

Polyfōnía

Ένας διαδραστικός ψηφιακός επεξεργαστής φωνής
με ανίχνευση κίνησης

Ναταλία Κωτσάνη

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Jazz και Νέες Τεχνολογίες

19 Ιουλίου 2019

Overview

- 1 Εισαγωγή
 - Polyfōnía
 - Γιατί Max/Msp;
 - Καινοτομία
- 2 Διαδραστική ερμηνεία της μουσικής και Νέες Τεχνολογίες
 - Επίπεδα διάδρασης
 - Παράμετροι διάδρασης του Polyfōnía
- 3 Φωνητικοί επεξεργαστές
 - Κατηγοριοποίηση
 - Live Looping
 - Harmonizer
- 4 Συστήματα και Αισθητήρες Ανίχνευσης Κίνησης
 - Συγκριτική Μελέτη Ανιχνευτών Κίνησης
 - Επίπεδα διάδρασης
 - Παραδείγματα
 - Συγκριτική Μελέτη Ανιχνευτών Κίνησης
 - Leap Motion

Εισαγωγή: Polyfōnía

Φωνητικοί επεξεργαστές

Μπορούν να αξιοποιηθούν σε πραγματικό χρόνο από τους τραγουδιστές, διευρύνοντας τα όρια των δυνατοτήτων της ανθρώπινης φωνής;

Polyfōnía: ένα σύστημα ηχητικών βρόχων πραγματικού χρόνου (live looping) και εναρμόνισης φωνής (harmonizer) το οποίο ο χρήστης ελέγχει μέσω κινήσεων των δακτύλων του¹.

Συγκεκριμένες χειρονομίες οδηγούν σε:

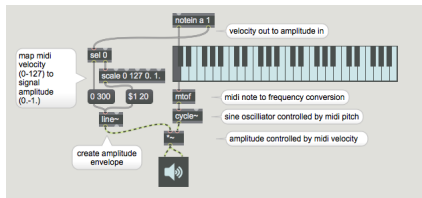
- εκκίνηση, διακοπή ή αναπαραγωγή της ηχογράφησης για το live looping
- ανάπτυξη αρμονιών φωνής, σύμφωνα με την παραμετροποίηση
- ενεργοποίηση, απενεργοποίηση του συστήματος

I: ηχητικό σήμα ενός μικροφώνου - O: το επεξεργασμένο ηχητικό σήμα.

¹ Η υλοποίηση πραγματοποιήθηκε στο Max/Msp/Jitter.

Εισαγωγή: Γιατί Max/Msp;

- πλούσια βιβλιοθήκη για φωνητικούς επεξεργαστές (πχ. ανάλυση σήματος πραγματικού χρόνου (real-time signal-processing))
- εύκολη δημιουργία εφαρμογών με χειρισμό σε πραγματικό χρόνο
- ταυτόχρονη δημιουργία γραφικού περιβάλλοντος (Graphical User Interface, GUI)
- patch αλληλεπίδρασης με το Leap Motion, από το IRCAM



Καινοτομία

Μέχρι σήμερα...

- υπάρχοντα συστήματα επεξεργαστών φωνής: προσαρμοσμένα στις ανάγκες των ερμηνευτών μουσικών οργάνων
- χειρισμός τους από τους τραγουδιστές: ως εμπόδιο στην εκφραστικότητα, την επικοινωνία με το κοινό και προϋποθέτει την αναπροσαρμογή της στάσης του σώματός τους
- παραμετροποίηση σε μία διάσταση
- έλεγχος: συχνά γίνεται από τους τεχνικούς ήχου και όχι από τους ίδιους τους ερμηνευτές

Ο χειρισμός του Polyfōnία μέσω ανίχνευσης κίνησης:

- πραγματοποιείται με τα χέρια και δεν προϋποθέτει την ύπαρξη οθόνης
- είναι βιωματικός, αίρει την ασυνέχεια (discontinuity) στην επικοινωνία ανθρώπων - υπολογιστή
- έχει τρισδιάστατο χαρακτήρα, όπως η «τρισδιάστατη εικόνα» του ήχου
- μπορεί να πραγματοποιηθεί σε πραγματικό χρόνο από τον τραγουδιστή σύμφωνα με την τεχνική και την αισθητική του.

