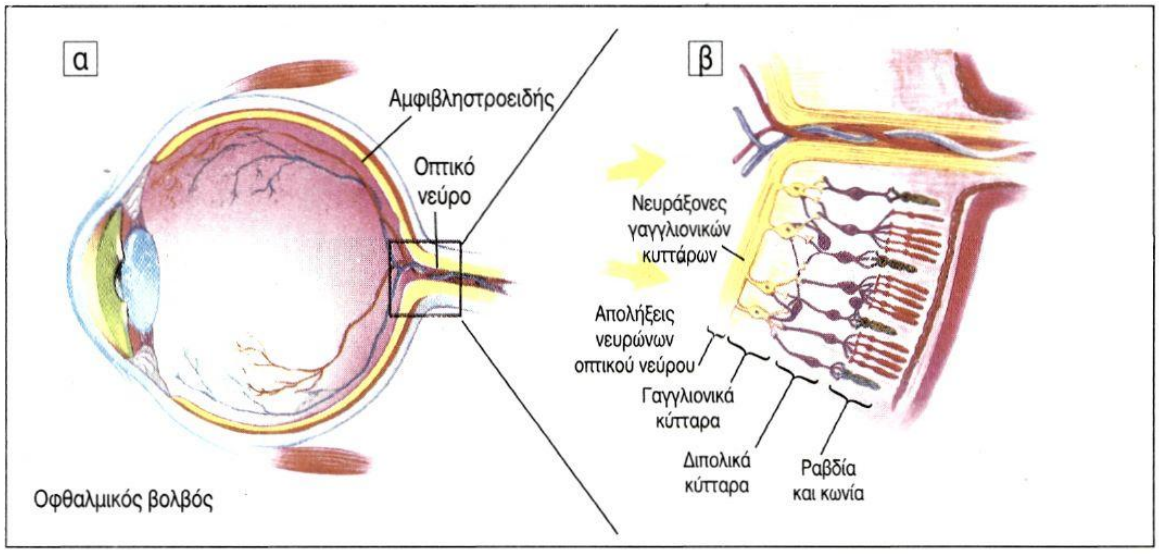
Φως ονομάζεται η ηλεκτρομαγνητική ακτινοβολία η οποία ανιχνεύεται από το ανθρώπινο μάτι και αντιλαμβάνεται ως αίσθηση αυτής. Η ζώνη του ορατού φωτός είναι μία στενή ζώνη του φάσματος της ηλεκτρομαγνητικής ακτινοβολίας, στην οποία είναι ευαίσθητο το αισθητήριο της όρασης των ζωντανών οργανισμών. Το ορατό φάσμα του φωτός αποτελείται από διάφορα μήκη κύματος που όλα μαζί μας δίνουν το λευκό φως. Αν περάσουμε τις ακτίνες του ήλιου από ένα πρίσμα θα δούμε να αναλύεται σε όλα τα χρώματα της ίριδας, όπως απέδειξε και ο Sir Isaac Newton. Το μήκος κύματος αυτών των ακτινοβολιών κυμαίνεται από 400 μέχρι 700nm περίπου. [1]



5.2 Πως αντιλαμβανόμαστε τα χρώματα;

Ο τρόπος με τον οποίο αντιλαμβανόμαστε το χρώμα είναι μια υποκειμενική κατάσταση. Υπάρχει μια αναλογία ως προς την αντίληψη του ήχου και του χρώματος. Όπως οι περισσότεροι ήχοι είναι σύνθετοι τόνοι, που αποτελούνται από πολλές συνιστώσες, κάτι παρόμοιο ισχύει με το χρώμα. Σπάνια συναντάται στη φύση μονοχρωματικό φως. Συνήθως όταν βλέπουμε ένα χρώμα, αυτό είναι μίγμα από πολλά μήκη κύματος σε διαφορετικές αναλογίες. Το **χρώμα** είναι ουσιαστικά μια αίσθηση που δημιουργείται στον εγκέφαλο από μέρος της αλληλουχίας των ηλεκτρικών ώσεων που φθάνουν σε αυτόν μέσω του οπτικού νεύρου. Οι ηλεκτρικές ώσεις που αφορούν πληροφορία χρώματος προέρχονται από εξειδικευμένους φωτοευαίσθητους υποδοχείς του ματιού, τα κονία, που αντιδρούν το καθένα στην ανίχνευση φωτός συγκεκριμένου εύρους μήκους κύματος.

5.3 Ραβδία και Κονία



Υπολογίζεται ότι στον αμφιβληστροειδή καθενός από μας, ο οποίος έχει διάμετρο 2,5 περίπου εκατοστών, υπάρχουν γύρω στα 132. 000. 000 εξειδικευμένα νευρικά κύτταρα από τα οποία 125. 000. 000 έχουν σχήμα ραβδιού διαμέτρου 0,002 mm και λέγονται **ραβδία.** Τα υπόλοιπα 7.000.000 λέγονται **κονία**. Καθένα τους έχει σχήμα που θυμίζει κάπως το χωνί και μέγιστη διάμετρο 0, 006 mm.[11] Τα ραβδία και τα κονία είναι οι φωτοαισθητήρες. Τα ραβδία είναι υπεύθυνα για την αντίληψη του αμυδρού φωτός ενώ τα κονία (ή κωνικά κύτταρα) για την αντίληψη των χρωμάτων.

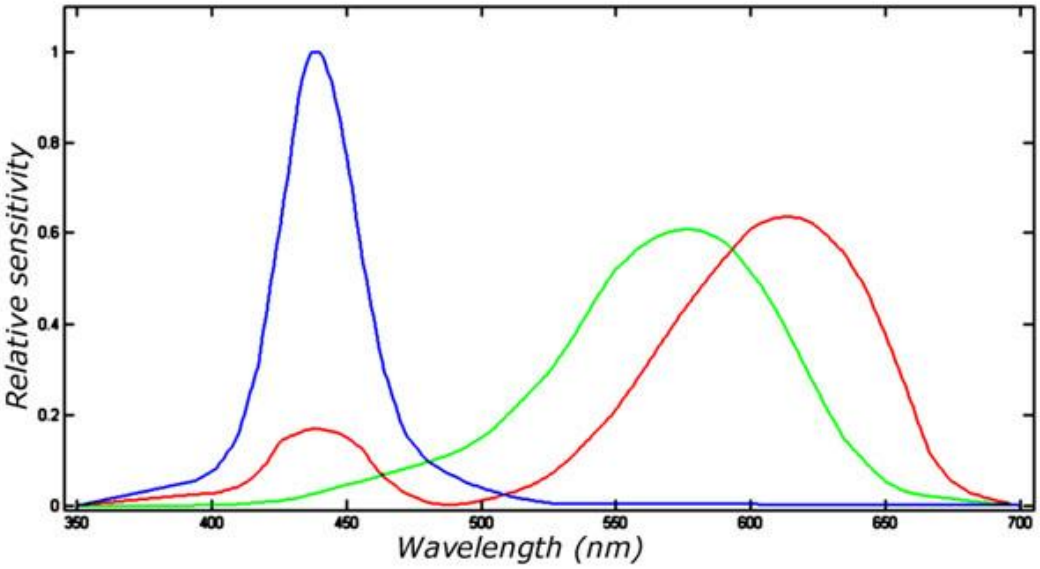
Η αναγνώριση των χρωμάτων λοιπόν γίνεται στον αμφιβληστροειδή μόνο από τα κονία τα οποία μετατρέπουν την ηλεκτρομαγνητική ακτινοβολία σε πληροφορία. Αυτό γίνεται απορροφώντας το φως με 3 διαφορετικές κατηγορίες κωνικών κυττάρων. Η διαφορά τους

έγκειται στην ικανότητα να μετατρέπουν ηλεκτρομαγνητική ακτινοβολία συγκεκριμένου μήκους κύματος σε πληροφορία για να την επεξεργαστεί ο εγκέφαλος. Αυτές είναι οι εξής.

- S-κονία: Είναι ευαίσθητα σε φωτόνια μικρού μήκους κύματος και παρουσιάζουν μέγιστη ευαισθησία σε μήκος κύματος περίπου 420nm. Στην ουσία είναι ευαίσθητα στο μπλε φως.

- M-κονία: Είναι ευαίσθητα σε φωτόνια μεσαίου μήκους κύματος και παρουσιάζουν μέγιστη ευαισθησία σε μήκος κύματος περίπου 530nm και χειρίζονται το πράσινο φως.

- L-κονία: Είναι ευαίσθητα σε φωτόνια μεγάλου μήκους κύματος και παρουσιάζουν μέγιστη ευαισθησία σε μήκος κύματος περίπου 560nm, που είναι το κυρίαρχο μήκος κύματος του κόκκινου φωτός.



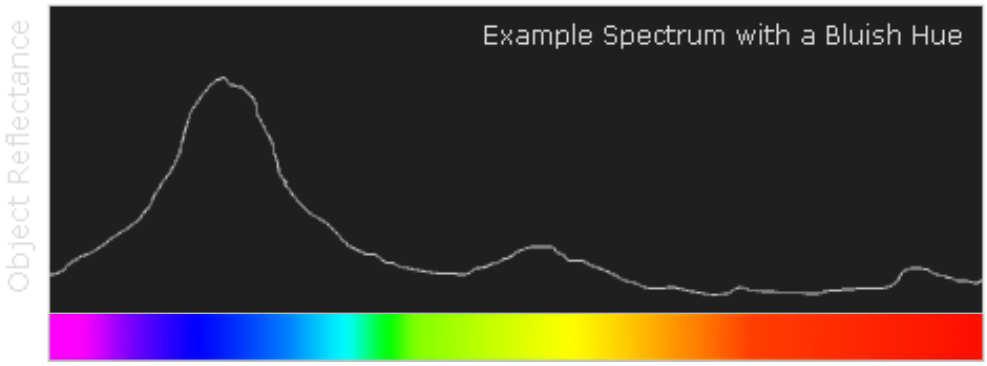
Στο διάγραμμα φαίνεται η απόκριση των S,M, L κονιών με μπλε πράσινο και κόκκινο χρώμα αντίστοιχα

5.4 Χαρακτηριστικά αντίληψης των χρωμάτων

Όπως στους αρμονικούς τόνους υπάρχει μια θεμελιώδη συχνότητα την οποία διακρίνει το αυτί, έτσι και στα χρώματα διακρίνεται το κυρίαρχο μήκος κύματος. Ακολουθούν τα κύρια χαρακτηριστικά με τα οποία περιγράφουμε το χρώμα, στα οποία δεν υπάρχει ακριβής ελληνική μετάφραση που να περιγράφει τις έννοιες με πληρότητα, γι'αυτό θα γίνεται χρήση των αγγλικών όρων.

- **Hue (χροιά)**: Έχει την ίδια σχέση που έχει η νότα με τη θεμελιώδη συχνότητα του ήχου που την παράγει. Είναι δηλαδή η ψυχοσωματική αντίδραση στο κυρίαρχο μήκος κύματος ενός χρώματος και είναι ουσιαστικά το στοιχείο που καθορίζει την κύρια ονομασία που θα δώσουμε σε ένα χρώμα. Υπάρχουν, όμως, περισσότερες από 100 αποχρώσεις χρωμάτων. Το κόκκινο π.χ. δεν είναι ένα κόκκινο. Υπάρχει κόκκινο με απόχρωση πορτοκαλί, κόκκινο με απόχρωση μωβ κλπ. Οι αποχρώσεις ορίζονται από τη θέση τους πάνω στην περιφέρεια του χρωματικού κύκλου.

Χρωματικός κύκλος



Παράδειγμα φασματικής ανάλυσης χρώματος με μπλε απόχρωση(Hue)

- Brightness (λάμψη): Είναι μια ιδιότητα που χρησιμοποιείται για να περιγράψει πηγές φωτός. Από ψυχοσωματική έννοια, ανταποκρίνεται είναι η αντιληπτή *ένταση* ενός αυτό-φωτειζόμενου αντικειμένου. Από χοντρική χρωματική άποψη πρόκειται για μια έννοια που περιγράφει την ποσότητα μαύρου που έχει αναμειχθεί με το εκάστοτε χρώμα. Λειτουργεί συνάρτησει του μήκος κύματος, επομένως χρώματα με την ίδια χροιά(hue) παρουσιάζουν διαφορετική λάμψη (brightness) στο ανθρώπινο μάτι.

- Lightness (φωτεινότητα): Είναι έννοια σχετική με το brightness, όπως υποδεικνύει και η ονομασία και αναφέρεται στην αντιληπτή αχρωματική ένταση ενός αντανακλώντος αντικειμένου. Ένας ακόμα ορισμός που εξηγεί την έννοια καλύτερα είναι ότι αποτελεί τη λάμψη(brightness) ενός οποιοδήποτε αντικειμένου σε σχέση με τη λάμψη ενός αντικειμένου που φαίνεται ως λευκό υπό τις ίδιες ακριβώς συνθήκες.

- Colorfulness (πληρότητα χρώματος): Μια έννοια που περιγράφει το περιεχόμενο χροιάς (hue content) μιάς οποιαδήποτε χρωματικής διέγερσης. Με την αύξηση της φωτεινότητας (luminance) αυξάνεται με ανάλογο ρυθμό.

- Chroma: Δεν μπορεί να χρησιμοποιηθεί ελληνικός όρος με ταυτόσημη έννοια. Είναι έννοια σχετική με την πληρότητα χρώματος(colourfulness) όπως ακριβώς και η φωτεινότητα(lightness) για τη λάμψη(brightness). Δηλαδή είναι η πληρότητα χρώματος μιας οποιαδήποτε οπτικής διέγερσης σε σχέση με την πληρότητα χρώματος μιας οπτικής διέγερσης που αντιλαμβάνεται ως λευκή στις ίδιες συνθήκες.

- Saturation (κορεσμός): Είναι η ψυχοσωματική αντίληψη της "καθαρότητας" ενός χρώματος, δηλαδή το πόσο λευκό ή γκρι είναι αναμειγμένα με αυτό

5.5 Χρωματικά μοντέλα(Color models) και Χρωματικοί χώροι (color spaces)

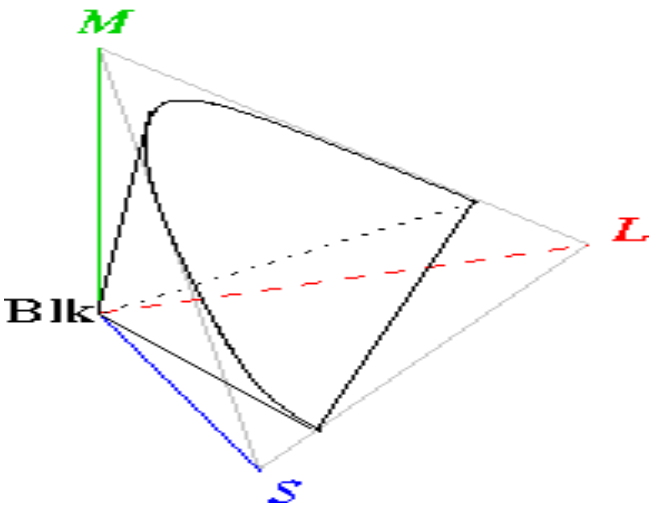
Για την περιγραφή και χρήση των χρωμάτων έχουν αναπτυχθεί διάφορες προσεγγίσεις περιγραφής τους, τα Χρωματικά μοντέλα. Χρωματικό μοντέλο ονομάζεται η περιγραφή ενός συνόλου χρωμάτων με συνδυασμούς λίγων βασικών, συνήθως 3 ή 4. **Βασικά** (ή **πρωτογενή**) χρώματα είναι αυτά τα οποία όταν συνδυαστούν μεταξύ τους μπορούν να παράγουν όλα τους δυνατούς χρωματισμούς. Τα χρώματα που προκύπτουν από τον συνδυασμό των βασικών λέγονται **δευτερογενή**. Γίνεται κατανοητό, λοιπόν, ότι χρησιμοποιούν μέθοδο παρόμοια με το ανθρώπινο μάτι.

