

## Διαδικτυακή Πλατφόρμα Εκτίμησης Συναισθημάτων από Ηχητικά Δεδομένα

Κωνσταντίνος Δρόσος  
Τμ. Τεχνών Ήχου & Εικόνας  
Ιόνιο Πανεπιστήμιο  
kdrosos@ionio.gr

Ανδρέας Γιαννακούλοπουλος  
Τμ. Τεχνών Ήχου & Εικόνας  
Ιόνιο Πανεπιστήμιο  
agiannak@ionio.gr

Ανδρέας Φλώρος  
Τμ. Τεχνών Ήχου & Εικόνας  
Ιόνιο Πανεπιστήμιο  
floros@ionio.gr

### ΠΕΡΙΛΗΨΗ

Μία από τις κύριες διεργασίες στο πεδίο αναγνώρισης συναισθημάτων από ηχητικά δεδομένα είναι η συλλογή της συναισθηματικής επισήμανσης. Λόγω της φύσης των δεδομένων αλλά και του πλήθους αυτών, κάθε τέτοια διεργασία μπορεί να είναι πολύπλοκη οργανωτικά και χρονοβόρος διαδικασία. Στην παρούσα εργασία παρουσιάζεται μία διαδικτυακή πλατφόρμα για την συλλογή συναισθηματικών επισημάνσεων στην οποία γίνεται χρήση του ευρύτερα καθιερωμένου *Self Assessment Manikin (SAM)*. Η πλατφόρμα επιτρέπει στον χρήστη να αναπαράγει τους προς αξιολόγηση ήχους, καταγράφει τα αποτελέσματα και παρέχει επισημάνσεις σχετικά με κύρια σημεία της διεργασίας, όπως την επανάληψη του πειράματος από τον ίδιο χρήστη.

### Web Based Platform for Emotion Evaluation from Audio Data

#### ABSTRACT

*One of the main process in emotions estimation from audio data is the ground truth data collection from human listeners. Due to the nature of the data (sound data) and the amount of them, any such process may endure many hours. In the proposed paper, we present a web based platform for the collection of the ground truth data using the well-established self assessment manikin (SAM). The platform allows the user to reproduce the associated sounds on demand, records the answers and offers annotations related to key features, like the repeat of the process by the same user.*

## Εισαγωγή

Τα τελευταία χρόνια έχει αναπτυχθεί η έρευνα γύρω από την ανίχνευση συναισθημάτων τόσο σε μουσική, όσο και σε ηχητικά γεγονότα εν γένει. Ήδη υπάρχουν έρευνες σχετικά με την ενσωμάτωση των αποτελεσμάτων σε πληθώρα εφαρμογών, όπως μία νέα κατηγοριοποίηση της μουσικής που δεν θα ακολουθεί την παραδοσιακή βάση καλλιτέχνη, είδους ή/και συγκροτήματος [1], ή η “έξυπνη” επιλογή μουσικών κομματιών βασισμένη σε συναισθήματα [2] ή, ακόμα, και εφαρμογές για “έξυπνα” σπίτια [3]. Κάτι που είναι κοινό σε όλες τις έρευνες σχετικά με την ανίχνευση συναισθημάτων είναι η επισήμανση του συναισθήματος στα δεδομένα που χρησιμοποιούνται. Αυτή μπορεί είτε να είναι πρωτογενής, ερωτώντας απευθείας τους συμμετέχοντες σε τέτοια πειράματα, είτε δευτερογενής, βασισμένη σε ήδη επισημασμένα δεδομένα.