Επίπεδα διάδρασης

Στη διαδραστική τέχνη (interactive art), ο Edmonds κατέταξε τα επίπεδα διάδρασης μεταξύ του κοινού και των καλλιτεχνικών έργων στις εξής κατηγορίες:

- **στατική (static):** το έργο δε μεταβάλλεται, απλώς παρατηρείται
- **δυναμική-παθητική (dynamic-passive):** το έργο περιλαμβάνει εσωτερικό μηχανισμό που του επιτρέπει τη μεταβολή του είτε ανεξάρτητα ή σε σχέση με εξωτερικούς παράγοντες ή δεδομένα αισθητήρων (αλλαγές προβλέψιμες ή ακολουθούν μία άγνωστη σύνθετη μη-γραμμική συνάρτηση ή στοχαστικές)
- **δυναμική διαδραστική (dynamic-interactive):** το κοινό διαδραματίζει σημαντικό ρόλο στην πρόκληση των μεταβολών
- **μεταβαλλόμενη δυναμική-διαδραστική (dynamic-interactive, varying):** οι μεταβολές λόγω ανθρώπινου παράγοντα ή μέσω λογισμικού (χρήσης τεχνικών μηχανικής μάθησης).

Επίπεδα διάδρασης

Κατά τον Edmonds η έρευνα στην τέχνη αναδεικνύει ενδιαφέρουσες πτυχές της διάδρασης, εξίσου σημαντικές και για τη μελλοντική έρευνα της διάδρασης ανθρώπου - υπολογιστή, όπως:

- οι μέθοδοι αξιολόγησης κάθε οντότητας που αλληλεπιδρά έχουν διαφορετικές παραμέτρους και οπτικές γωνίες αξιολόγησης
- η φύση της ενασχόλησης με ένα σύστημα διάδρασης μπορεί να είναι μεταβαλλόμενη (ανάγκη μελέτης φάσεων ενασχόλησης και μεταβάσεων)
- κάθε φάση ενασχόλησης και οι μεταβάσεις αποτελούν αντικείμενο έρευνας
- τα κριτήρια σχεδιασμού μπορεί να διαφέρουν ανάλογα με τη φάση διάδρασης.

Παράμετροι διάδρασης του Polyfōnía

Οι παράμετροι διάδρασης Polyfōnía, σχεδιάστηκαν έτσι ώστε:

- ο χειρισμός να μπορεί να γίνεται με κινήσεις χεριών, χωρίς την ύπαρξη καλωδίων, γαντιού ή συσκευής μεγάλου όγκου
- να μπορεί να γίνει ανίχνευση και των δύο χεριών
- οι κινήσεις που θα επιλεγθούν να είναι σαφείς, να διευκολύνουν το χειριστή στην ανάκλησή τους και να μην απαιτούν την τροποποίηση της ευθύγραμμης στάσης του ερμηνευτή
- να μην εμπεριέχουν έντονα κινησιολογικά στοιχεία, παρεμποδίζοντας την έκφραση του ερμηνευτή
- ο χειρισμός να μπορεί να πραγματοποιηθεί και σε συνθήκες χαμηλού φωτισμού
- η απόκριση να είναι άμεση και επιτυχής, χωρίς να απαιτείται από τον ερμηνευτή να επιβλέπει τη διαδικασία μέσω κάποιας οθόνης

Φωνητικοί επεξεργαστές

Κατά τον Udo Zolner, οι ψηφιακές μονάδες επεξεργασίας (DAFX) μπορούν να κατηγοριοποιηθούν ως εξής

- Filters and delays (resampling)
- Modulators and demodulators
- Non-linear processing
- Spatial effects
- Time-segment processing
- Time-frequency processing
- Source-filter processing
- Adaptive effects processing
- Spectral processing
- Time and frequency warping
- Virtual analog effects
- Automatic mixing
- Source separation

Βρόχοι πραγματικού χρόνου

Βρόχοι πραγματικού χρόνου (Live Looping)

Ηχογράφηση ηχητικών σημάτων ή μουσικών θεμάτων και αναπαραγωγή τους, με ενδεχόμενη παραμετροποίηση από τους μουσικούς, έτσι ώστε να κατασκευαστούν μουσικές δομές ή ηχητικά περιβάλλοντα (sound textures).