Οι χρωματικοί χώροι είναι οι βάσεις της διαχείρισης του χρώματος. Περιλαμβάνουν ομάδες διαφορετικών χρωμάτων τα οποία περιγράφονται από χρωματικά μοντέλα. Δηλαδή με λίγα λόγια, όταν ένα χρωματικό μοντέλο συσχετιστεί με κάποια ακριβή περιγραφή, η οποία ερμηνεύει τις συνιστώσες του(components), τότε οι ομάδες χρωμάτων που έχουμε ως αποτέλεσμα ονομάζονται χρωματικοί χώροι. Οι ομάδες αυτές μπορεί να βασίζονται σε ορισμένα βασικά χρώματα από τον συνδυασμό των οποίων σε συγκεκριμένες ποσότητες προκύπτουν τα υπόλοιπα ή να βασίζονται σε μεταβλητές τις οποίες χρησιμοποιούν χρωματικά μοντέλα.

Η τριχρωματική φύση της χρωματικής αντίληψης στην όραση επιτρέπει σε οποιαδήποτε χρώμα να απεικονιστεί ως ένας συνδυασμός των 3 αυτών βασικών χρωμάτων. Η τριχρωματικότητα της όρασης είναι γραμμική κι αυτό φαίνεται κι από την εξίσωση που περιγράφει τη διαδικασία, όπου για μια οπτική διέγερση S έχουμε την χρωματική απεικόνιση Cs.

$Cs(S) = C1(l1) + C2(l2) + C3(l3)$

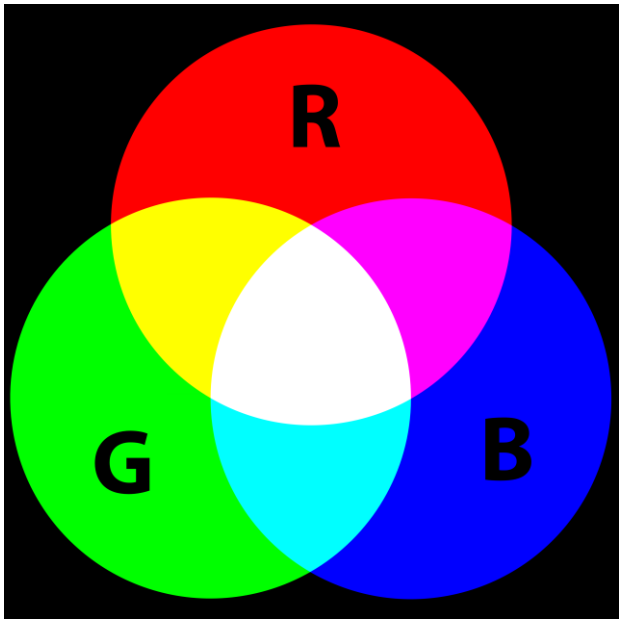
C1,C2,C3 είναι σταθεροί αριθμοί ενώ I1,I2,I3 είναι τα βασικά χρώματα μπλε, πράσινο και κόκκινο. Ένα διάγραμμα το οποίο απεικονίζει τον χρωματικό χώρο που αντιλαμβάνεται το μάτι είναι το εξής.



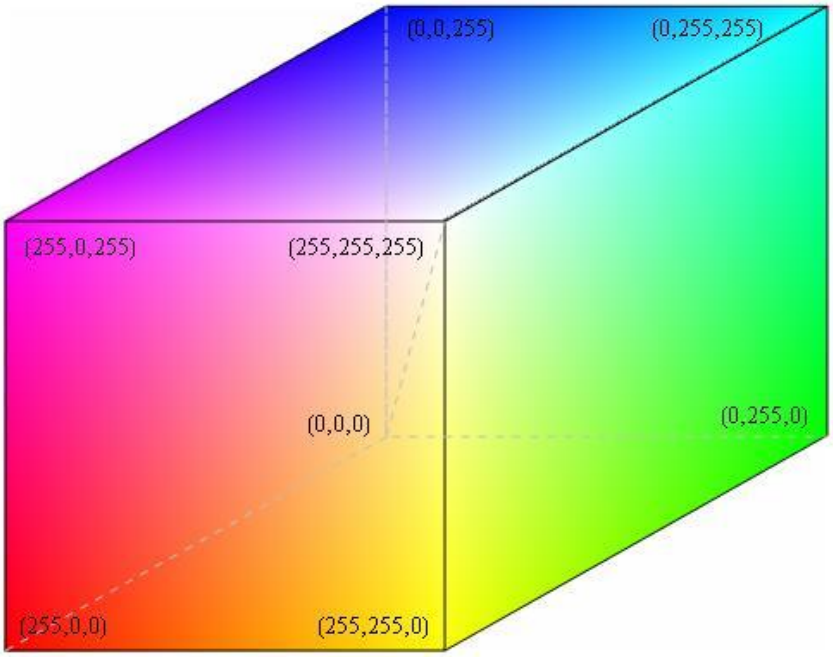
Τα S,M,L συμβολίζουν τις συνιστώσες που αντιστοιχούν στο μπλε, πράσινο και κόκκινο αντίστοιχα. Για το μαύρο χρώμα ισχύει (S,M,L)=(0,0,0) (Blk). Επομένως είναι φανερή η αντιστοιχία με ένα τρισδιάστατο Ευκλείδιο χώρο, όπου οι άξονες x,y,z είναι οι οπτικές διεγέρσεις που δέχεται το μάτι για τα τους 3 τύπους κονιών που διαθέτει. [Βικιπαίδεια]

Με παρόμοιο τρόπο γίνεται η περιγραφή διάφορων χρωματικών μοντέλων (και σε επέκταση χρωματικών χώρων) που έχουν δημιουργηθεί, για να γίνει δυνατή η περιγραφή των χρωμάτων με μαθηματική μορφή, κατάλληλη για την επεξεργασία τους από ψηφιακά μέσα. Στους ηλεκτρονικούς υπολογιστές, η εικόνα αναλύεται σε αρκετά ομοιόμορφα στοιχεία, τα οποία ονομάζονται **εικονοστοιχεία(pixel)**. Τα χρωματικά μοντέλα ουσιαστικά αφορούν την κωδικοποίηση χρώματος ενός εικονοστοιχείου Για τους σκοπούς της παρούσας διπλωματικής εργασίας θα αναλύσουμε 2 από τους βασικότερους χρωματικούς χώρους.

5.5.1 Χρωματικό μοντέλο και Χρωματικός χώρος RGB



Το χρωματικό μοντέλο αυτό είναι ακρωνύμιο των λέξεων Red, Green, Blue (Κόκκινο, Πράσινο, Μπλε), δηλαδή των 3 βασικών χρωμάτων που αποτελούν τις 3 συνιστώσες-μεταβλητές του μοντέλου και είναι αποτέλεσμα προσθετικής μίξης. Ο χρωματικός χώρος RGB καλύπτει ένα αρκετά μεγάλο μέρος του ορατού φάσματος καθώς χρησιμοποιεί ως βασικά χρώματα που αντιλαμβάνεται ευκολότερα το ανθρώπινο μάτι. Οι τρεις μεταβλητές (R,G,B) έχουν τιμές 0-255, οι οποίες συμβολίζουν την ένταση τους κάθε χρώματος στη μίξη. Επομένως για τη ψηφιακή αναπαράσταση τους χρειαζόμαστε 8 δυαδικά ψηφία (2^8=256) και κάθε χρώμα μπορεί να παρασταθεί από μία τριάδα αριθμών. Για (0,0,0) με τις ελάχιστες δυνατές τιμές έχουμε μαύρο χρώμα ενώ για (255,255,255) έχουμε λευκό.

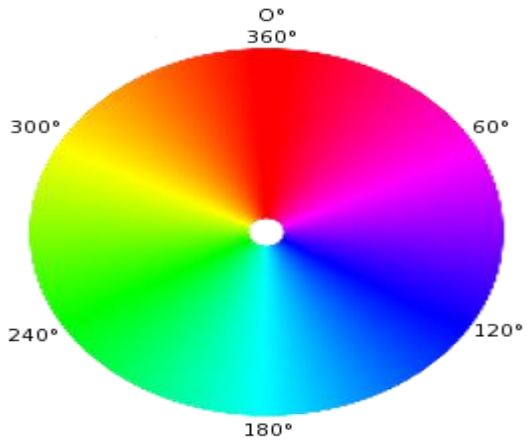


Γραφική αναπαράσταση του Χρωματικού χώρου RGB

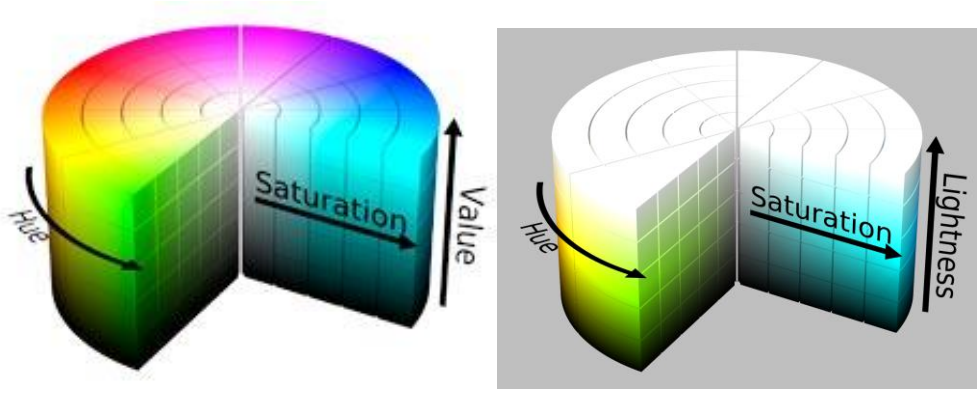
5.5.2 Χρωματικό μοντέλο και Χρωματικός χώρος HSV και HSL

Τα χρωματικά μοντέλα HSV και HSL είναι πιο ανθρώπινα με την έννοια ότι οι μεταβλητές τους μοιάζουν να απαντούν σε ερωτήσεις όπως τι χρώμα είναι, αν είναι έντονο, σκούρο ή ανοιχτό. Αυτό γιατί οι μεταβλητές εδώ είναι Hue (απόχρωση, χροιά), Saturation (κορεσμός) και Value, όπου εδώ έχει τον ορισμό της λάμψης(Brightness) που δόθηκε προηγουμένως και Lightness (φωτεινότητα) για HSV και HSL αντίστοιχα. Πρόκειται στην ουσία για μη-γραμμικό μετασχηματισμό των συνιστωσών του RGB μοντέλου σε πολικές συντεταγμένες.

Η χροιά έχει ίδια έννοια και στα 2 μοντέλα και είναι η θέση στο χρωματικό κύκλο, έχοντας τιμές 0°-360°.



Ο Κορεσμός(Saturation) , η λάμψη(Value) και η φωτεινότητα(Lightness) παίρνουν τιμές από 0-100%.



Γραφική αναπαράσταση Χρωματικών χώρων HSV και HSL

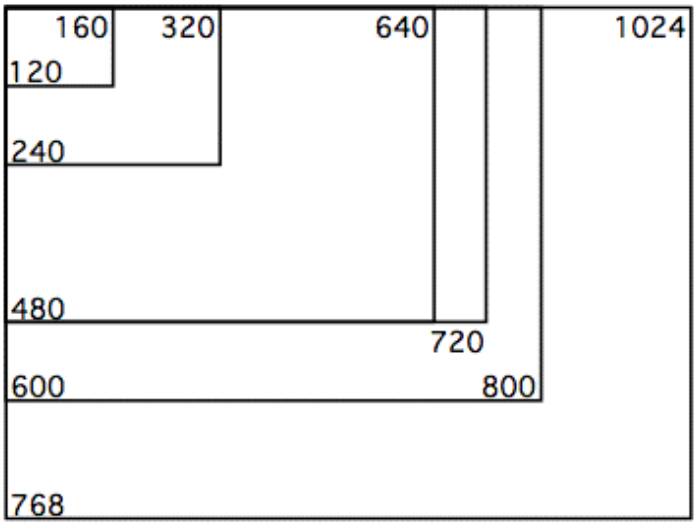
Τα μοντέλα αυτά ξεκίνησαν να χρησιμοποιούνται για εφαρμογές σε γραφικά Η/Υ και πια γίνεται χρήση τους σε λογισμικά επεξεργασίας και ανάλυσης εικόνας αλλά και υπολογιστική όραση. Για την επεξεργασία του βίντεο στον αλγόριθμο, χρησιμοποιήθηκε το μοντέλο HSL.

5.6 Επεξεργασία βίντεο στο MAX/MSP

Η επεξεργασία δεδομένων βίντεο γίνεται με εργαλεία που ανήκουν σε ένα πρόσθετο βοηθητικό πρόγραμμα του MAX/MSP, το Jitter. Για να γίνει κατανοητή η μεθοδολογία που χρησιμοποιήθηκε ακολουθεί επεξήγηση του τρόπου διαχείρισης βίντεο από το Jitter.

