

## **Αναγνώριση κατασκευαστικών χαρακτηριστικών και εξαγωγή μουσικής πληροφορίας από ηχογραφήσεις μπεντίρ.**

Νικόλαος Βρύζας<sup>1,\*</sup>, Βασίλειος Μπουντουράκης<sup>2</sup>, Αντώνης Παγώνης<sup>3</sup>

<sup>1</sup>Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

<sup>2</sup>Πανεπιστήμιο Aalto

<sup>3</sup>Pagonis Percussions

\*nvryzas@auth.gr

### **ΠΕΡΙΛΗΨΗ**

*Το μπεντίρ είναι ένα παραδοσιακό μεμβρανόφωνο κρουστό όργανο, η κατασκευή και η τεχνική του οποίου βασίζεται στην προφορική παράδοση. Η παρούσα εργασία αποσκοπεί στη διερεύνηση εργαλείων, αρχιτεκτονικών και μοντέλων για την αυτοματοποιημένη εξαγωγή πληροφορίας σχετικά τα κατασκευαστικά χαρακτηριστικά και τη μουσική εκτέλεση μέσα από καταγραφές μπεντίρ. Χρησιμοποιείται ένα σύνολο ηχογραφήσεων από μπεντίρ διαφορετικών γεωμετρικών χαρακτηριστικών και υλικών κατασκευής, παγμένων με διαφορετικές τεχνικές και είδη χτυπημάτων. Αξιολογούνται μοντέλα αναγνώρισης των διαφορετικών μουσικών φθόγγων και κατασκευαστικών χαρακτηριστικών, ενώ προτείνεται ένα σύστημα αυτόματης μουσικής μεταγραφής ηχογραφήσεων και η χρήση του συνόλου δεδομένων ως μια βιβλιοθήκη ήχων για τον προγραμματισμό εικονικού μουσικού οργάνου.*

### ***Construction attributes recognition and music information retrieval from frame drum recordings***

#### **ABSTRACT**

*The frame drum (bendir) is a traditional membrane percussion instrument, the construction and technique of which is based on oral tradition. This work aims to investigate tools, architectures and models for the automated extraction of information about the construction attributes and the musical performance through bendir recordings. A set of recordings of frame drums of different geometries and construction materials, played with different techniques and types of strikes, is used. Models for the recognition of different musical tones and construction attributes are evaluated, while a system for automatic music transcription of recordings and the use of the dataset as a sound library for programming virtual musical instruments are proposed.*

## 1. Εισαγωγή

Το μπεντίρ είναι ένα παραδοσιακό μουσικό όργανο που συναντάται στη μουσική παράδοση πολλών περιοχών της Μεσογείου, Ανατολικής Ευρώπης και Ασίας. Πρόκειται για ένα κυκλικό μεμβρανόφωνο κρουστό. Αποτελείται από ένα κυκλικό στεφάνι, κατά κανόνα φτιαγμένο από ξύλο, και μία δερμάτινη (ή συνθετική σε πιο πρόσφατες προσεγγίσεις) μεμβράνη. Τα κατασκευαστικά χαρακτηριστικά του μπεντίρ μπορεί να απέχουν ως προς τις διαστάσεις (διάμετρο και βάθος του στεφανιού), την επιλογή είδους ξύλου, και το είδος του δέρματος. Οι τεχνικές της κατασκευής του στηρίζονται σχεδόν κατ' αποκλειστικότητα στην προφορική παράδοση.

Η παρούσα εργασία εντάσσεται σε μία προσπάθεια αφενός συστηματοποίησης της διαδικασίας κατασκευής του μπεντίρ με τη χρήση νέων τεχνολογικών προσεγγίσεων, αφετέρου την αξιοποίηση τεχνικών μηχανικής μάθησης για την αυτοματοποίηση της εξαγωγής μουσικής πληροφορίας από ηχογραφήσεις μουσικών εκτελέσεων σε μπεντίρ. Το κίνητρο είναι η σύνδεση της λαϊκής παράδοσης με τις νέες τεχνολογικές δυνατότητες, η διερεύνηση των παραδοσιακών αντιλήψεων σχετικά με τις τεχνικές κατασκευής του οργάνου και η επέκταση των δυνατοτήτων του για τις ανάγκες ενός σύγχρονου μουσικού.