Τα πλεονεκτήματα του live-looping είναι:

- απουσία χρονικής κβαντοποίησης (temporal quantization), και ανάδειξη φυσικών μουσικών μοτίβων
- μπορεί να αξιοποιηθεί σε οποιοδήποτε μουσικό ιδίωμα, καθώς πραγματοποιείται με οποιαδήποτε είσοδο ηχητικού σήματος, ανεξαρτήτως της μετρικής της δομής (time-signature)
- οι μουσικοί μπορούν να αξιοποιήσουν προηχογραφημένα θέματα ή να επιλέξουν αυτοσχεδιαστικές προσεγγίσεις
- επιτρέπει την πραγματοποίηση γρήγορων μουσικών προσχεδίων (song's sketching)

Συστήματα Εναρμόνιση φωνής

Σύστημα εναρμόνισης φωνής (Harmonizer)

Μίξη ενός ηχητικού σήματος με ένα ή περισσότερα μετατονισμένα (pitch-shifted) στιγμιότυπα του αρχικού σήματος.

Έξυπνη αρμονία (Smart harmony, Intelligent Harmonization)

Η εναρμόνιση πραγματοποιείται από μία συχνότητα εισόδου και τα μελωδικά «συμφραζόμενα» (context)

Παράμετροι συστήματος εναρμόνιση φωνής του Polyfōnía

Το σύστημα εναρμόνισης του Polyfōnía, σχεδιάστηκε έτσι ώστε να ικανοποιεί τις εξής προδιαγραφές:

- να μπορεί μία φωνή να ενεργοποιείται και να απενεργοποιείται με μία απλή χειρονομία (λειτουργία διακόπτη), και όχι να παραμένει ενεργή μόνο όσο διατηρείται μία χειρονομία
- να μπορούν να δημιουργηθούν όλα τα πιθανά διαστήματα του δωδεκάφθογγου συστήματος χωρίς να απαιτείται η ύπαρξη οθόνης για τον έλεγχο της κατάστασης του συστήματος
- να μπορούν να ελεγχούν παράμετροι όπως η ένταση κάθε φωνής (volume control) και το τονικό ύψος (pitch shifting)

Επαυξημένα μουσικά όργανα

Τα επαυξημένα μουσικά όργανα (augmented musical instruments):

- δημιουργούνται με την τοποθέτηση αισθητήρων σε ακουστικά ή ηλεκτρικά όργανα
- επεκτείνουν τις δυνατότητες του εκάστοτε μουσικού οργάνου ελέγχοντας είτε δημιουργικές διεργασίες ή ψηφιακούς επεξεργαστές
- προϋποθέτουν εξοικείωση των μουσικών με τον τρόπο λειτουργίας τους

Οι πιο δημοφιλείς προσεγγίσεις όσον αφορά στα επαυξημένα όργανα είναι:

- η τοποθέτηση και χρήση αισθητήρων για την προσθήκη δυνατοτήτων ελέγχου, διευρύνοντας την «παραδοσιακή τεχνική» του οργάνου (για παράδειγμα η τοποθέτηση κουμπιών στο σώμα ενός μουσικού οργάνου)
- η τοποθέτηση αισθητήρων για τον εντοπισμό των τεχνικών κινήσεων του μουσικού.

- Steinberg, IC Air: μία επέκταση με ανίχνευση κίνησης για το Cubase
- Elena Jessop, VAMP: ελεγκτής σε μορφή γαντιού για το χειρισμό φωνητικών σε πραγματικό χρόνο
- Berthaut et al.: εικονικό περιβάλλον διάδρασης με χρήση τρισδιάστατων οπτικοακουστικών αντικειμένων
- Smallwood, Cook et al., Ημισφαιρικά ηχεία (hemispherical speakers): σύστημα 6 ηχείων που μιμούνται τον τρόπο με τον οποίο τα ακουστικά όργανα διαδίδουν τον ήχο σε όλες τις κατευθύνσεις