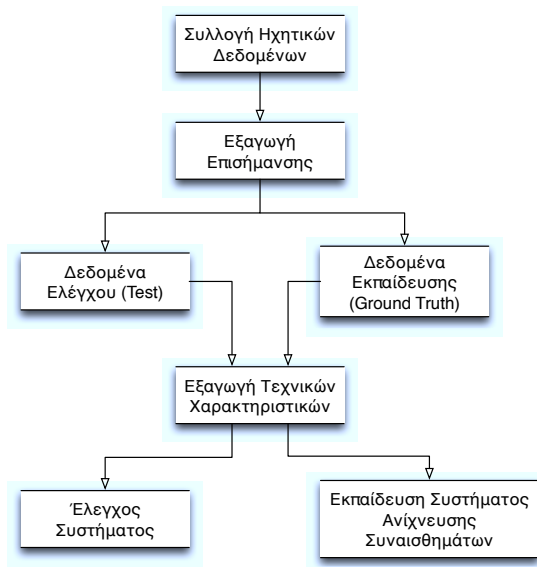
Στην περίπτωση της πρωτογενούς έρευνας, η διαδικασία της επισήμανσης των ηχητικών δεδομένων έχει ιδιαίτερη βαρύτητα μιας και θα αποτελέσει την βάση για την ομάδα δεδομένων που θα χρησιμοποιηθούν για την εκπαίδευση του συστήματος αναγνώρισης. Λόγω της φύσης της διαδικασίας αυτής, απαιτείται χρόνος και φυσική παρουσία των συμμετεχόντων αλλά και των διοργανωτών. Έτσι, η συλλογή το μέγιστου δυνατού πλήθους απαντήσεων δυσχεραίνεται. Με την ανάπτυξη όμως νέων μεθόδων επισήμανσης των ηχητικών δεδομένων, αλλά και την χρήση των υφιστάμενων τεχνολογιών διαδικτύου, φαίνεται ότι είναι δυνατή η κατάργηση των ορίων που εισάγονται, κυρίως, από την απαίτηση για φυσική παρουσία. Όμως, αν και είναι χρήσιμη η διεξαγωγή ενός απομακρυσμένου πειράματος, θα πρέπει να δοθεί ιδιαίτερη προσοχή στην αντικατάσταση, στο μέτρο που είναι εφικτό και δυνατό, των διοργανωτών του πειράματος.

Στην παρούσα εργασία παρουσιάζεται μία διαδικτυακή πλατφόρμα για πειράματα εκτίμησης συναισθημάτων από ηχητικά δεδομένα με την χρήση της μεθόδου Self Assessment Manikin (SAM). Η πλατφόρμα έχει όλες τις δυνατότητες ελέγχου των διαδικασιών του χρήστη που θα μπορούσαν, δυνητικά, να αλλοιώσουν τα αποτελέσματα του πειράματος, αλλά και προσφέρει άμεση προσπέλαση των αποτελεσμάτων και αντικατάσταση των δεδομένων, σε περίπτωση επαναχρησιμοποίησης. Η υπόλοιπη εργασία είναι διαρθρωμένη ως εξής: στην Ενότητα 1 παρουσιάζονται οι μέθοδοι επισήμανσης ηχητικών δεδομένων καθώς και τα αντίστοιχα συναισθηματικά μοντέλα, ενώ στην Ενότητα 2 παρουσιάζεται η αναπτυχθείσα πλατφόρμα. Τέλος, στην Ενότητα 3 αναπτύσσονται τα συμπεράσματα.

### 1 Διαδικασίες επισήμανσης ηχητικών δεδομένων και συναισθηματικά μοντέλα

Μία κοινή οργάνωση των ερευνών στο πεδίο της ανίχνευσης συναισθημάτων από ηχητικά μοντέλα είναι αυτή που παρουσιάζεται στο Σχήμα 1.1. Όπως μπορεί να φανεί, τα δεδομένα που θα χρησιμοποιηθούν για την διαδικασία ανίχνευσης θα επισημανθούν, σε κάθε περίπτωση (εκπαίδευσης ή ελέγχου), για την εκμείωση του συναισθήματος που εγείρουν στον ακροατή. Η διαδικασία αυτή μπορεί να παρακαμφθεί εάν χρησιμοποιηθούν ήδη επισημασμένα δεδομένα. Όμως, αν και τέτοια δεδομένα θα μπορούσαν να επιταχύνουν την ολοκλήρωση του πειράματος, η υλοποίηση αυτού εξαρτάται από παράγοντες που δεν μπορούν να ελεγχθούν από

τους διοργανωτές. Ο κυριότερος, ειδικά στην περίπτωση των ηχητικών γεγονότων, είναι η διαθεσιμότητα τέτοιων δεδομένων, όπου, ενδέχεται, να μην μπορεί να καλύψει τις ανάγκες του πειράματος.