Κάθε βίντεο αποτελείται από έναν αριθμό διαδοχικών καρέ, συνήθως 24 καρέ ανά δευτερόλεπτο, τα οποία είναι ουσιαστικά εικόνες, αν τα εξετάσουμε ξεχωριστά. Αποσαφηνίζοντας τη διαδικασία επεξεργασίας μιας εικόνας, λοιπόν, μπορεί κατ'επέκταση να

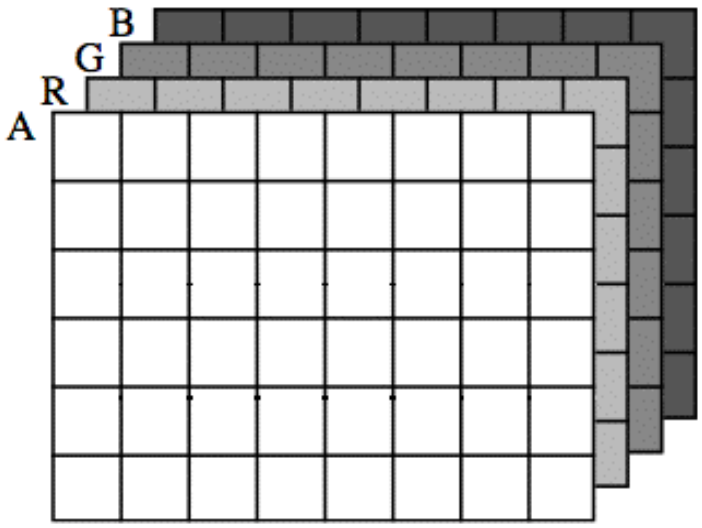
γίνει κατανοητή η επεξεργασία ενός βίντεο. Μια εικόνα απεικονίζεται στον υπολογιστή μέσω εικονοστοιχείων(pixel), για τα οποία έχει ήδη γίνει αναφορά, και των οποίων ο αριθμός ποικίλει ανάλογα με την ανάλυση. Το κάθε εικονοστοιχείο(pixel) είναι μια μικρή μονάδα που αναπαριστά ένα χρώμα σε ένα συγκεκριμένο σημείο της οθόνης. Στην οθόνη του υπολογιστή η σύνηθης ανάλυση είναι 1024x768, δηλαδή 1024 εικονοστοιχεία(pixel) κατά πλάτος και 768 κατά ύψος. Όσο μεγαλύτερα είναι αυτά τα νούμερα, τόσο καλύτερη ανάλυση έχουμε σε μια εικόνα.



Επομένως για τη ψηφιακή απεικόνιση της εικόνας, χρειαζόμαστε ένα δισδιάστατο πίνακα, με μέγεθος ανάλογο με την επιθυμητή ανάλυση. Όπως αναφερθηκε προηγουμένως, υπάρχουν διάφοροι τρόποι με τους οποίους απεικονίζεται το χρώμα στους υπολογιστές, τα λεγόμενα χρωματικά μοντέλα. Το χρωματικό μοντέλο που χρησιμοποιείται πρωταρχικά στο jitter είναι το RGB, δηλαδή κάθε χρώμα μπορεί να αναπαρασταθεί με 3 τιμές για το κόκκινο, πράσινο και μπλε αντίστοιχα (Red, Green, Blue). Επομένως τα στοιχεία του πίνακα αυτού, τα οποία αντιστοιχούν στο κάθε εικονοστοιχείο της εικόνας και είναι ουσιαστικά τετραδιάστατα διανύσματα, γιατί χρειάζονται 4 αριθμοί για να έχουμε πληροφορία για το χρώμα και την διαφάνεια/αδιαφάνεια(alpha channel) ενός εικονοστοιχείου.

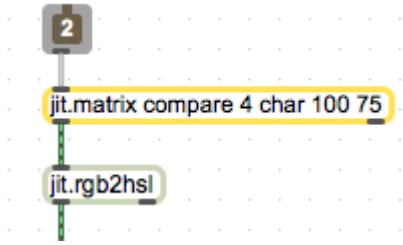
A: 255 R: 218 G: 111 B: 218	A: 255 R: 218 G: 111 B: 218	A: 254 R: 218 G: 112 B: 217	etc.
A: 255 R: 218 G: 111 B: 217	A: 255 R: 218 G: 111 B: 217	A: 254 R: 218 G: 112 B: 217	etc.
etc.	etc.	etc.	etc.

Η απεικόνιση αυτή ονομάζεται ARGB και στην ουσία χρησιμοποιούνται 4 πίνακες ίδιου μεγέθους για την αναπαράσταση μιας εικόνας.



Οι τιμές βρίσκονται στην έκταση 0-255 και άρα χρειαζόμαστε 8 bits αποθηκευτικού χώρου για τον κάθε αριθμό (2^8=256 διαφορετικές τιμές). Για τις τιμές ARGB, λοιπόν, χρησιμοποιείται ο γνωστός τύπος μεταβλητής char, ο οποίος συνήθως χρησιμοποιείται για απεικόνιση αλφαβητικών χαρακτήρων, αλλά και συμβόλων, όμως στο Jitter, αρκεί να γνωρίζει ότι χρειάζεται να δεσμεύσει 8 bit στη μνήμη για την αποθήκευση του. Σε περίπτωση, βέβαια, μονοχρωματικής (ασπρόμαυρης) εικόνας μας αρκεί μονάχα μία τιμή για κάθε εικονοστοιχείο, άρα γίνεται χρήση ενός απλού δισδιάστατου πίνακα.

Το Jitter είναι μια επιπρόσθετη εργαλειοθήκη για το MAX/MSP, της οποίας η χρήση γίνεται στο ίδιο περιβάλλον με τα υπόλοιπα αντικείμενα του MAX. Η διαφορά έγκειται στο ότι τα αντικείμενα του jitter έχουν το πρόθεμα "jit".



Το πρώτο αντικείμενο, το οποίο ονομάζεται **jit.matrix** είναι το βασικότερο εργαλείο του Jitter. Εκεί αποθηκεύεται η δισδιάστατη μήτρα για μια εικόνα ή ένα καρέ ενός βίντεο αντίστοιχα. Συνήθως χρησιμοποιούμαι ονομασίες για τον κάθε πίνακα έτσι ώστε να μπορεί να χρησιμοποιηθεί σε διάφορα σημεία με απλή ανάκληση του ονόματος του και χωρίς να δεσμεύσει μνήμη για καινούριο πίνακα το Jitter. Ο αριθμός 4 υποδεικνύει ότι χρησιμοποιούνται 4 τιμές για κάθε στοιχείο, στην περίπτωση έγχρωμης εικόνας δηλαδή. Στη περίπτωση μονοχρωματικής εικόνας θα είχαμε τον αριθμό 1. Το char τέλος είναι ο τύπος μεταβλητής που χρησιμοποιείται για το κάθε εικονοστοιχείο, δεσμεύοντας 8 bit για κάθε τιμή τα οποία μας αρκούν για τις λειτουργίες που θα εκτελεστούν. Οι τελευταίοι 2 αριθμοί υποδεικνύουν τον αριθμό των εικονοστοιχείων που χρησιμοποιείται κατά πλάτος και κατά ύψος αντίστοιχα. Συνολικά λοιπόν για την ανάλυση της συγκεκριμένης εικόνας έχουμε 100x75=7500 εικονοστοιχεία. Το επόμενο αντικείμενο, **jit.rgb2hsl**, πραγματοποιεί μετατροπή των δεδομένων RGB στο χρωματικό μοντέλο HSL.

6. Επεξήγηση αλγορίθμου

Η αλγοριθμική σύνθεση που πραγματεύεται η παρούσα εργασία, όπως αναφέρθηκε προηγουμένως, κάνει χρήση του Video Sonification. Πρόκειται για έναν αλγόριθμο, ο οποίος για διαφορετικά βίντεο, παράγει μουσική που είναι συνυφασμένη κατά κάποιο τρόπο με το περιεχόμενο του. Ακόμα, όμως και με χρήση του ίδιου οπτικού υλικού, μπορούν να παρατηρηθούν διαφορές στο ηχητικό αποτέλεσμα, καθώς έχει εισαχθεί το στοιχείο της τυχαιότητας.

Οι παράμετροι που χρησιμοποιούνται από το κάθε καρέ αφορούν το χρώμα και είναι οι Hue(χροιά), Saturation(κορεσμός), Lightness(φωτεινότητα). Γίνεται αντιληπτό, επομένως, πως για την επεξεργασία της εικόνας γίνεται χρήση του χρωματικού μοντέλου HSL. Παρόλο που το jitter αποθηκεύει τα δεδομένα σε RGB μορφή αρχικά, όπως φάνηκε από το προηγούμενο παράδειγμα, η μετατροπή σε HSL είναι αρκετά απλή.

Ο λόγος που χρησιμοποιήθηκε το μοντέλο HSL, είναι ότι οι συνιστώσες του μπορούν να συνδεθούν πιο άμεσα με έννοιες όπως χρώμα, φωτεινότητα, καθαρότητα χρώματος. Σε αντίθεση με το RGB, όπου π.χ. το ποσοστό του κόκκινου χρώματος σε μια εικόνα δεν έχει κάποια απτή σημασία. Επομένως, με αυτόν τον τρόπο έγινε ευκολότερη η συσχέτιση τους με μουσικούς όρους όπως μουσικό φθόγγο(pitch), ένταση (volume), αλλά και ρυθμό(tempo).

Αρχικά ανιχνεύεται ο μέσος ρυθμός μεταβολής του βίντεο από καρέ σε καρέ και αυτή η τιμή αντιστοιχίζεται με την κατάλληλη τιμή για το tempo του μουσικού κομματιού. Το tempo ορίζεται ως ένα είδος μεταβλητής παγκόσμιας εμβέλειας(global range) για το patcher μας και με αυτόν τον τρόπο μπορούμε να διαχειριστούμε το ρυθμικό αποτέλεσμα της μουσικής σε κάθε subpatcher.

Στη συνέχεια, για κάθε καρέ λαμβάνονται οι μέσες τιμές για τα Hue, Saturation, Lightness, όχι όμως για όλα τα εικονοστοιχεία. Μέσω ενός είδος φίλτρου, επιλέγουμε αυτά τα οποία είναι είτε τα φωτεινότερα είτε τα σκοτεινότερα, ανάλογα με το που βρίσκεται η σημαντικότερη πληροφορία κάθε φορά. Ορίζοντας μια τιμή για το όριο επιθυμητής φωτοβολίας προηγουμένως, υπολογίζονται μόνο αυτά τα οποία έχουν μεγαλύτερη ή μικρότερη τιμή αντίστοιχα από το όριο.

Η τιμή των H,S,L στο MAX κυμαίνεται από 0-255, όπως και με τα R,G,B. Αυτές οι τιμές αντιστοιχούν σε διαφορετικές περιοχές στο χρωματικό κύκλο. Συγκεκριμένα όμως με βάση την εκάστοτε μέση τιμή του Hue έχουν επιλεχθεί 5 διαφορετικές μουσικές κλίμακες, ή πιο συγκεκριμένα Τρόποι, οι οποίες αντιστοιχούν αριθμητικά σε 5 διαφορετικές περιοχές του χρωματικού κύκλου. Η επιλογή αυτής της αντιστοίχισης, βασίστηκε σε γενικές γραμμές στη γενική ψυχολογική επίπτωση που προκαλούν τα χρώματα της κάθε περιοχής, όπως και τη γενική αίσθηση που προκαλεί ο κάθε Τρόπος.

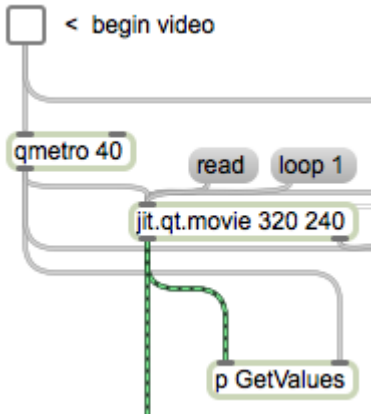
Αφού επιλέχθεί ο Τρόπος, οι μέσες τιμές H,S,L χρησιμοποιούνται για την επιλογή ενός μουσικού φθόγγου(pitch) που ανήκει στην κλίμακα αυτή μέσω μιας συναρτησιακής σχέσης 3 μεταβλητών.

Pitch=Hue-Saturation+2*Lightness

Θα πρέπει να σημειωθεί ότι το pitch λαμβάνει midi τιμές, οι οποίες όπως έχει αναφερθεί προηγουμένως, κυμαίνονται από 0-127. Επομένως, οι τιμές πριν χρησιμοποιηθούν έχουν μετατραπεί σε κατάλληλη κλίμακα. Συγκεκριμένα το Hue σε αυτή την εξίσωση έχει τιμή 0-30, το Saturation 15-0(αντιστρόφως ανάλογα από την τιμή του) και το Lightness 0-10.

Η τιμή του pitch χρησιμοποιείται για τη βασική μελωδία του ηχητικού αποτελέσματος. Επιπρόσθετα, έχουμε κίνηση του μπάσου, η οποία περιέχει ρυθμικές παραλλαγές, οι οποίες εξαρτώνται κυρίως από τυχαιότητα, αλλά και από τη συχνότητα αλλαγών χρωματικής περιοχής. Τέλος, υπάρχουν και κρουστά, τα οποία ενισχύουν το ρυθμικό κομμάτι του αποτελέσματος και περιέχουν κι αυτά τυχαιότητα.

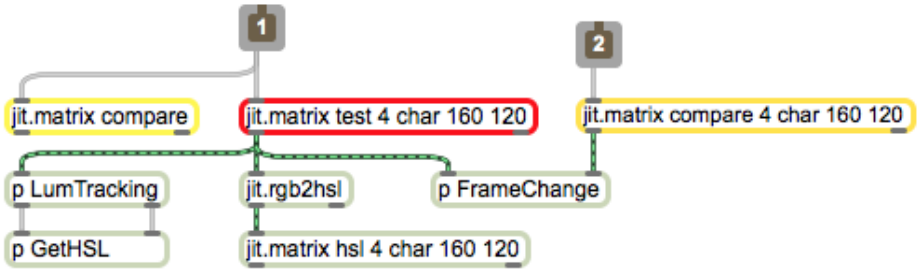
6.1 Εξαγωγή δεδομένων από βίντεο



Στο παραπάνω στιγμιοτύπο(screenshot) φαίνεται ένα κομμάτι από το βασικό patcher στο MAX. Πατώντας το **read** επιλέγεται το επιθυμητό βίντεο και κλείνοντας το διακόπτη(toggle), με το σχόλιο "begin video" ξεκινάει η αναπαραγωγή του σε ξεχωριστό παράθυρο. Το **jit.qt.movie 320 240** είναι υπεύθυνο για την αναπαραγωγή του οπτικού υλικού και έχουμε ορίσει από πριν την επιθυμητή ανάλυση (320x240 pixels). Το **qmetro 40** στέλνει *bang* κάθε 40 ms, επομένως με αυτόν τον τρόπο έχουμε 25 καρέ ανά δευτερόλεπτο, αριθμός αρκετά ικανοποιητικός για τη πλειοψηφία των βίντεο. Τέλος, υπάρχει επιλογή και για αναπαραγωγή πλήρης οθόνης, πιέζοντας το πλήκτρο “Esc”.