Συγκριτική Μελέτη Ανιχνευτών Κίνησης

- **Kinect της Microsoft (2010):** αισθητήρας ανίχνευσης κίνησης για κονσόλες παιχνιδιών και τους υπολογιστές με λειτουργικό Microsoft Windows. Απερρίφθει λόγω της διακοπής της κυκλοφορίας του, της παλιάς τεχνολογίας του και το χαρακτήρα του ως ανιχνευτή κίνησης με την ευρύτερη έννοια, χωρίς να δίνεται έμφαση στις κινήσεις των δακτύλων.
- **Αλγόριθμος Open Pose, Carnegie Mellon (2018):** το πρώτο σύστημα πραγματικού χρόνου που μπορεί να πραγματοποιήσει ανίχνευση κίνησης του ανθρώπινου σώματος, προσώπου, χεριών και ποδιών μέσω του εντοπισμού 153 σημείων, με είσοδο της εικόνας από την κάμερα ενός υπολογιστή. Απερρίφθη λόγω των υψηλών απαιτήσεων σε υπολογιστικούς πόρους σε περίπτωση χρήσης διαφορετικής κάρτας γραφικών από NVIDIA GPU ή AMD GPU (8 GB RAM).
- **Leap motion:** η αναγνώριση των κινήσεων πραγματοποιείται μέσω αλγορίθμων που δέχονται ως είσοδο τα ανεπεξέργαστα δεδομένα του αισθητήρα και αναλύουν τις δισδιάστατες εικόνες με στόχο την ανάκτηση της τρισδιάστατης αναπαράστασης. Επιλέχθηκε, ως ένας από τους πιο αξιόπιστους, σύγχρονους και οικονομικούς ελεγκτές, μικρών διαστάσεων, με χαμηλές απαιτήσεις υπολογιστικών πόρων (2 GB RAM) και δωρεάν διαθέσιμο πακέτο λογισμικού ανάπτυξης (Software Development Kit).

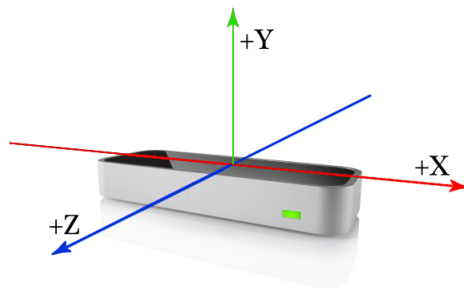
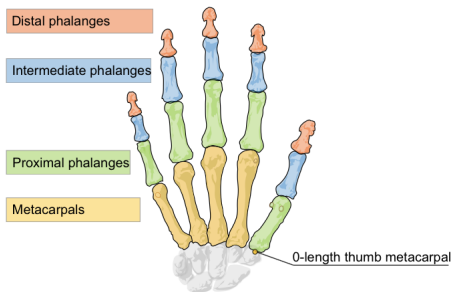
Leap Motion

- αποτελείται από δύο κάμερες και τρία υπέρυθρα leds: ανιχνεύουν τις ακτίνες μήκους κύματος 850 nanometers που βρίσκονται εκτός του ορατού φάσματος για τον άνθρωπο
- ευρυγώνιους φακούς του: μεγάλος χώρος διάδρασης σε σχήμα ανεστραμμένης πυραμίδας (8 cubic feet)
- ο ελεγκτής διαβάζει τα δεδομένα από τους αισθητήρες στη μνήμη του και πραγματοποιεί ρυθμίσεις ανάλυσης
- τα δεδομένα μεταφέρονται μέσω της θύρας USB, στο λογισμικό ανίχνευσης του Leap Motion.