Η τυπική διαδικασία ακουστικής μελέτης μουσικών οργάνων περιλαμβάνει παράλληλα με τις ακουστικές μετρήσεις την υποκειμενική αξιολόγηση μέσω της αντιληπτικής ικανότητας χρηστών [1]. Σε αυτή την περίπτωση είναι κομβική η θέσπιση μιας κοινής γλώσσας με περιγραφείς που μπορούν να είναι μετρήσιμοι από τους μηχανικούς ήχου και αντιληπτοί από τους μουσικούς ή τους ακροατές [2]. Κοινά ακουστικά χαρακτηριστικά μπορούν να παρουσιάζονται με διαφορετική ορολογία στα πειράματα ακρόασης [3]. Τα πειράματα ακρόασης είναι πολύ σημαντικά γιατί σε πολλές περιπτώσεις φαίνεται η αδυναμία ειδικών του πεδίου να αντιληφθούν μεταβολές σε εύκολα μετρήσιμα μεγέθη, όπως στις θεμελιώδεις συχνότητες κρουστών οργάνων [4].

Το πεδίο της εξαγωγής μουσικής πληροφορίας από μουσικές ηχογραφήσεις, αφορά την αυτόματη δημιουργία παρτιτούρας από μουσικές εκτελέσεις. Στην περίπτωση των κρουστών οργάνων, ιδιαίτερο ενδιαφέρον έχει, εκτός από τη χρονική εύρεση των ηχητικών συμβάντων κρούσης (onsets) [5], η αναγνώριση του συγκεκριμένου κρουστού οργάνου σε κάθε χτύπημα (Automatic Drum Transcription - ADT) [5]-[8]. Σε ένα τυπικό σετ τυμπάνων (drum set), υπάρχει ένας συγκεκριμένος αριθμός διαφορετικών κρουστών που πρέπει να αναγνωριστούν, τα οποία αντιστοιχούν σε διαφορετικές κλάσεις σε ένα μοντέλο ταξινόμησης μηχανικής μάθησης [6], [7]. Σε πολλές περιπτώσεις, είναι χρήσιμο να υπάρχουν ευρύτερα συστήματα ταξινόμησης που να επιτρέπουν την αναγνώριση ενός εκτεταμένου σετ κρουστών. Τα συνελκτικά νευρωνικά δίκτυα αποτελούν μια τυπική επιλογή για την αναγνώριση κρουστών [5]-[8]. Στην περίπτωση του μπεντίρ, όπως θα αναλυθεί σε επόμενη ενότητα, το σχήμα ταξινόμησης δεν αφορά έναν αριθμό από διακριτά μουσικά όργανα, αλλά διαφορετικές τεχνικές παιξίματος του ίδιου του οργάνου.

Η εργασία δομείται ως εξής: Στην ενότητα 2 παρουσιάζονται τα βασικά χαρακτηριστικά του οργάνου και της τεχνικής του που είναι υπό διερεύνηση, καθώς και το σύνολο δεδομένων που έχει δημιουργηθεί για την ανάλυσή τους. Στην ενότητα 3 παρουσιάζονται οι προτεινόμενες εφαρμογές και τα εργαλεία, καθώς και

η αξιολόγησή τους. Στην ενότητα 4 συνοψίζονται τα συμπεράσματα και παρατίθενται οι μελλοντικές ερευνητικές επεκτάσεις

## 2. Βασικές χαρακτηριστικές παράμετροι σε μία καταγραφή μπεντίρ

Κατά την ανάλυση του προβλήματος, αναγνωρίστηκαν δύο βασικοί άξονες που καθορίζουν την ακουστική συμπεριφορά σε μία καταγραφή μπεντίρ. Το πρώτο αφορά τα κατασκευαστικά χαρακτηριστικά του οργάνου και το δεύτερο τον τρόπο παιξίματος και τις διαφορετικές τεχνικές.



Σχήμα 2.1 Δύο μπεντίρ με διαφορετικά κατασκευαστικά χαρακτηριστικά

### 2.1 Κατασκευαστικά χαρακτηριστικά και παράμετροι

Καθώς η κατασκευή του μπεντίρ βασίζεται στην προφορική παράδοση και αποτελεί στοιχείο πολιτιστικής κληρονομιάς μεταξύ διαφορετικών λαών και περιοχών, δεν υπάρχει κάποια αυστηρά καθορισμένη παραδοχή για τα κατασκευαστικά χαρακτηριστικά του οργάνου. Στην πιο απλή εκδοχή του, πρόκειται για ένα ξύλινο στεφάνι στο οποίο είναι στερεωμένη μία δερμάτινη μεμβράνη. Σε πολλές περιπτώσεις, είναι πιθανό το στεφάνι να έχει υποδοχές στις οποίες τοποθετούνται μικρά μεταλλικά κύμβαλα, οι ζίλιες. Οι τρόποι πρόσδεσης της μεμβράνης στο στεφάνι μπορούν να διαφέρουν και επηρεάζουν τον τρόπο κουρδίσματος του οργάνου σε συγκεκριμένη τονικότητα.