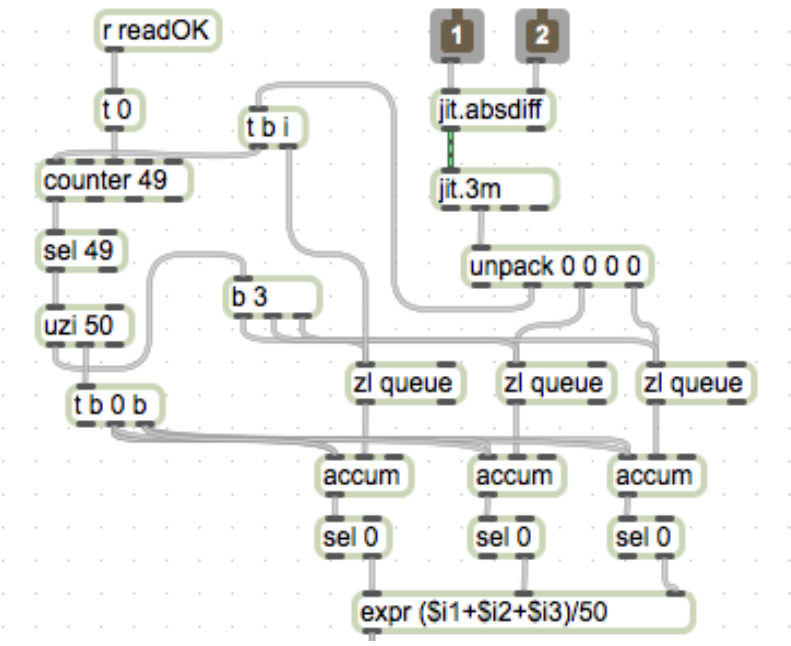
Το subpatcher **p Getvalues** είναι υπεύθυνο για την εξαγωγή των επιθυμητών δεδομένων. Όπως φαίνεται διαθέτει 2 εισόδους, η αριστερή λαμβάνει τα στοιχεία του κάθε καρέ σε μορφή πίνακα και η δεξιά δέχεται bang από το qmetro 40.

6.1.1 Εύρεση του ρυθμού χρωματικής μεταβολής του βίντεο

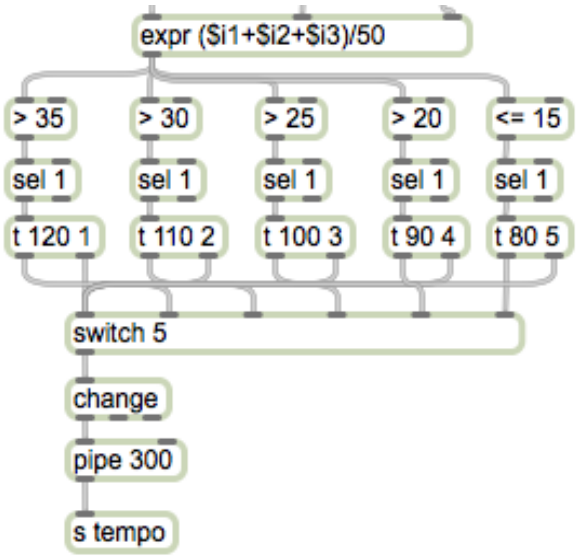


Το πρώτο βήμα στον αλγόριθμο είναι η ανίχνευσή του μέσου ρυθμού χρωματικής μεταβολής των καρέ ενός βίντεο. Η διαδικασία αυτή πραγματοποιείται στο **p FrameChange** και πετυχαίνεται με σύγκριση των τιμών R,G,B ενός καρέ σε σχέση με το προηγούμενο και μέσω ενός είδους ανάδρασης. Όπως φαίνεται στο παραπάνω στιγμιότυπο, ο πίνακας **jit.matrix compare** χρησιμοποιείται σε 2 διαφορετικά σημεία. Μέσω της εισόδου 1 αποθηκεύονται τα στοιχεία του καρέ στην είσοδο 2 δέχεται *bang* έτσι ώστε να εξαχθούν τα δεδομένα που είναι ήδη αποθηκευμένα στον πίνακα.

Το **p FrameChange** στη συνέχεια δέχεται αυτά τα δεδομένα από την δεξιά είσοδο και στην αριστερή δέχεται τα δεδομένα του επόμενου καρέ. Έτσι προχωράει σε σύγκριση όπως φαίνεται από το παρακάτω στιγμιότυπο, όπου μπορούμε να αναλύσουμε τη λειτουργία του.



Το **jit.absdiff** δημιουργεί έναν πίνακα που αποτελείται από τις απόλυτες διαφορές 2 πινάκων. Στην περίπτωση μας, από την είσοδο 2 δέχεται αρχικά τον πίνακα compare, στη δεξιά είσοδο του αντικειμένου, η οποία, όπως στα περισσότερα αντικείμενα του MAX, αποθηκεύει τον πίνακα έτσι ώστε να εκτελέσει την πράξη μόλις δεχτεί κάποιο μήνυμα στην αριστερή του είσοδο. Ο πίνακας που δέχεται εκεί είναι ο **test**, ο οποίος όπως είδαμε πριν, περιέχει τα στοιχεία του επόμενου καρέ. Χρησιμοποιώντας το **jit.3m** παίρνουμε τη μέση τιμή των διαφορών για τα R,G,B και η διαδικασία αυτή επαναλαμβάνεται για τα επόμενα 49 καρέ, όπου οι διαφορές αθροίζονται, προστίθενται μεταξύ τους και εξάγεται η μέση τιμή τους.



Στη συνέχεια, όπως φαίνεται έχουν τεθεί κάποια όρια (thresholds) και μέσω λογικών ελέγχων, στέλνονται οι ανάλογες διακριτές τιμές για την ταχύτητα(tempo), οι οποίες έχουν οριστεί από πριν. Το **change** μεταφέρει την τιμή μόνο αν είναι διαφορετική από την προηγούμενη και το **pipe 300** καθυστερεί τη μετάδοση της κατά 300 ms. Με αυτόν τον τρόπο αποφεύγουμε τυχόν υπερβολικά γρήγορες αλλαγές στην ταχύτητα, οι οποίες θα ήταν κακόηχες.

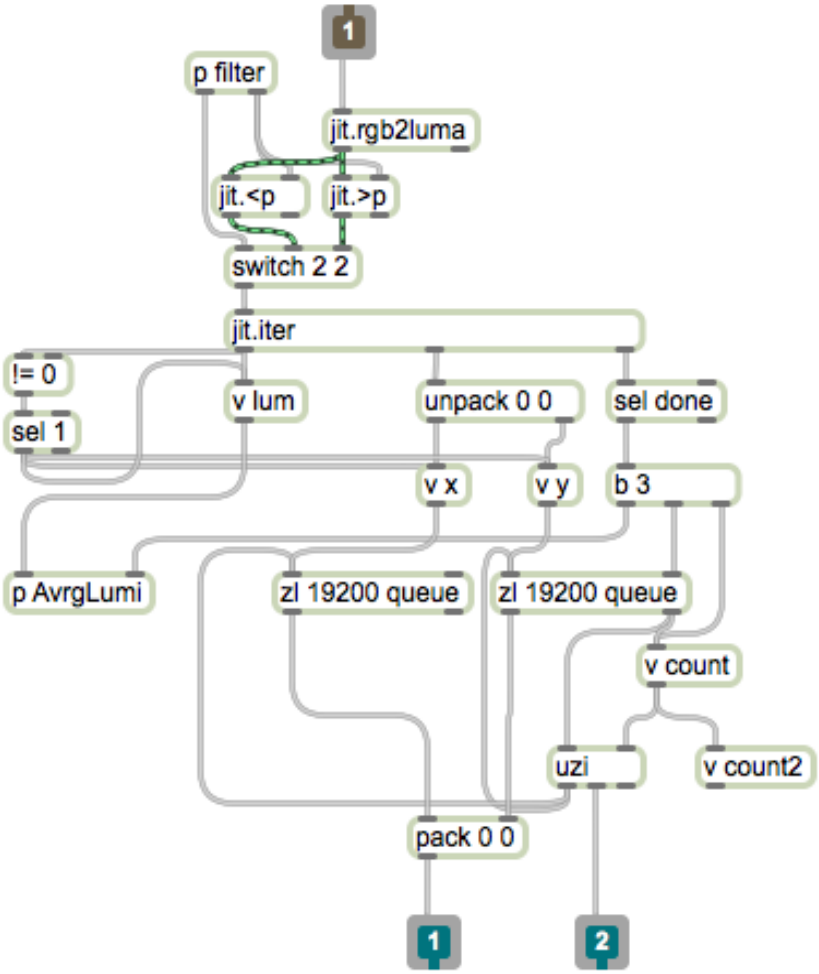
6.1.2 Εντοπισμός σημαντικών εικονοστοιχείων(pixel)

Το επόμενο βήμα είναι η ανίχνευση των εικονοστοιχείων που περιέχουν τις τιμές Hue, Saturation, Value που μας ενδιαφέρουν σε κάθε καρέ του βίντεο. Αυτό πραγματοποιείται μέσω του **p LumTracking** όπου αρχικά τα ARGB δεδομένα που είναι αποθηκευμένα στον πίνακα test μετατρέπονται σε μονοχρωματικά. Η μετατροπή γίνεται εύκολα με αντικείμενο του jitter, το jit.rgb2luma, το οποίο κάνει χρήση του τύπου:

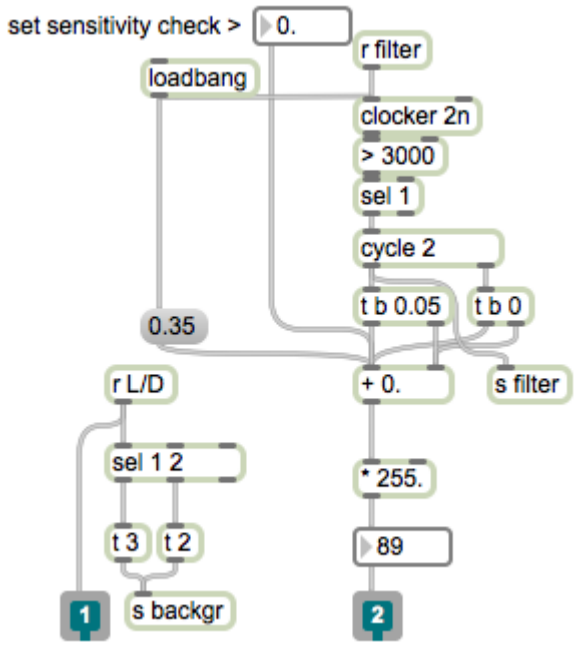
L = (.299 * R value) + (.587 * G value) + (.114 * B value).

Όπου L το Luminosity, και R,G,B οι γνωστές τιμές για κόκκινο, πράσινο, μπλε. Σε αυτή την ψηφιακή αναπαράσταση της μονοχρωματικής εικόνας, αντιστοιχεί μόνο μια τιμή για κάθε εικονοστοιχείο.

μετατροπή έγχρωμης εικόνας σε μονοχρωματική, όπου διατηρείται η φωτοβολία(luminosity) της αρχικής



Ακολουθεί μια διαδικασία φιλτραρίσματος, μέσω της οποίας επιλέγονται τελικά μόνο τα εικονοστοιχεία που έχουν τη μεγαλύτερη(ή μικρότερη για εικόνες με φωτεινό φόντο) φωτοβολία(luminosity) σε κάθε καρέ. Η επιλογή αυτή οφείλεται στο γεγονός ότι η σημαντικότερη πληροφορία σε κάθε εικόνα βρίσκεται σε αυτά τα εικονοστοιχεία. Στο **p filter** ορίζονται οι παράμετροι για τον έλεγχο και καθορίζεται αν πρόκειται για εικόνα με ανοιχτόχρωμο ή σκουρόχρωμο φόντο.



Στο **p filter** έχει οριστεί αρχικά ως συντελεστής ευαισθησίας το 0.35, που αντιστοιχεί στην τιμή 89 για τιμές 0-255. Ο συντελεστής μπορεί να μεταβληθεί από το χρήστη, αλλά και μεταβάλλεται με το χρόνο μέσω του **clocker 2n** (η λειτουργία του θα αναλυθεί παρακάτω), αυξομειώνοντας την τιμή του κατά 0.05.

Με αυτόν τον τρόπο το **switch** επιλέγει την είσοδο από το κατάλληλο στοιχείο (jit. <p ή jit. >p αντίστοιχα). Η αριστερή είσοδος χρησιμοποιείται για έλεγχο, σε αντίθεση με την πλειοψηφία των αντικειμένων στο MAX. Για τις τιμές 0, 1, 2 στην είσοδο ελέγχου, η έξοδος αντιστοιχεί με την κατάλληλη είσοδο(με 0 δεν υπάρχει έξοδος).

Στη συνέχεια τα δεδομένα αποθηκεύονται σε νέο κατάλληλο πίνακα και με χρήση του **jit.iter** εξάγεται το κάθε στοιχείο ξεχωριστά. Μέσω του **unpack 0 0**, τα v x, v y (values), αποθηκεύουν προσωρινά τις συντεταγμένες του κάθε εικονοστοιχείου προς εξέταση, ενώ το **v lum** την τιμή που αντιστοιχεί για τη φωτοβολία(luminosity). Στο != 0 πραγματοποιείται ο λογικός έλεγχος για μη μηδενικά στοιχεία, σε θετική περίπτωση το αποτέλεσμα είναι 1, ενώ σε αρνητική 0.

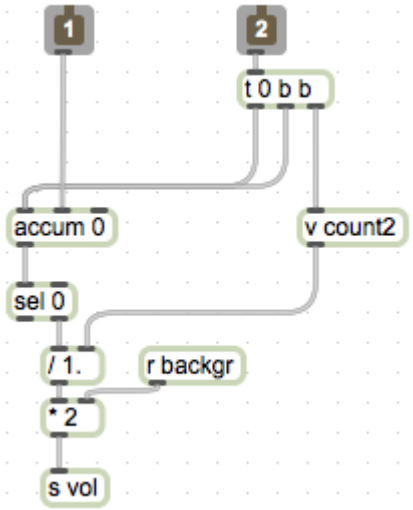
Το **sel 1** έχει παρόμοια λειτουργία με τη γνωστή συνθήκη "if … then...", από άλλες προγραμματιστικές γλώσσες, καθώς αν δεχτεί την τιμή 1 παράγει bang από την αριστερή έξοδο. Σε διαφορετική περίπτωση εξάγει τα δεδομένα απαράλλαχτα από την δεξιότερη έξοδο, η οποία στην προκειμένη περίπτωση δε χρησιμοποιείται. Με αυτόν τον τρόπο, αποθηκεύονται οι συντεταγμένες κάθε μη μηδενικού στοιχείου στα **zl 19200 queue**. Εκεί μπορούν να αποθηκευτούν λίστες, όπου το μέγιστο μέγεθος ορίζεται στην ονομασία τους (19200 στοιχεία) λειτουργούν με (First In, First Out) και διαθέτουν δείκτη για την καταμέτρηση και ταξινόμηση των αποθηκευμένων στοιχείων.