Leap Motion

Το Leap Motion:

- είναι σε θέση να δίνει πληροφορία για κάθε δάκτυλο του αριστερού ή του δεξιού χεριού (thumb, index, middle, ring, pinky), και καθένα από αυτά αποτελείται από τέσσερα οστά (metacarpal, proximal phalanx, intermediate phalanx, distal phalanx), (εκτός του αντίχειρα που αποτελείται από τρία)
- μπορεί να εντοπίσει εργαλεία (tools) και να τα διακρίνει από τα δάκτυλα, ως μακρύτερα και λεπτότερα από αυτά
- ως εργαλεία ανιχνεύει αποκλειστικά λεπτά κυλινδρικά αντικείμενα
- αναγνωρίζει συγκεκριμένα μοτίβα κινήσεων
- έχει δεξιόστροφο, ορθογώνιο σύστημα συντεταγμένων του ελεγκτή είναι με σημείο αναφοράς το κέντρο της κορυφής του ελεγκτή



Leap Motion Skeletal Tracking in Max

Max/Msp Patch «Leap Motion Skeletal Tracking in Max»

Εφαρμογή ανοιχτού κώδικα για το Max/Msp, βασισμένο στο Leap Motion SDK V2 Skeletal Tracking Beta και (εμπνευσμένο από το aka.leapmotion του Masayuki Akamatsu). Αναπτύχθηκε από το Ινστιτούτο IRCAM (Institut de Recherche et Coordination Acoustique/Musique).

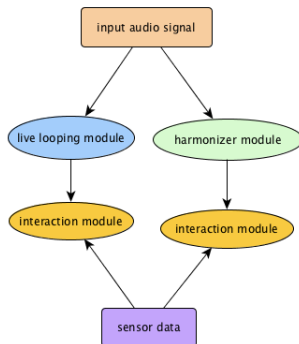
Τα τεχνικά χαρακτηριστικά του Patch περιλαμβάνουν:

- Full Skeletal Tracking
- Automatic left/right hand routing
- Automatic finger routing
- Build-in Leap Motion gesture recognition
- Simple record/replay based on MuBu

Polyfōnía: Αρχιτεκτονική

Το σύστημα Polyfōnía αποτελείται από τα εξής υποσυστήματα:

- ηχητικών βρόχων πραγματικού χρόνου (live looping module)
- εναρμόνισης φωνής (harmonizer module)
- αλληλεπίδρασης (interaction - module)



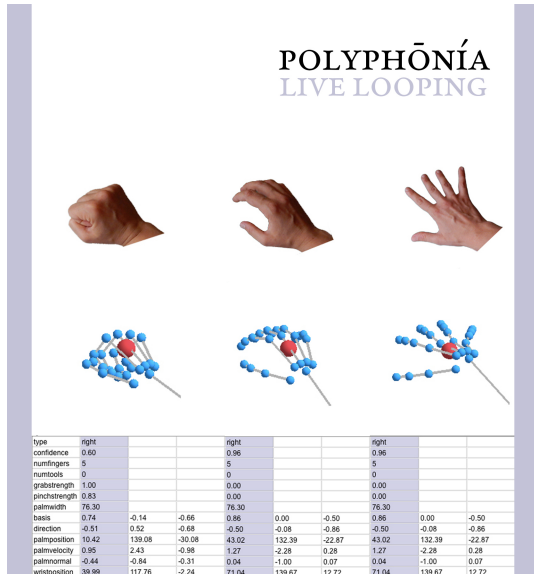
Polyfōnía: Harmonizer Module

Το σύστημα εναρμόνισης φωνής πραγματοποιεί μετατόνιση (pitchshifting) του ηχητικού σήματος εισόδου μέσω του αντικειμένου του Max/MSP «gizmo».

Το ηχητικό σήμα εισόδου:

- 1 δρομολογείται μέσω του αντικειμένου «gate» του Max/Msp σε 12 κανάλια
- 2 μετατρέπεται σε ψηφιακή μορφή μέσω του αντικειμένου «adc» του Max/Msp ώστε να μπορεί να δωθεί ως είσοδος στους ψηφιακούς επεξεργαστές
- 3 κάθε ένα από τα 12 κανάλια δρομολογείται στο αντικείμενο «gizmo» το οποίο πραγματοποιεί τη μετατόνιση με είσοδο x , όπου x ο αριθμός των ημιτονίων ($x \in \mathbb{N}, x \in [1, 12]$)
- 4 η έξοδος από το gizmo δρομολογείται σε 12 διακόπτες (κάθε ένας από τους οποίους ενεργοποιεί ή απενεργοποιεί το εκάστοτε κανάλι) αλλά και στο σήμα εξόδου

Polyfōnía: Live Looping Module



Polyfōnía: Live Looping Module