Στην απλή περίπτωση του μπεντίρ με μεμβράνη, χωρίς ζίλιες, τα βασικά χαρακτηριστικά που μπορούν να παραμετροποιηθούν αφορούν τη γεωμετρία και τα υλικά κατασκευής. Ως προς τη γεωμετρία, μπορούμε να έχουμε όργανα με διαφορετική διάμετρο της μεμβράνης και βάθος του στεφανιού. Τα υλικά αφορούν τα ξύλα που θα χρησιμοποιηθούν για την κατασκευή του στεφανιού και το δέρμα της μεμβράνης. Σε πολλά σύγχρονα όργανα, μπορεί να επιλέγονται και συνθετικά

υλικά. Οι έμπειροι οργανοπαίχτες έχουν προτιμήσεις αξιολογούν τα συγκεκριμένα χαρακτηριστικά ως προς την εργονομία, το οπτικό, αλλά και το ακουστικό αποτέλεσμα.

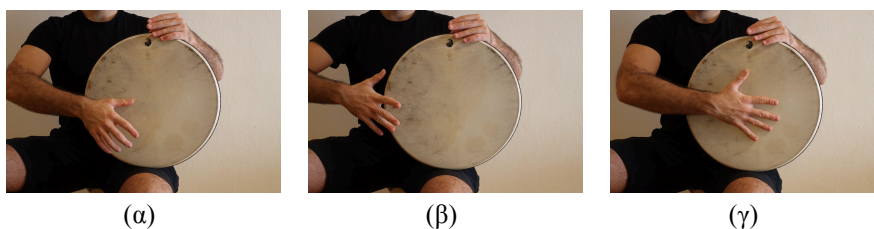
### 2.2 Τεχνικές παιζίματος

Μία παράμετρος που διαφοροποιεί τον ήχο του μπεντίρ είναι η θέση και το αντίστοιχο κράτημα του οργάνου. Στην όρθια θέση, το όργανο κρατιέται από τη λαβή σε ελεύθερη θέση με το ένα χέρι, ενώ το άλλο χρησιμοποιείται για τα χτυπήματα. Στην καθιστή θέση, το όργανο εφάπτεται στο σώμα του/της οργανοπαίχτη/τριας, πράγμα που επηρεάζει και την ακουστική του απόκριση, όπως φαίνεται στο Σχήμα 2.2.



Σχήμα 2.2 Κράτημα του μπεντίρ σε (α) όρθια και (β) καθιστή θέση

Το βασικό λεξιλόγιο του οργάνου αποτελείται από τρία χαρακτηριστικά χτυπήματα, τα οποία έχουν διακριτό ήχο και χρησιμοποιούνται για την εκτέλεση μουσικών μοτίβων, το *ντουμ*, το *τεκ* και το *σλαπ*. Το χτύπημα *τεκ* είναι μία κρούση κοντά στο στεφάνι του οργάνου, ενώ το χτύπημα *ντουμ* είναι μια κρούση ενδιάμεσα στο στεφάνι και το κέντρο της μεμβράνης. Στο χτύπημα *σλαπ*, όλη η επιφάνεια της παλάμης χτυπάει τη μεμβράνη και παραμένει σε επαφή με τη μεμβράνη μετά το χτύπημα. Στο Σχήμα 2.3 φαίνονται τα τρία χαρακτηριστικά χτυπήματα.



Σχήμα 2.3 Χτυπήματα (α) ντουμ (β) τεκ (γ) σλαπ

### 2.3 Σύνολο δεδομένων καταγραφών δειγμάτων

Για τις ανάγκες της έρευνας σχετικά με την εξαγωγή πληροφορίας για τα κατασκευαστικά χαρακτηριστικά και τη μουσική εκτέλεση, δημιουργήθηκε ένα