Όταν προσπελάσει όλα τα στοιχεία του πίνακα το **jit.iter** στέλνει το μήνυμα **done** από την δεξιότερη έξοδο, το οποίο έχει ως αποτέλεσμα να στείλει 3 bang σε 3 διαφορετικά σημεία. Το ένα είναι το **v count**, όπου έχει αποθηκευτεί ο αριθμός των μη μηδενικών στοιχείων στο zl queue. Το άλλο είναι το **uzi**, αντικείμενο που στέλνει πολλαπλά bang στη σειρά. Στην προκειμένη περίπτωση, στέλνει bang στα zl queue, τόσα όσα έχουν αποθηκευτεί στο **v**

count. Με αυτόν τον τρόπο στέλνουμε τις συντεταγμένες των μη μηδενικών εικονοστοιχείων στο **pack 0 0**, το οποίο απλά στέλνει τις x, y τιμές ως σύνολο 2 αριθμών(το αντίθετο του unpack 0 0). Το **uzi**, με το πέρας της διαδικασίας στέλνει ένα τελευταίο bang από τη μεσαία έξοδο, το οποίο χρησιμοποιείται αργότερα ως σήμα ελέγχου.

6.1.3 Υπολογισμός μέσης φωτοβολίας(luminosity) - επιλογή έντασης(volume)

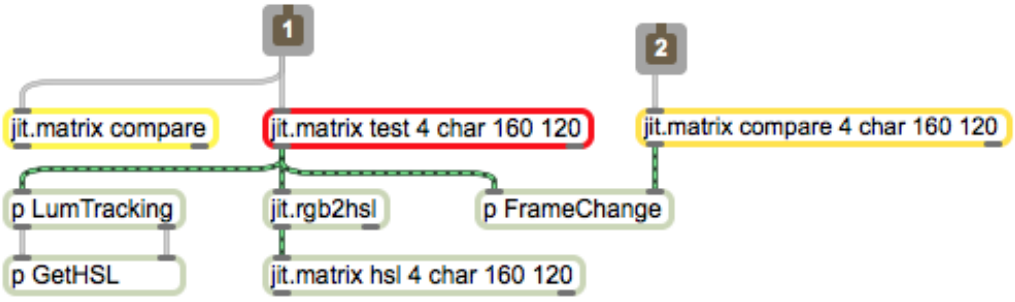
Το τρίτο bang πηγαίνει στο subpatcher **p AvrgLumi**, το οποίο υπολογίζει τη μέση τιμή φωτοβολίας(luminosity) για τα εικονοστοιχεία που μας ενδιαφέρουν, η οποία χρησιμοποιείται αργότερα για την παραμετροποίηση της έντασης (volume).



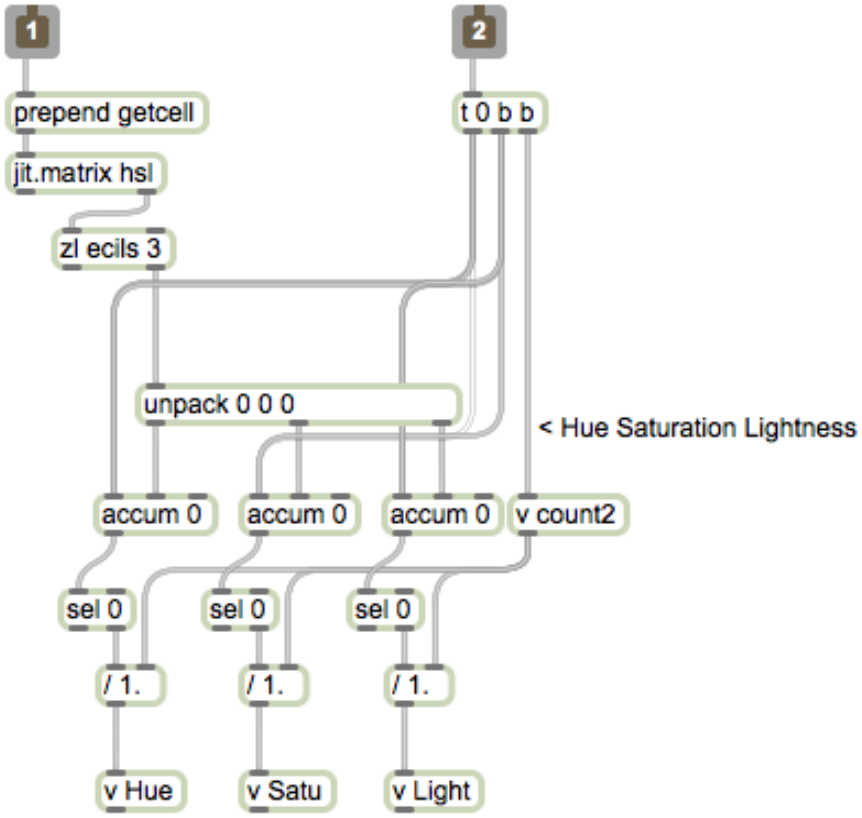
Το **accum 0** προσθέτει τις τιμές φωτοβολίας(luminosity) από την είσοδο 1 και όταν δεχτεί το bang από το t 0 b b, στέλνει το συνολικό άθροισμα, το οποίο στη συνέχεια διαιρείται από τον αριθμό των στοιχείων που βρίσκεται αποθηκευμένο στο **v count2**. Το τελικό αποτέλεσμα πολλαπλασιάζεται με έναν συντελεστή (2 ή 3), ανάλογα με τον τύπο του βίντεο (ανοιχτόχρωμο/σκουρόχρωμο) βίντεο. Σε αυτό το σημείο θα πρέπει να προσθέσουμε για την επικοινωνία των αντικειμένων στο MAX δεν είναι πάντα απαραίτητη η φυσική σύνδεση τους. Τα αντικείμενα send/ receive (συντομογραφία s/r) χρησιμοποιούνται ως πομποί και δέκτες αντίστοιχα για τη μεταφορά μηνυμάτων σε διαφορετικά σημεία ενός patcher. Αρκεί να μοιράζονται το ίδιο όνομα. Επομένως εδώ το **r backgr** στέλνει την πληροφορία που ορίζει το συντελεστή από κάποιο άλλο σημείο του patcher και το **s vol** μεταφέρει την τιμή για την ένταση σε κάποιο άλλο σημείο του patcher, εκεί όπου γίνεται η αναπαραγωγή του ήχου.

6.1.4 Υπολογισμός μέσων τιμών H, S, L για τα ενδιαφερόμενα εικονοστοιχεία

Η διαδικασία που πραγματοποιείται στη συνέχεια είναι η εξαγωγή των HSL δεδομένων για τα επιλεγμένα εικονοστοιχεία. Η λειτουργία αυτή πραγματοποιείται στο **p GetHSL** το οποίο δέχεται τις εισόδους του από το **p LumTracking** όπως φαίνεται κι από το παρακάτω στιγμιότυπο.



Για να επιτευχθεί αυτό, όμως αρχικά τα δεδομένα μας αποθηκεύονται επίσης στον πίνακα **jit.matrix hsl**, αλλά σε μορφή HSL, όπως υποδεικνύει και η ονομασία του, ύστερα από κατάλληλη μετατροπή μέσω του **jit.rgb2hsl**. Ας δούμε τώρα λοιπόν τη λειτουργία του **GetHSL**.



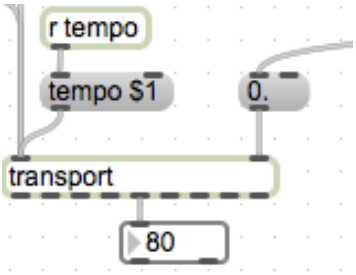
Η είσοδος 1 μεταφέρει τις συντεταγμένες του στοιχείου που μας ενδιαφέρει από το pack 0 0 του p LumTracking. Με την εντολή **prepend getcell** ουσιαστικά ζητείται από τον πίνακα hsl

να εξάγει τις τιμές Hue, Saturation και Value ως λίστα από την δεξιά έξοδο. Μέσω του **zl ecils 3**(μας δίνει μόνο τα 3 τελευταία στοιχεία μιας λίστας) και του unpack 0 0 0, ξεχωρίζουμε τις τιμές αυτές, στις οποίες υπολογίζεται η μέση τιμή με τρόπο παρόμοιο με αυτόν που αναλύθηκε προηγουμένως. Στη συνέχεια αυτές αποθηκεύονται προσωρινά ως values στα αντίστοιχα αντικείμενα που φέρουν το όνομα τους.

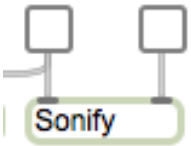
6.2 Επεξεργασία δεδομένων και αναπαραγωγή ήχου

6.2.1 Ορισμός Tempo

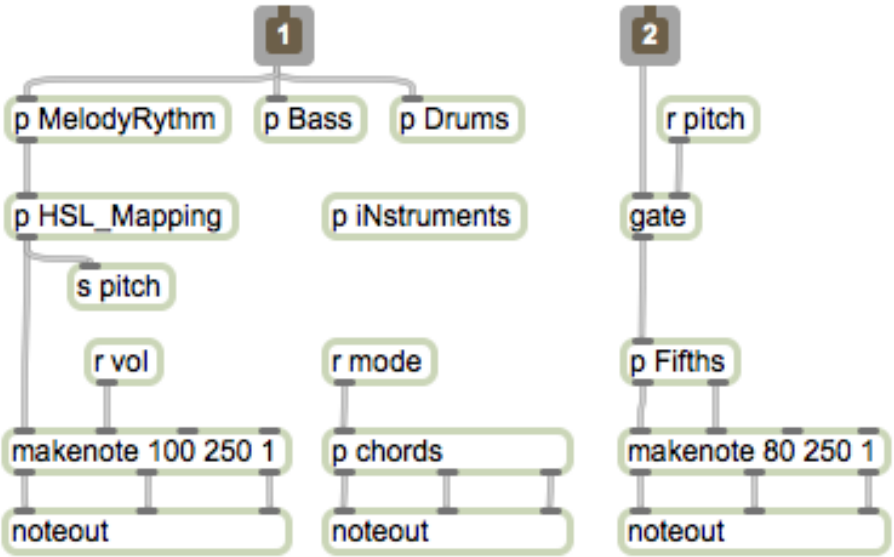
Εφόσον έχουν εξαχθεί τα απαραίτητα δεδομένα από το βίντεο, γίνεται επεξεργασία αυτών έτσι ώστε να υπάρξει αντιστοιχία με μουσικές παραμέτρους. Η κυριότερη λειτουργία είναι ο ορισμός του tempo για τη σύνθεση, κάτι που εξαρτάται όπως είπαμε από το ρυθμό αλλαγής των καρέ του βίντεο. Για το χειρισμό του tempo, το MAX διαθέτει ένα ειδικό αντικείμενο, το **transport**.



Στέλνοντας εδώ, λοιπόν, το αποτέλεσμα του FrameChange, μέσω των s tempo/ r tempo, ορίζουμε τους παλμούς ανά λεπτό(bpm) για τα τέταρτα. Η επεξεργασία των δεδομένων και το sonification πραγματοποιούνται στο subpatcher με όνομα **Sonify**.

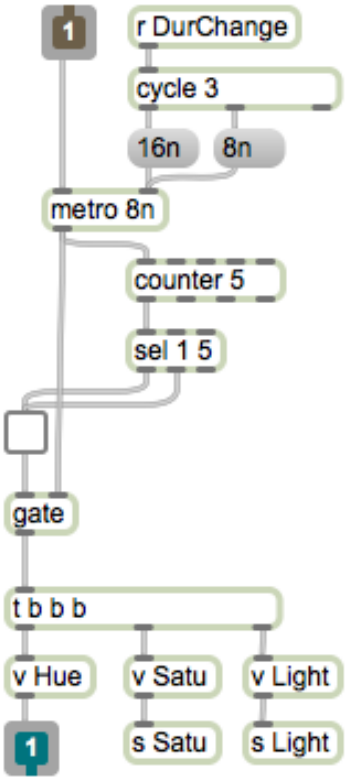


Παρακάτω φαίνεται το περιεχόμενο του **Sonify**.



6.2.2 Ρυθμός μελωδίας

Θα ξεκινήσουμε με την ανάλυση του πρώτου supatcher πάνω αριστερά, του **p MelodyRythm**.

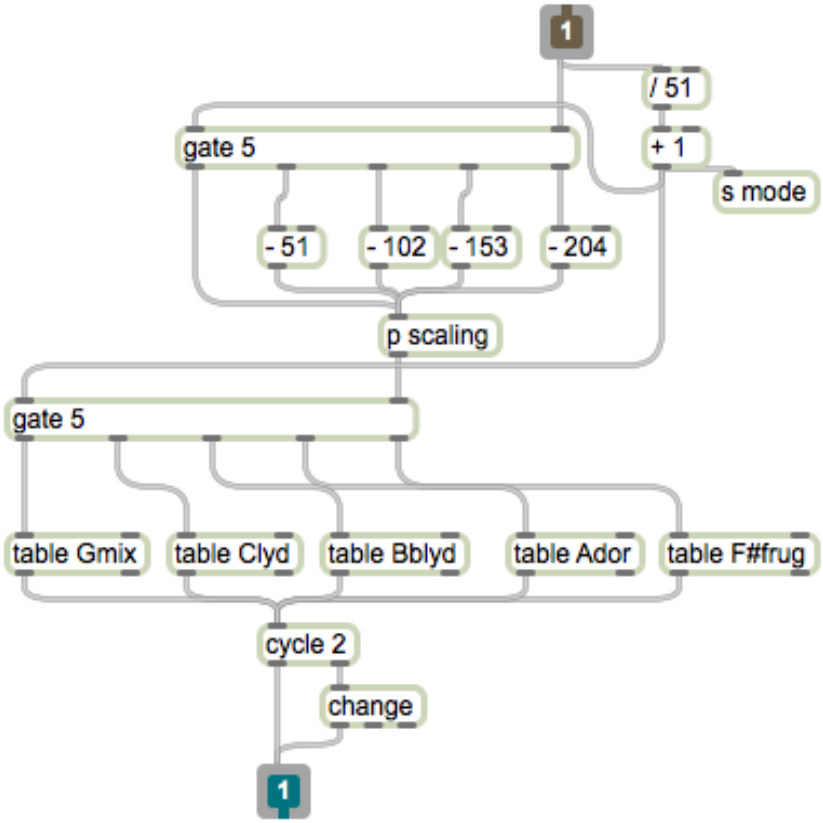


Σε αυτό το σημείο θα πρέπει να σημειωθεί πως στο MAX, οι μουσικές αξίες συμβολίζονται με έναν αριθμό και το χαρακτήρα n. Έτσι το ολόκληρο είναι το 2n, το τέταρτο 4n, το όγδοο 8n κ.ο.κ. Επομένως γίνεται κατανοητό, ότι μόλις τεθεί σε λειτουργία το metro 8n, στέλνει bang για κάθε όγδοο που αντιστοιχεί στο tempo που έχει οριστεί από το transport. Το **counter** είναι ένας απλός μετρητής, ο οποίος με κάθε bang αυξάνεται η τιμή εξόδου, ξεκινώντας από το 0 ως το 5 κ.ο.κ. Το **gate**, όπως και το switch, δέχεται το σήμα ελέγχου στην αριστερή είσοδο. Με την τιμή 1, τα δεδομένα από τη δεξιά είσοδο μεταφέρονται αυτούσια στην έξοδο, ενώ με την τιμή 0 δεν υπάρχει έξοδος. Ο συνδυασμός των **counter 5, sel 1 5** και **gate** έχει το εξής αποτέλεσμα. Για κάθε όγδοο του **metro**, δέχονται *bang* οι τιμές των Hue, Saturation , Lightness για τις τιμές 1 ως 4 του **counter.**

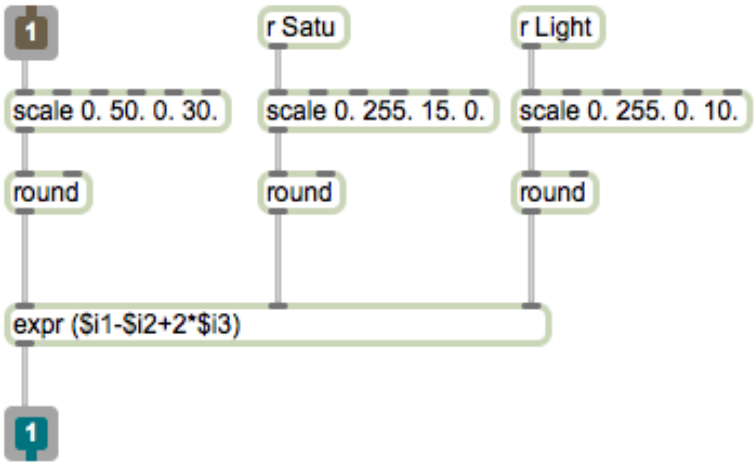
Υπάρχει μια επιπρόσθετη παραλλαγή σε αυτή τη διαδικασία, όπου το metro αλλάζει το ρυθμό bang (16n), μέσω κατάλληλου σήματος από το **r DurChange**. Το **cycle 3**, εξάγει τα δεδομένα εισόδου αυτούσια σε κάθε έξοδο εναλλάξ. Η τρίτη έξοδος δε χρησιμοποιείται, επομένως για κάθε τρίτο δεδομένο που δέχεται, διατηρείται η προηγούμενη κατάσταση.