Σχήμα 1.1 Η διαδικασία ανίχνευσης συναισθημάτων από ηχητικά δεδομένα

### 1.1 Διαδικασίες επισήμανσης

Η επισήμανση των ηχητικών δεδομένων πραγματοποιείται με την παρουσίαση, στον ακροατή, ενός πλήθους από διαθέσιμες επιλογές. Η επιλογή μίας εξ αυτών, ή και περισσότερων, αναλόγως το μοντέλο συναισθημάτων που χρησιμοποιείται, είναι και αυτή που χαρακτηρίζει το συναίσθημα που εγείρεται στον ακροατή από το προς εξέταση ηχητικό αρχείο. Σε κάθε περίπτωση όμως, ο/οι διοργανωτής-ες του πειράματος θα πρέπει να καθοδηγήσουν τον χρήστη σχετικά με την σημασία των διαθέσιμων επιλογών.

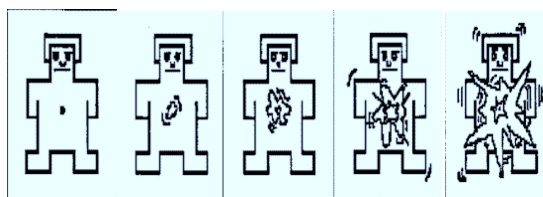
Σε περιπτώσεις που χρησιμοποιούνται απλές λεκτικές περιγραφές, η επεξήγηση είναι αρκετά απλή και μπορεί να μην χρειαστεί. Όμως αυτό σημαίνει την υιοθέτηση συγκεκριμένων μοντέλων συναισθημάτων με ό,τι περιορισμούς αυτό συνεπάγεται.

Έτσι, σε πειράματα από διάφορους κλάδους, όπως αυτός των νευροεπιστημών, μπορεί κανείς να συναντήσει λεκτικές περιγραφές όπως: Χαρά και Φόβος [4]. Ή σε αντίστοιχες έρευνες στο πεδίο του Music Emotion Recognition (MER) και της μουσικής ψυχολογίας (Music Psychology) απαντώνται τόσο οι προηγούμενες λεκτικές περιγραφές όσο και χρήση επιθέτων, όπως vigorous, graceful, soothing, dreaming [5].

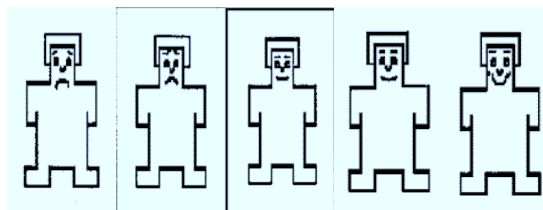
Παρόλα αυτά, δεν υπάρχουν κοινές λεκτικές περιγραφές για την επισήμανση. Γεγονός που συντελεί σε μία μη συνέχεια των αποτελεσμάτων μεταξύ των ερευνών, λόγω των διαφορετικών, έστω και σε ελάχιστο βαθμό, νοημάτων που υπάρχουν σε συνώνυμες λέξεις [6].

Στην εργασία [7] παρουσιάζεται ένας διαφορετικός τρόπος επισήμανσης. Αυτός ονομάζεται Self Assessment Manikin (SAM). Χρησιμοποιεί σειρές από κατάλληλες εικόνες, αντί για λεκτικές περιγραφές, και ο “χρήστης” καλείται να επιλέξει είτε μία από αυτές, είτε δύο συνεχόμενες, εάν η συναισθηματική του κατάσταση αντικατοπτρίζεται επακριβώς από μία ή είναι μεταξύ δύο εικόνων, αντιστοίχως.