σύνολο δεδομένων με πολλαπλές ετικέτες, που μπορεί να αξιοποιηθεί σε διαφορετικά προβλήματα. Στην πρώτη έκδοση του συνόλου, ηχογραφήθηκαν έξι μπεντίρ διαφορετικών χαρακτηριστικών. Τα μπεντίρ ηχογραφήθηκαν στο στούντιο οπτικοακουστικής παραγωγής του Εργαστηρίου Ηλεκτρονικών ΜΜΕ του Τμήματος Δημοσιογραφίας του ΑΠΘ χρησιμοποιώντας ταυτόχρονα δύο μικρόφωνα, ένα δυναμικό μικρόφωνο Shure SM57 και ένα πυκνωτικό Rode NT1000. Κάθε μπεντίρ ηχογραφήθηκε και στις δύο θέσεις παιξίματος με περίπου 50 δείγματα για κάθε είδος χτυπήματος. Στη συνέχεια η βάση εμπλουτίστηκε με επιπλέον ηχογραφήσεις σε διαφορετικούς χώρους και με διαφορετικές συσκευές ηχογράφησης. Έτσι δημιουργήθηκε μία βάση ~15000 δειγμάτων πολλαπλών ετικετών, επισημειωμένα ως προς α) τη διάμετρο β) το βάθος γ) το ξύλο του στεφανιού δ) τη θέση παιξίματος (όρθια ή καθιστή), ε) το χτύπημα (ντουμ, τεκ, σλαπ) και, στ) το μικρόφωνο καταγραφής.

### 3. Τεχνικές και εφαρμογές εξαγωγής πληροφορίας από τις καταγραφές

#### 3.1 Αναγνώριση κατασκευαστικών χαρακτηριστικών

Το σύνολο δεδομένων που δημιουργήθηκε αξιοποιήθηκε για τη διεξαγωγή πειράματος ακρόασης MUSHRA [9]. Εννιά εθελοντές με ειδικευση είτε στη μουσική είτε στην Ηλεκτροακουστική απάντησαν ένα σετ ερωτήσεων σε σχέση με την αναγνώριση τριών βασικών χαρακτηριστικών που συνηθίζονται στη βιβλιογραφία στα πειράματα υποκειμενικής αξιολόγησης (brightness, reverberance, harmonicity). Η μόνη παράμετρος που εμφάνισε στατιστικά σημαντική επίδραση στην κατανομή των απαντήσεων είναι η διάμετρος [9].

#### 3.2 Μεταγραφή μουσικής πληροφορίας και εικονικά μουσικά όργανα

Η δεύτερη εφαρμογή που προτείνεται στο πλαίσιο της συγκεκριμένης έρευνας αφορά τη μεταγραφή της μουσικής πληροφορίας από καταγραφές μουσικών εκτελέσεων. Στην πράξη μας ενδιαφέρει:

- Ο χρονικός εντοπισμός των ηχητικών συμβάντων - χτυπημάτων (onset detection)
- Για κάθε χτύπημα, η αναγνώριση του είδους χτυπήματος (ντουμ, τεκ, σλαπ)
- Η αναγνώριση της θέσης παιξίματος (όρθια καθιστή). Αυτό πραγματοποιείται για κάθε χτύπημα αλλά αναμένεται να είναι το ίδιο για το σύνολο των χτυπημάτων (χρονική ολοκλήρωση).
- Η αναγνώριση του κουρδίσματος (pitch) του οργάνου. Όμοια, πραγματοποιείται για κάθε χτύπημα αλλά αναμένεται να είναι το ίδιο για το σύνολο των χτυπημάτων (χρονική ολοκλήρωση).

Για την αναγνώριση των χτυπημάτων και της θέσης παιξίματος έχουν εκπαιδευτεί δύο μοντέλα διδιάστατων Συνελκτικών Νευρωνικών Δικτύων που δέχονται ως είσοδο τα φασματογραφήματα κλίμακας Μελ 56 ζωνών για κάθε χτύπημα. Η ακρίβεια αναγνώρισης είναι της τάξης 94.5% για την αναγνώριση χτυπήματος και 89.2% για την αναγνώριση θέσης [10]. Για την περίπτωση θέσης, προτείνεται η χρονική ολοκλήρωση για το σύνολο της καταγραφής.