6.2.3 Παραμετροποίηση HSL δεδομένων

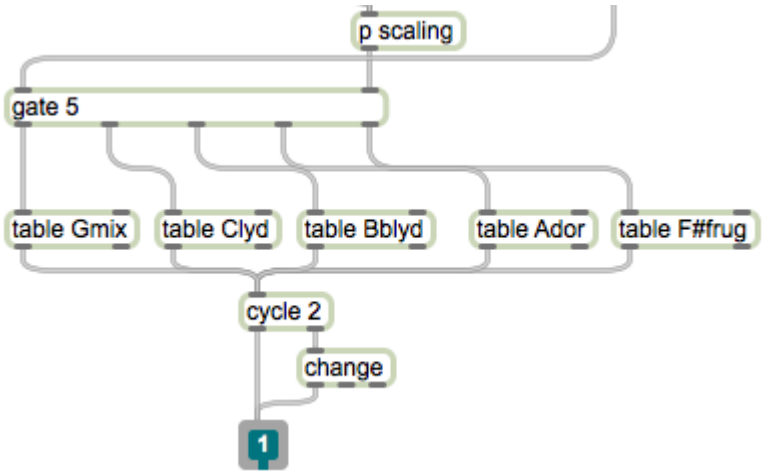
Στο **p HSL_Mapping** επιτελείται το σημαντικότερο κομμάτι του αλγορίθμου. Εδώ, ανάλογα με τη χρωματική περιοχή στην οποία βρίσκεται η εκάστοτε μέση τιμή για το Hue, επιλέγεται κατάλληλα ένας από τους 5 Τρόπους που έχουν οριστεί. Οι πέντε τρόποι είναι ο Σολ μιξολύδιος (Gmix), Ντο Λύδιος (Clyd), Σι ύφεση Λύδιος(Bblyd), Λα Δώριος (Ador) και Φα δίεση Φρύγιος (F#frug). Οι περιοχές τιμών του Hue στις οποίες αντιστοιχεί ο καθένας είναι 0-50, 51-101, 102-152, 153-203, 204-255. Για τον κάθε τρόπο αντιστοιχούν δηλαδή 51 τιμές εκτός από την τελευταία όπου έχουμε 52.



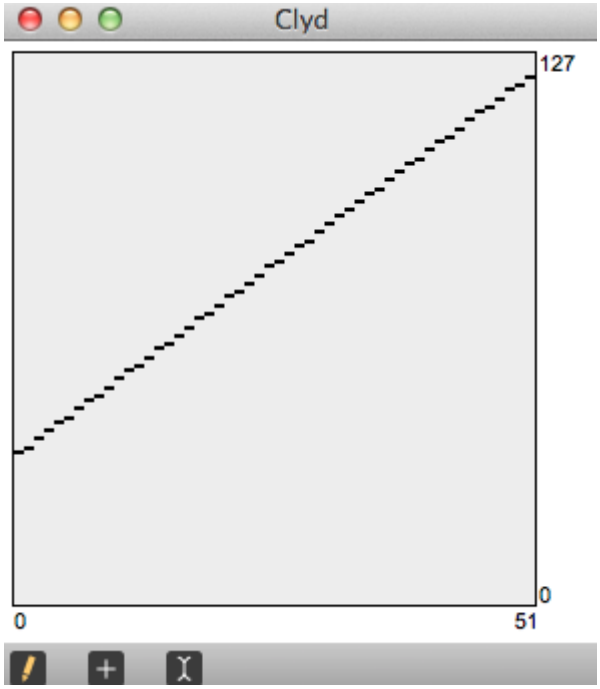
Στην είσοδο έχουμε τη μέση τιμή Hue που έχει υπολογιστεί από προηγούμενες διαδικασίες. Η τιμή αυτή αρχικά διαιρείται ακέραια με τον αριθμό 51, στη συνέχεια προστίθεται μία μονάδα κι έτσι ενεργοποιείται η κατάλληλη έξοδος για τα στοιχεία gate 5. Ας δούμε όμως και τη λειτουργία του **p scaling**.



Οι τιμές για τα Lightness και Saturation έχουν σταλεί από τις διαδικασίες που εκτελούνται στο **p MelodyRythm**. Στη είσοδο παίρνουμε μόνο την τιμή του Hue. Στις τιμές του Hue γίνεται αντιστοίχιση(scaling) ώστε να κυμαίνονται από 0-30, του Saturation από 0-15 και του Lightness 0-10. Παρατηρούμε ότι για το Saturation(κορεσμός) η αντιστοίχιση είναι αντιστρόφως ανάλογη. Αυτό γιατί στην συνάρτηση όπου υπολογίζεται η τελική τιμή για το μουσικό φθόγγο (pitch) η τιμή του Saturation αφαιρείται σε αντίθεση με τις άλλες 2 που λειτουργούν προσθετικά. Έτσι, όσο "καθαρό" είναι ένα χρώμα, τόσο λιγότερη ποσότητα αφαιρείται και τόσο πιο υψίσυχνος φθόγγος επιλέγεται. Κατά τα άλλα, το Lightness πολλαπλασιάζεται με το συντελεστή 2, κι έτσι για φωτεινότερα καρέ έχουμε πάλι πιο υψίσυχνους και ευδιάκριτους ήχους.



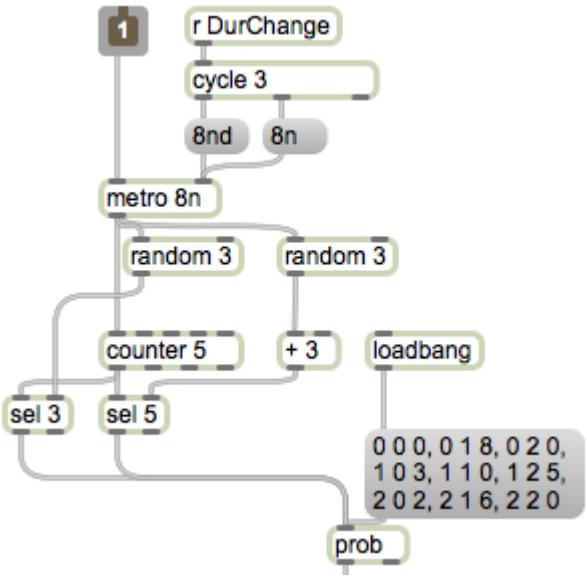
Η τιμή που υπολογίζεται από τη συνάρτηση H-S+2*L οδηγείται μέσω του gate 5 στον κατάλληλο πίνακα αντιστοίχισης με τον Τρόπο που αντιστοιχεί στην περιοχή χρωμάτων που βρισκόμαστε κάθε φορά. Οι πίνακες λειτουργούν ως συνάρτηση y(x), όπου x στην προκειμένη περίπτωση είναι οι τιμές από 0-51 και y οι τιμές που έχουμε αντιστοιχίσει για τον κάθε Τρόπο. (30-127).



Τέλος το **change** μεταφέρει την τιμή εισόδου μόνο αν είναι διαφορετική από την αμέσως προηγούμενη τιμή που έχει λάβει. Με αυτόν τον τρόπο αποφεύγονται τυχόν πολλές επαναλήψεις της ίδιας μουσικής νότας και εισάγονται περισσότερες παύσεις.

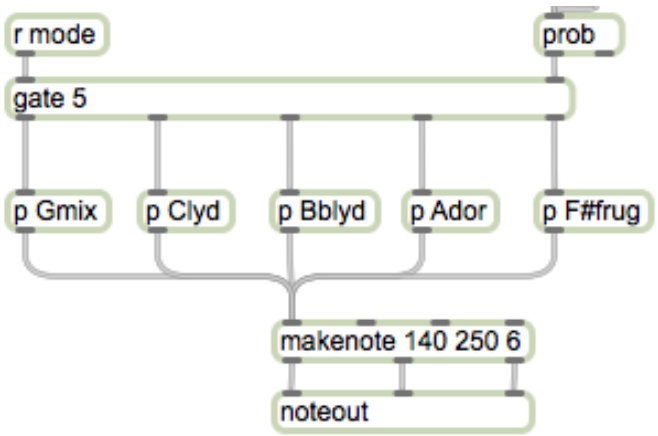
6.2.4 Κίνηση μπάσου

Υπεύθυνο για τη δημιουργία του μπάσου είναι το p Bass. Παρακάτω φαίνεται το πρώτο σκέλος της λειτουργίας του.

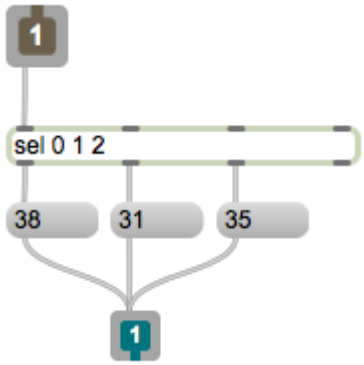


Όπως φαίνεται και σε αυτήν την περίπτωση χρησιμοποιείται το **metro 8n** για το ρυθμικό κομμάτι, με μια παρόμοια παραλλαγή. Η διαφορά που υπάρχει σε αυτήν την ρυθμική παραλλαγή είναι ότι εδώ αντί για 16n (δέκατο-έκτο) έχουμε 8nd (παρεστιγμένο όγδοο). Το παρεστιγμένο προσθέτει χρονική διάρκεια ίση με τη μισή της βασικής αξίας στην οποία προστίθεται. Στη συνέχεια βλέπουμε ότι πάλι χρησιμοποιείται counter 5 ως μετρητής και 2 sel. Έτσι ανά 6 μετρήσεις του counter έχουμε πάντα 2 "χτυπήματα" για το μπάσο, καθώς θέλουμε γενικά να είναι λιγότερο συχνά από αυτά της μελωδικής γραμμής(από p MelodyRythm).

Το **prob**, του οποίου η ονομασία προέρχεται από το probability(πιθανότητα), μπορεί να έχει 3 τιμές ως έξοδο (1, 2, 3). Δέχεται μια λίστα όπου θέτουμε τα ειδικά βάρη για μετάβαση από την κάθε κατάσταση σε οποιαδήποτε άλλη. Με αυτόν τον τρόπο με κάθε bang η έξοδος εξαρτάται από την προηγούμενη κατάσταση. Πρόκειται στην ουσία για προσομοίωση μοντέλου αλυσίδας Markov μηδενικής τάξης, 3 καταστάσεων. Στην προκείμενη περίπτωση π.χ. έχουμε θέσει [0 0 0, 0 1 8, 0 2 0]. Αυτό σημαίνει ότι αν η προηγούμενη έξοδος ήταν 0, ή επόμενη έξοδος μπορεί να είναι μόνο 1, καθώς η μετάβαση από 0 σε 1 και 0 σε 2 έχουν ειδικό βάρος 0.