Το SAM έχει χρησιμοποιηθεί ήδη σε μία πληθώρα ερευνών και μπορεί να παράγει επιστημονικά ικανές να χρησιμοποιηθούν με πλήθος διαφορετικά μοντέλα συναισθημάτων. Οι εικόνες για τις δύο από τις τρεις σειρές, όπου είναι και αυτές που θα χρησιμοποιηθούν στην πλατφόρμα που παρουσιάζεται, απεικονίζονται στα Σχήματα 1.2 και 1.3.



Σχήμα 1.2 Η σειρά εικόνων του SAM για την κατηγορία απαντήσεων *Arousal*



Σχήμα 1.3 Η σειρά εικόνων του SAM για την κατηγορία απαντήσεων *Valence*

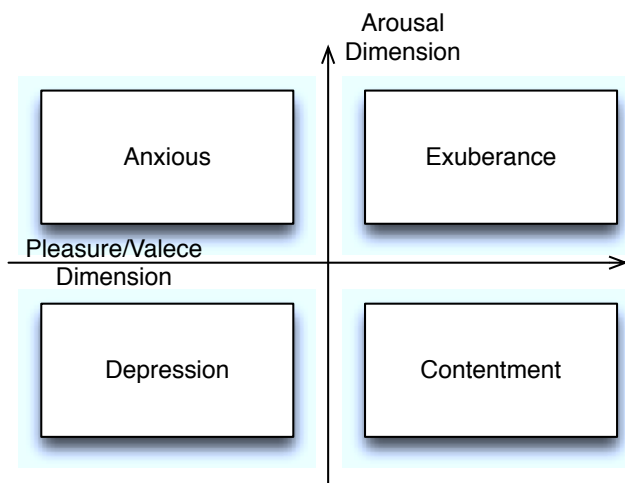
### 1.2 Συναισθηματικά μοντέλα

Τα συναισθηματικά μοντέλα μπορούν να διαχωριστούν σε δύο “αφηρημένες” (abstract) κατηγορίες. Τα διακριτά και τα συνεχή [8]. Στα μεν πρώτα, ανήκουν μοντέλα τα οποία χρησιμοποιούν συγκεκριμένες λέξεις για την περιγραφή των συναισθημάτων, όπως το μοντέλο των βασικών συναισθημάτων (Basic Emotions) [9] ή το μοντέλο της λίστας επιθέτων (List of Adjectives) [10].

Όπως μπορεί να φανεί, τα μοντέλα αυτά συνάδουν με τις μεθόδους επισήμανσης που χρησιμοποιούν συγκεκριμένες λεκτικές περιγραφές, πράγμα που σημαίνει ότι το πρόβλημα της συνέχειας μεταξύ των αποτελεσμάτων από διαφορετικές έρευνες παραμένει. Εντούτοις, πολλές έρευνες, και ειδικά στο πεδίο των νευροεπιστημών, τείνουν να χρησιμοποιούν το μοντέλο των βασικών συναισθημάτων ενώ έρευνες στο πεδίο της μουσικής ψυχολογίας (Music Psychology) το μοντέλο της λίστας επιθέτων.

Στον αντίποδα των διακριτών μοντέλων βρίσκονται τα συνεχή. Σε αυτά, γίνεται χρήση ενός ν-διάστατου χώρου με κάθε διάσταση να αντιπροσωπεύει μία συναισθηματική συνιστώσα της τελικής συναισθηματικής κατάστασης. Κάθε συναίσθημα, αναπαρίσταται ως ένα σημείο στο χώρο αυτό. Οι περισσότερο χρησιμοποιούμενες διαστάσεις είναι της ευχαρίστησης και της διέγερσης αν και σε έρευνες χρησιμοποιούνται και περισσότερες διαστάσεις, όπως αυτή της επικράτησης [8, 11].