Η εξαγωγή της παραπάνω πληροφορίας επιτρέπει δύο βασικές εφαρμογές:

#### 1. Αυτόματη εξαγωγή μουσικής παρτιτούρας από καταγραφή

Μπορεί να καταγραφούν τα χτυπήματα και οι χρόνοι εντοπισμού για τη μεταγραφή της μουσικής εκτέλεσης σε μουσική σημειογραφία

#### 2. Χρήση των δειγμάτων της βάσης ως εικονικό όργανο

Ουσιαστικά ο χρήστης μπορεί να χρησιμοποιήσει ένα οποιοδήποτε μπεντίρ ως αισθητήρα εισόδου και έπειτα να επιλέξει από τη βάση την αναπαραγωγή της μουσικής εκτέλεσης με δείγματα από όργανα συγκεκριμένων χαρακτηριστικών (γεωμετρικών, υλικών). Ένα επιπλέον όφελος είναι ότι μπορεί έτσι σε έναν οποιοδήποτε χώρο, με όργανο οποιασδήποτε ποιότητας κατασκευής και οποιοδήποτε μικρόφωνο να δημιουργήσει ένα τελικό αρχείο ήχου με ηχογραφήσεις σε στουντιακές συνθήκες, καλής ακουστικής χώρου και χρήσης επαγγελματικού εξοπλισμού, με όργανα επαγγελματικών προδιαγραφών.

### 4. Συμπεράσματα και μελλοντικές επεκτάσεις

Παρουσιάστηκε μια μεθοδολογία για την αναγνώριση της επίδρασης διαφορετικών κατασκευαστικών χαρακτηριστικών στην ακουστική συμπεριφορά του παραδοσιακού μεμβρανόφωνου κρουστού οργάνου μπεντίρ, καθώς και μοντέλα αναγνώρισης των διαφορετικών χτυπημάτων της μουσικής εκτέλεσης. Προτείνεται ένα σύστημα για τη μεταγραφή της μουσικής πληροφορίας καταγραφών μπεντίρ, καθώς και τη χρήση του ως είσοδο για τη σύνθεση μουσικής με χρήση εικονικού μουσικού οργάνου.

Στις μελλοντικές ερευνητικές κατευθύνσεις βρίσκεται η δημιουργία ενός γραφικού περιβάλλοντος που απευθύνεται σε κατασκευαστές και έχει ως σκοπό την υποστήριξη του σταδίου του ελέγχου ποιότητας, και ενός περιβάλλοντος για μουσικούς που να μπορούν να αξιοποιήσουν τις προτεινόμενες δυνατότητες. Απαιτείται ένα στάδιο διερεύνησης της αξιοπιστίας αναγνώρισης σε καταγραφές πραγματικού κόσμου σε διαφορετικές συνθήκες. Για την ευρωστία του συστήματος σχεδιάζεται η διεύρυνση της βάσης με ηχογραφήσεις σε διαφορετικές συνθήκες, αλλά και μέσω συνθετικής επαύξησης των υπαρχόντων δεδομένων.

### 5. Αναφορές

[1] Fritz, C. and Dubois, D., “Perceptual evaluation of musical instruments: State of the art and methodology,” *Acta Acustica united with Acustica*, **101**(2), pp. 369–381, (2015)

[2] Fritz, C., Blackwell, A. F., Cross, I., Woodhouse, J., and Moore, B. C., “Exploring violin sound quality: Investigating English timbre descriptors and correlating resynthesized acoustical modifications with perceptual properties,” *The Journal of the Acoustical Society of America*, **131**(1), pp. 783–794 (2012)

[3] Brent, W., *Physical and perceptual aspects of percussive timbre*, University of California, San Diego (2010)

[4] Toulson, R. and Hardin, M., “Evaluating the accuracy of musicians and sound engineers in performing a common drum tuning exercise,” in *Audio Engineering Society Convention 149*, Audio Engineering Society (2020)

- [5] Wu, C. W., Dittmar, C., Southall, C., Vogl, R., Widmer, G., Hockman, J., ... & Lerch, A. "A review of automatic drum transcription." *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, **26(9)**, 1457-1483 (2018)
- [6] Vogl, R., Widmer, G., & Knees, P. "Towards multi-instrument drum transcription." arXiv preprint arXiv:**1806.06676**. (2018)
- [7] Cartwright, M., & Bello, J. P. "Increasing drum transcription vocabulary using data synthesis." In *Proc. International Conference on Digital Audio Effects (DAFx)* (pp. 72-79) (2018, September)
- [8] Blaszkę, M., & Kostek, B. "Musical instrument identification using deep learning approach." *Sensors*, **22(8)**, 3033 (2022)
- [9] Vryzas, N., Bountourakis, V., & Pagonis, A. "Towards a systematic approach to the design of the traditional bendir drum." In *Audio Engineering Society Convention 154*. Audio Engineering Society (2023)
- [10] Vryzas, N., Bountourakis, V., & Pagonis, A. "From live frame drum performance to music notation: a data-driven approach." In *Audio Engineering Society Convention 156*. Audio Engineering Society (2024)