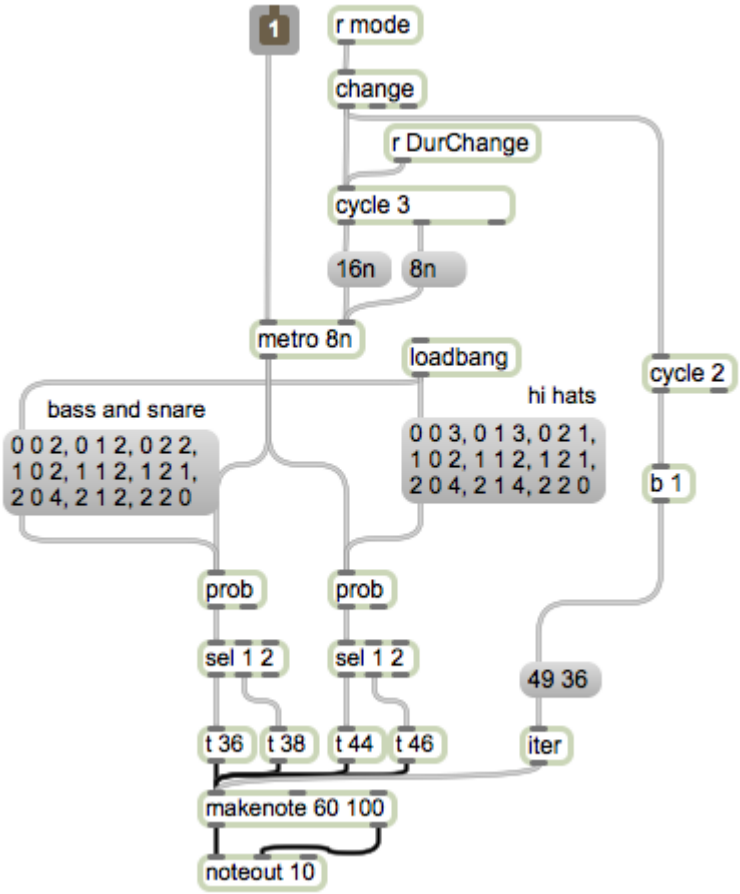


Στο δεύτερο σκέλος του subpatcher βλέπουμε ότι από μέσω του **r mode** ορίζεται η κατάλληλη έξοδος για το gate 5. Στη συνέχεια έχουμε 5 πανομοιότυπα subpatcher, στα οποία έχουν οριστεί 3 διαφορετικές νότες για κάθε Τρόπο και η επιλογή τους γίνεται με τυχαίο τρόπο, σύμφωνα με την έξοδο του prob. Παρακάτω βλέπουμε το **p Gmix.**



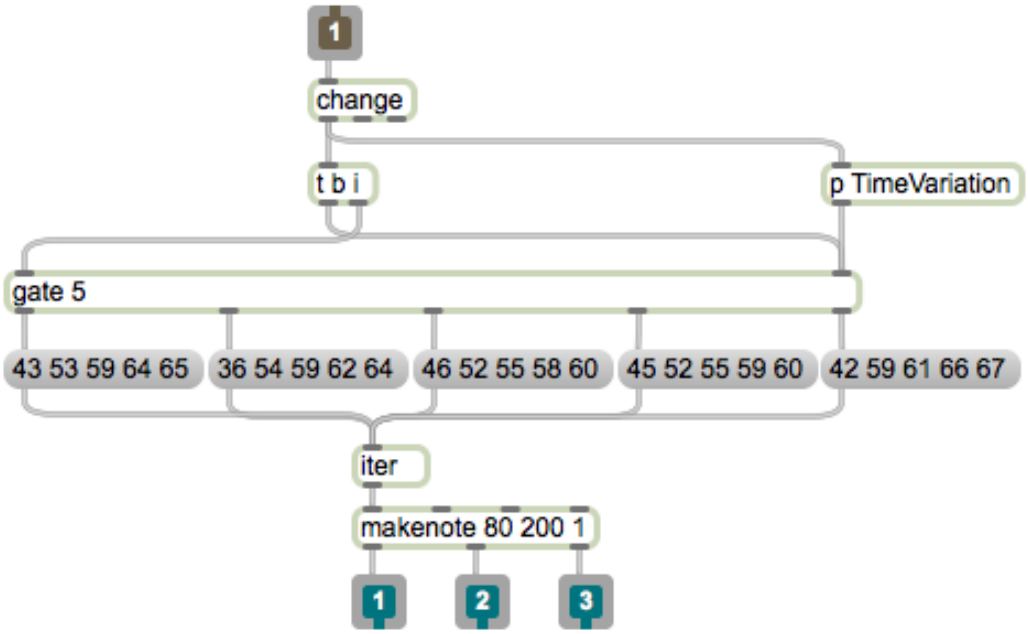
Εδώ φαίνεται πως για τις πιθανές εξόδους του **prob** (1,2,3), μέσω του **sel** επιλέγονται οι midi τιμές που έχουν οριστεί από πριν για το συγκεκριμένο τρόπο

6.2.5 Κρουστά



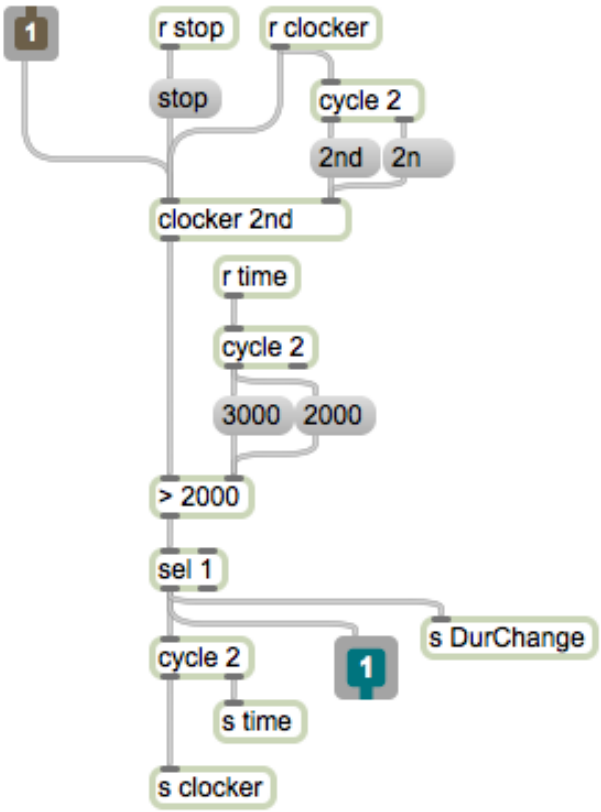
Στο **p Drums** χρησιμοποιείται η ίδια ρυθμική παραλλαγή με το **p MelodyRythm** και σε αυτήν την περίπτωση χρησιμοποιούνται 2 prob. Το πρώτο είναι υπεύθυνο για να επιλέγει με βάση τα ειδικά βάρη που έχουμε θέσει, ανάμεσα σε αναπαραγωγή midi ήχου που αντιστοιχεί σε "μπότα" (bass drum) και ταμπούρο(snare drum). Το δεύτερο επιλέγει ανάμεσα σε ήχο ανοιχτού πιατινιού ή κλειστού (open hi-hat, closed hi-hat). Και στις 2 περιπτώσεις η τιμή 0 χρησιμοποιείται σαν παύση κι αυτό φαίνεται από τα sel, τα οποία επιλέγουν μόνο ανάμεσα στις τιμές 1 και 2.

6.2.6 Συγχορδίες



Στην είσοδο δεχόμαστε το αποτέλεσμα του r mode, το οποίο καθορίζει τη συγχορδία που θα ακουστεί, ανάλογα πάλι με τη χρωματική περιοχή στην οποία βρισκόμαστε. Το change χρησιμεύει έτσι ώστε αυτό να γίνεται με κάθε αλλαγή Τρόπου. Για κάθε ένα από τους 5 Τρόπους έχει οριστεί μια συγχορδία, έτσι ώστε να έχει μεγαλύτερη αρμονική συνοχή το αποτέλεσμα. Το αντικείμενο **iter** χρησιμεύει έτσι ώστε να μπορούν να συνηχούν πάνω από 2 φθόγγοι ταυτόχρονα μέσω του makenote.

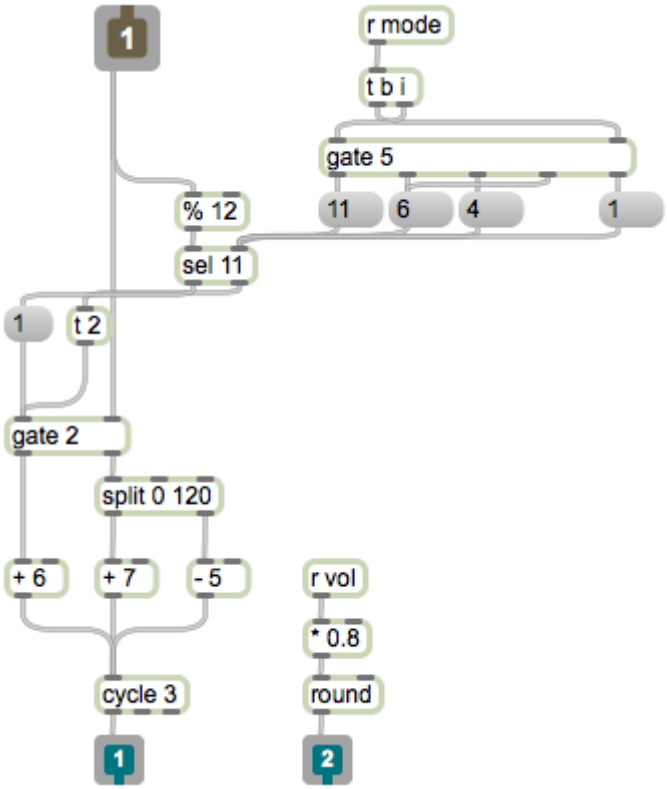
Το **p TimeVariation** με χρήση ειδικών χρονομετρητών επαναλαμβάνει τη συγχορδία ύστερα από διαφορετικά χρονικά σημεία, σε περίπτωση που το καρέ του βίντεο δεν έχει αλλάξει χρωματική περιοχή.



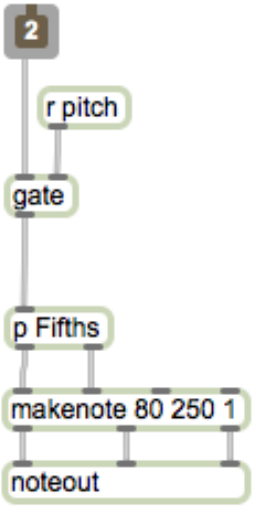
Το **clocker** είναι ο χρονομετρητής του MAX, ο οποίος με το που δεχτεί *bang* ξεκινάει τη χρονομέτρηση και το αποτέλεσμα της μέτρησης εξάγεται για κάθε παρεστιγμένο μισό (2nd). Αν η απόλυτη χρονική αξία του γεγονότος αυτού είναι μεγαλύτερη από 2000ms προκαλούνται διάφορες παραλλαγές στη χρονική μεταβολή του **timer**, αλλά και στα υπόλοιπα subpatchers όπως είδαμε προηγουμένως, μέσω του **s DurChange**.

6.2.7 Πέμπτες

Ο χρήστης έχει την επιλογή να προστεθεί και μια δεύτερη μελωδική φωνή, η οποία βρίσκεται σε διάστημα 5ης από την πρώτη φωνή δημιουργώντας έτσι ένα πιο μοντέρνο ηχητικό αποτέλεσμα και η λειτουργία αυτή επιτελείται μέσω του **p Fifths.**



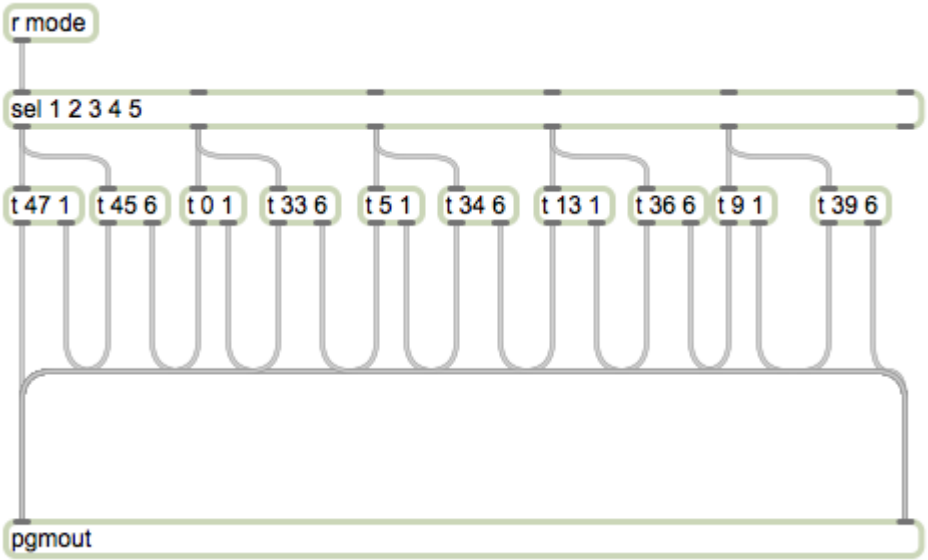
Για να έχουμε τη νότα που αντιστοιχεί σε διάστημα καθαρής πέμπτης πάνω από ένα οποιαδήποτε μουσικό φθόγγο, αρκεί να προσθέσουμε τον αριθμό 7 στο midi αριθμό που αντιστοιχεί στον αρχικό φθόγγο (ή να αφαιρέσουμε τον αριθμό 5 αν θέλουμε να πάρουμε την ίδια νότα μια οκτάβα χαμηλότερα). Όμως σε κάθε τρόπο από αυτούς που έχουμε επιλέξει, οι οποίοι διαθέτουν 7 διαφορετικούς μουσικούς φθόγγους σε διάφορες οκτάβες, υπάρχει ένας φθόγγος, όπου το διάστημα καθαρής πέμπτης δεν είναι εύηχο. Χρειάζεται διάστημα ελαττωμένης πέμπτης και άρα με midi όρους να προσθέσουμε τον αριθμό 6 αντί για 7. Με το % 12 βρίσκουμε το υπόλοιπο της ακεραίας διαίρεσης, από την αρχική νότα μας με τον αριθμό 12. Καθώς στο σύνολο ενός κλαβιέ υπάρχουν 12 διαφορετικοί φθόγγοι συνολικά οι οποίοι επαναλαμβάνονται σε υψηλότερες οκτάβες, με το υπόλοιπο εντοπίζουμε για ποιο φθόγγο πρόκειται.



Μέσω του **r mode** και του **gate**, λαμβάνεται η πληροφορία για τον Τρόπο που χρησιμοποιείται και στη συνέχεια το sel χρησιμοποιεί τον κατάλληλο αριθμό(που αντιστοιχεί στη νότα που χρειάζεται ελαττωμένη πέμπτη) στον οποίο θα προσθέσει τον αριθμό 6 αντί για τον αριθμό 7. Το **split 0 120** χρησιμεύει έτσι ώστε το τελικό αποτέλεσμα να είναι στα όρια των midi αριθμών (0-127). Όταν η βασική μελωδία αντιστοιχεί σε αριθμό μεγαλύτερο από 120, τότε αφαιρείται το 5 αντί να προστεθεί το 7, έχοντας ως αποτέλεσμα τον ίδιο φθόγγο σε χαμηλότερη οκτάβα. Τέλος το **cycle 3** χρησιμοποιείται έτσι ώστε να ακούγεται η πέμπτη ανά 3 νότες της βασικής μελωδίας.