Λόγω της συνεχούς απεικόνισης των συναισθημάτων ως σημεία σε ένα χώρο, παρέχεται η δυνατότητα για δημιουργία περιοχών που αντιστοιχούν σε λεκτικές περιγραφές συναισθημάτων από τα διακριτά μοντέλα. Μία τέτοια αντιστοίχιση φαίνεται και στο Σχήμα 1.4



Σχήμα 1.4 Το επίπεδο AV με ομαδοποίηση περιοχών και αντιστοίχιση διακριτών συναισθημάτων

## 2 Η προτεινόμενη πλατφόρμα

Η πλατφόρμα που παρουσιάζεται αναπτύχθηκε με την χρήση των τεχνολογιών - γλωσσών προγραμματισμού Java Script, PHP, MySQL και HTML. Αναγνωρίζει τις ενέργειες που έχει εκτελέσει ο χρήστης (client) μέσω της χρήσης cookies, σε συνδυασμό με την γλώσσα προγραμματισμού Java Script, και την χρήση του μηχανισμού των sessions. Επίσης, και λόγω της εκτεταμένης χρήσης της γλώσσας Java Script, έχει προβλεφθεί η περίπτωση να μην έχει ενεργοποιηθεί η χρήση αυτής της γλώσσας στον browser του χρήστη, όπου και τυπώνεται σχετικό μήνυμα. Έτσι, ο χρήστης δεν μπορεί να εκτελέσει “ανανέωση” (refresh) και “επιστροφή” (back) κατά το δοκούν, διασφαλίζοντας την ελεγχόμενη, και μη επιβλεπόμενη από άνθρωπο, ροή του πειράματος.

## 2.1 Αρχιτεκτονική

Η αρχιτεκτονική της πλατφόρμας, σε επίπεδο προγραμματισμού και οργάνωσης της λειτουργικότητας, ακολουθεί τη λογική των συσχετίσεων που έχουν αποτυπωθεί κατά τον σχεδιασμό στην υποκείμενη σχεσιακή βάση δεδομένων. Παρά το γεγονός ότι στη διεπαφή χρήστη (UI) η πλατφόρμα εμφανίζει ένα γραμμικό μοντέλο ακολουθίας ενεργειών, η αρχιτεκτονική δομή είναι αμιγώς προσαρμοσμένη στον πρωταρχικό στόχο της δημιουργίας της που είναι η διεξαγωγή του πειράματος κατά τον βέλτιστο δυνατό τρόπο σύμφωνα με τις πειραματικές προδιαγραφές. Ο στόχος αυτός εξυπηρετείται από τον σχεδιασμό της βάσης δεδομένων και των συσχετίσεων ανάμεσα στις επιμέρους οντότητες.

Στο επίκεντρο του σχεδιασμού βρίσκεται ο συμμετέχοντας που εκτελεί το πείραμα το οποίο στην ουσία αποτελείται από μια συλλογή ήχων η οποία παρουσιάζεται γραμμικά, στο UI, ως μια βηματική ακολουθία σε κάθε στάδιο της οποίας ο “χρήστης” καλείται να ακούσει έναν ήχο από τη συλλογή και να επιλέξει τα εικονίδια που εκφράζουν το συναίσθημα που βίωσε. Η σύνδεση του πρώτου με τη συλλογή επιτυγχάνεται με τη συσχέτιση του μοναδικού προσδιοριστικού κάθε συλλογής με τον αντίστοιχο συμμετέχοντα μέσω ένα ξένου κλειδιού στον πίνακα των συμμετεχόντων. Η γραμμική ακολουθία που αντιστοιχεί στο UI εμφανίζεται στον σχεδιασμό της βάσης δεδομένων, επιτρέποντας κατά αυτόν τον τρόπο την ολοκλήρωση του συγκεκριμένου πειράματος ακόμα και αν για οποιονδήποτε λόγο διακοπεί, ακριβώς από το σημείο της διακοπής. Το πεδίο αυτό αποτελεί το σημείο τομής ανάμεσα στη γραμμική απεικόνιση του UI και στη σχεσιακή οργάνωση της αρχιτεκτονικής δομής της πλατφόρμας. Με την ολοκλήρωση του πειράματος τα αποτελέσματα αναλύονται σε ξεχωριστές γραμμές που περιλαμβάνουν το μοναδικό προσδιοριστικό του ήχου και τις επιλογές που έκανε ο συμμετέχοντας από κάθε σύνολο εικονιδίων για τον συγκεκριμένο ήχο και αποθηκεύονται υπό μορφή νέας εγγραφής στον πίνακα. Η τελευταία αυτή μετεγγραφή και επαναποθήκευση των αποτελεσμάτων επιτρέπει την πολλαπλή στατιστική επεξεργασία των μετρήσεων, ως προς τους συμμετέχοντες, τις συλλογές των ήχων και, κυρίως, τον κάθε ήχο ξεχωριστά.