6.2.8 Επιλογή μουσικού οργάνου αναπαραγωγής

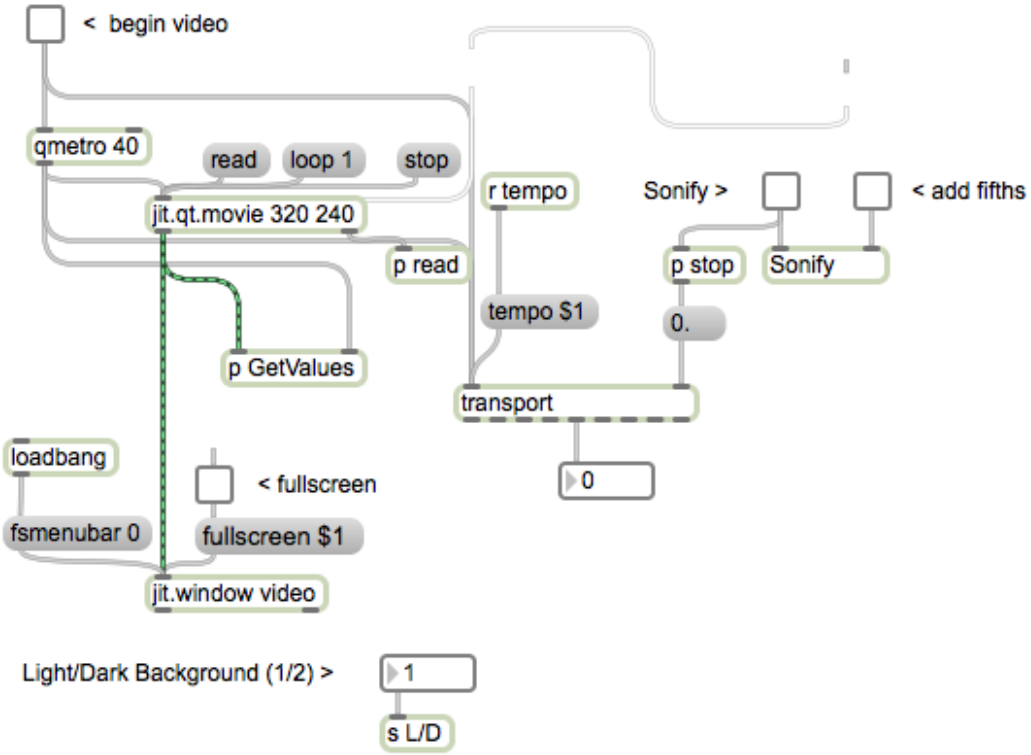
Τέλος για να γίνεται καλύτερα αντιληπτή η επιλογή Τρόπου και η χρωματική περιοχή στην οποία βρίσκονται τα σημαντικά εικονοστοιχεία του βίντεο, επιλέχθηκαν 5 ξεχωριστοί ήχοι για τη μελωδία του κάθε Τρόπου και 5 ήχοι για το μπάσο αντίστοιχα. Η διαδικασία επιλογής λαμβάνει χώρα στο **p iNstruments**.



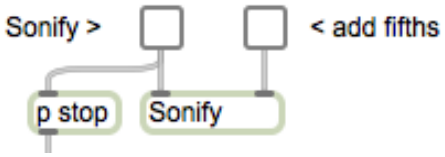
Η επιλογή γίνεται πολύ απλά μέσω του r mode και του sel. Αφού επιλεχθεί ο αριθμός που αντιστοιχεί στον Τρόπο στέλνονται 2 ζευγάρια αριθμών σε κάθε περίπτωση. Στο κάθε ζευγάρι ο αριστερός αριθμός αντιστοιχεί στο όργανο επιλογής, ενώ ο δεξιός στο midi κανάλι όπου θα χρησιμοποιηθεί αυτός. Για τη μελωδία χρησιμοποιούμαι το κανάλι 1 ενώ για το μπάσο το κανάλι 6. Το **pgmout** είναι το αντικείμενο που δέχεται αυτούς τους αριθμούς και κάνει τις επιθυμητές μετατροπές, έτσι ώστε τα noteout να αναπαράγουν ήχους με χρήση των συγκεκριμένων μουσικών οργάνων για το αντίστοιχο κανάλι.

6.3 Μορφή παρουσίασης στο βασικό patcher

Στο παρακάτω στιγμιότυπο βλέπουμε ολόκληρη τη δομή του βασικού patcher που επιτελεί όλες τις λειτουργίες που συζητήθηκαν προηγουμένως



Αφού λοιπόν ο χρήστης επιλέξει το επιθυμητό βίντεο κλείνοντας τους αντίστοιχους διακόπτες ξεκινάει η ηχοποίηση του. Ο δεύτερος διακόπτης χρησιμεύει για την αναπαραγωγή πέμπτων, στις οποίες αναφερθήκαμε προηγουμένως



Τέλος, επιλέγει αν το φόντο που υπάρχει είναι φωτεινό η σκοτεινό έτσι ώστε να γίνουν οι κατάλληλες μετατροπές στον αλγόριθμο μέσω του **s L/D**.

7. Προτάσεις για μελλοντική επέκταση του αλγορίθμου

Ο αλγόριθμος που χρησιμοποιήθηκε για το Video Sonification αποτελεί μια πρώτη απλή σχετικά προσέγγιση του ζητήματος. Μπορούν να προστεθούν επιπρόσθετα στοιχεία ανάλογα με την κατεύθυνση που θα ακολουθηθεί ως προς τη σύνθεση του κομματιού. Οι κατευθύνσεις αυτές διαχωρίζονται με βάση 2 διαφορετικά κριτήρια που αφορούν τους κανόνες αλγοριθμικής σύνθεσης που χρησιμοποιούνται.

Ανάλογα με το είδος αρμονίας που επιθυμεί ο συνθέτης, το αποτέλεσμα μπορεί να είναι

- Μουσική με χρήση τονικών κέντρων
- Ατονική μουσική

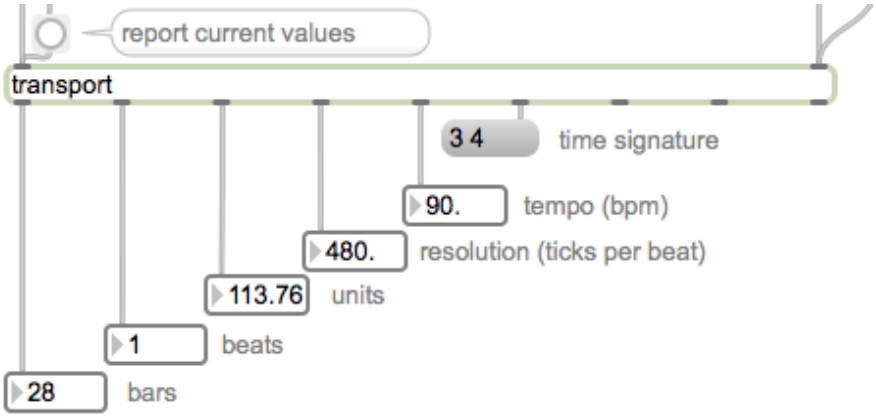
Ανάλογα με το είδος των ήχων που θα χρησιμοποιηθούν

- Μουσική με χρήση Αρμονικών ήχων
- Μουσική με χρήση Μη-αρμονικών τόνων και *Θόρυβο*

7.1 Με χρήση αρμονικών ήχων-Τονικών κέντρων

Στην παρούσα εργασία, ο αλγόριθμος σχεδιάστηκε έτσι ώστε η σύνθεση να κάνει χρήση τονικών κέντρων και αρμονικών ήχων. Ακολουθώντας λοιπόν αυτήν την κατεύθυνση, το επόμενο βήμα θα ήταν μια περεταίρω παραμετροποίηση του ηχητικού αποτελέσματος, με χρήση διάφορων μουσικών ρευμάτων της σύγχρονης δημοφιλούς μουσικής. Η επιλογή αυτή θα μπορεί να γίνεται είτε από το χρήστη, ή ιδανικότερα παραμετροποιώντας κάποια χαρακτηριστικά του βίντεο έτσι ώστε να γίνεται αυτή η επιλογή αυτόματα. Τα ρεύματα αυτά έχουν ορισμένα ιδιαίτερα χαρακτηριστικά, όσον αφορά τη μελωδία, το ρυθμό αλλά και τα όργανα που χρησιμοποιούνται για κάθε είδος. Για να γίνει πιο κατανοητή η επέκταση αυτή θα αναλυθεί ένα παράδειγμα.

Αν γινόταν π.χ. επιλογή για βαλς, ως ύφος του κομματιού, τότε όσον αφορά το ρυθμικό κομμάτι, το μέτρο του κομματιού είναι θα έπρεπε να γίνει Τρία Τέταρτα (3/4). Αυτό μπορεί να γίνει πολύ εύκολα στο MAX μέσω του Transport, όπου μπορούν να οριστούν εύκολα το tempo, το επιθυμητό time signature, αλλά και άλλες δυνατότητες όσον αφορά το ρυθμικό σκέλος.



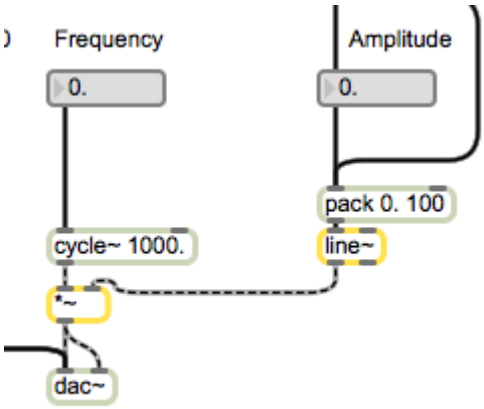
Το αρμονικό περιεχόμενο μπορεί να επιλεχθεί με χρήση πινάκων που θα περιέχουν αντιστοιχίσεις με νότες που ανήκουν στα επιθυμητά τονικά κέντρα και τέλος μπορεί να γίνει και επιλογή κατάλληλων ήχων για τα μουσικά όργανα αναπαραγωγής, χρησιμοποιώντας στην συγκεκριμένη περίπτωση πιάνο και κουαρτέτο εγχόρδων.

Παρόμοιες διαδικασίες θα ακολουθηθούν και για την επιλογή οποιαδήποτε άλλου μουσικού ύφους, δημιουργώντας έτσι ένα πιο ολοκληρωμένο αποτέλεσμα.

Όσον αφορά την κατεύθυνση ατονικής μουσικής, αλλά αρμονικών ήχων, μπορεί να γίνει χρήση στατιστικών μοντέλων, όπως αλυσίδες Markov, όπου η επόμενη νότα εξαρτάται από τις προηγούμενες. Τα δεδομένα του βίντεο μπορούν να χρησιμεύσουν ως παράμετροι για το επιλεγμένο στατιστικό μοντέλο που θα χρησιμοποιηθεί, έχοντας φυσικά το παράγοντα της τυχαιότητας.

7.2Εισαγωγή Μη-αρμονικών τόνων, Θορύβου μέσω του MSP

Η χρήση μη-αρμονικών τόνων και θορύβου σε μία σύνθεση συνήθως συνεπάγεται ατονική μουσική, χωρίς αυτό να είναι κάτι το υποχρεωτικό. Η σύνθεση μουσικής με χρήση μη-αρμονικών τόνων και θορύβου μπορεί να γίνει εύκολα εφικτή με το συμπληρωματικό εργαλείο του MAX, το MSP. Το MSP προσφέρει εργαλεία για την παραγωγή μη-αρμονικών τόνων, καθώς υπάρχει η δυνατότητα παραγωγής ήχων από ταλαντωτές, στους οποίους επιλέγουμε το πλάτος, τη συχνότητα αλλά και φάση. Τα αντικείμενα του MSP μοιάζουν αρκετά με του MAX, με τη διαφορά ότι η επικοινωνία μεταξύ τους γίνεται σε πολύ μεγαλύτερες ταχύτητες και ξεχωρίζουν από την κατάληξη ~.



Στο στιγμιότυπο φαίνεται ένας ταλαντωτής, ο **cycle~** με αρχική συχνότητα 1000Hz, η οποία μπορεί να μεταβληθεί κατάλληλα. Το **line~** αποτελεί ουσιαστικά την περιβάλλουσα της τελικής κυματομορφής και γι' αυτό τα αποτελέσματα τους πολλαπλασιάζονται μέσω του ***~**. Το **dac~** (digital to analog converter) είναι το αντικείμενο που μετατρέπει την ψηφιακή πληροφορία σε ήχο. Όπως φαίνεται διαθέτει 2 εισόδους, μία για το αριστερό και μία για το δεξί κανάλι εξόδου (Left, Right).

Γίνεται κατανοητό λοιπόν, ότι με χρήση του MSP διαθέτουμε περισσότερες παραμέτρους με τις οποίες μπορούμε να αντιστοιχίσουμε με τα στοιχεία του βίντεο. Μπορούν να χρησιμοποιηθούν πάνω από ένας ταλαντωτής, στους οποίους μπορούμε να διαμορφώσουμε κατάλληλα τη συχνότητα, το πλάτος ταλάντωσης, αλλά και τη διάρκεια του ήχου. Επίσης μπορούμε να πειραματιστούμε με τα στερεοφωνικά κανάλια του ήχου (panning) και γενικότερα να παραχθούν μη αρμονικοί τόνοι αλλά και ελεγχόμενος θόρυβος. Μέσω των εργαλείων που προσφέρει, μπορούμε να χρησιμοποιήσουμε οποιαδήποτε τεχνική σύνθεσης ήχου, από αυτές που αναφέρθηκαν σε προηγούμενη ενότητα (additive synthesis, amplitude modulation) και να δημιουργήσουμε έτσι ένα πιο σύγχρονο ηχητικό αποτέλεσμα.

Μια πιθανή και ενδιαφέρουσα τροποποίηση θα ήταν η εξής. Οι μέσες τιμές του Hue σε κάθε καρέ για τα εικονοστοιχεία που βρίσκονται πάνω από κάποιο επίπεδο φωτοβολίας (luminosity) μπορούν να παραμετροποιηθούν έτσι ώστε να χρησιμοποιηθούν ως συχνότητα ταλάντωσης ενός αριθμού ταλαντωτών. Οι μέσες τιμές του Hue για τα υπόλοιπα εικονοστοιχεία μπορούν να χρησιμοποιηθούν αντίστοιχα για άλλους ταλαντωτές, έχοντας 2 διαφορετικές μελωδίες που συνυπάρχουν με αυτόν τον τρόπο. Η παράμετρος του Saturation μπορεί να παραμετροποιηθεί με τη διάρκεια των ήχων και η παράμετρος Lightness με το πλάτος των ταλαντωτών και άρα την ένταση ουσιαστικά. Τέλος, ανάλογα με το που βρίσκονται στην οθόνη η πλειοψηφία των σημαντικών εικονοστοιχείων, μπορεί να διαμορφωθεί κατάλληλα ο ήχος στερεοφωνικά.

8. Βιβλιογραφία/Πηγές

[1] Βικιπαίδεια

[2] Communications of the ACM

http://cacm.acm.org/magazines/2011/7/109891-algorithmic-composition/fulltext

[3] John A. Maurer, Stanford University

https://ccrma.stanford.edu/~blackrse/algorithm.html

[4] Jack Belzer, Albert G. Holzman, Allen Kent

Encyclopedia of Computer Science and Technology: Volume 11 - Minicomputers to PASCAL

[5] John Bewley- Lejaren A. Hiller: Computer music pioneer

http://library.buffalo.edu/music/exhibits/hillerexhibitsummary.pdf

[6] Algorithmic Composition: A Gentle Introduction to Music Composition Using Common LISP and Common Music, University of Michigan

http://quod.lib.umich.edu/s/spobooks/bbv9810.0001.001/1:5?rgn=div1;view=fulltext

[7] Αλγοριθμική Σύνθεση Ήχου και Μουσικής- Κυριακή Υψηλάντη

[8] Algorithmic Composition: A Guide to Composing Music with Nyquist

[9] David Wright - Mathematics and Music

[10] Το Τονικό μουσικό Σύστημα - Αμάραντος Αμαραντίδης

[11] Ανδρέας Ιωάννου Κασσέτας- Πως αντιλαμβάνεται το μάτι το χρώμα
http://users.sch.gr/kassetas/xCOLOR27.htm

[12] MSP Documentation: How digital Audio works