## 2.2 Διαδικασία αρχικοποίησης πειράματος

Η διαδικασία της αρχικοποίησης μπορεί να χωριστεί σε δύο υπο-διαδικασίες. Σε αυτή που σχετίζεται με την πλατφόρμα και σε αυτή που προετοιμάζει κάθε χρήστη για την επιτυχημένη συμμετοχή του στο πείραμα.

Αναφορικά με την πλατφόρμα, κατά την είσοδο έκαστου χρήστη ζητείται ένα email, τόσο για λόγους επικοινωνίας, όσο και για ταυτοποίηση. Εάν ο χρήστης εισέρχεται για πρώτη φορά στην πλατφόρμα, τότε δημιουργείται αυτόματα μία λίστα με τους ήχους που θα κληθεί να επισημάνει και γίνεται η μετάβαση στο στάδιο της προετοιμασίας του χρήστη. Στην περίπτωση που ο ίδιος χρήστης επανέρχεται στην πλατφόρμα, τότε είτε πραγματοποιείται η μετάβαση στο στάδιο όπου υπήρχε η προηγούμενη διακοπή ή, και στην περίπτωση που έχει ολοκληρωθεί το πείραμα σε παρελθόντα χρόνο, παρουσιάζονται συγκεντρωτικά οι επισημάνσεις που έχει πραγματοποιήσει.

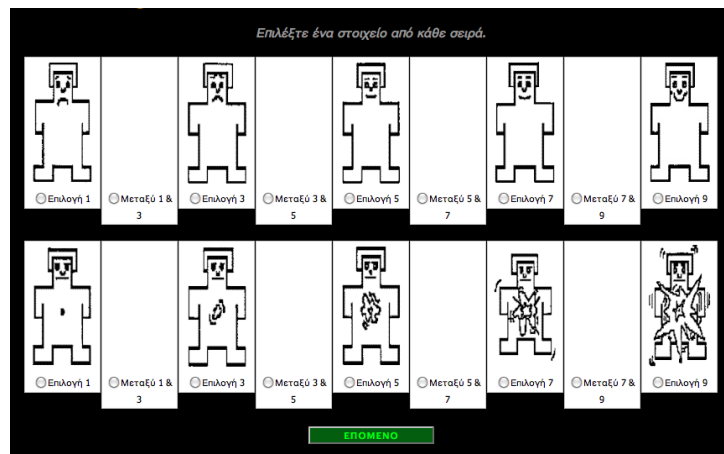
Η προετοιμασία του χρήστη περιλαμβάνει την αναπαραγωγή ενός video με την ενδεικνυόμενη χρήση της πλατφόρμας καθώς και την διαδικασία στάθμισης του επιπέδου αναπαραγωγής ήχου των ακουστικών. Η στάθμιση των ακουστικών γίνεται βάσει του μέγιστου ανεκτού, από τον χρήστη, επιπέδου αναπαραγωγής ήχου. για τον σκοπό αυτό γίνεται η αναπαραγωγή ενός τόνου συχνότητας 1kHz και ζητείται από τον χρήστη η ρύθμιση της στάθμης αναπαραγωγής στο μέγιστο ανεκτό επίπεδο.

### 2.3 Διαδικασία ακρόασης & επισήμανσης

Η διαδικασία ακρόασης βασίζεται στην τεχνολογία των sessions, διασφαλίζοντας ότι κάθε αναπαραγωγή ηχητικού αρχείου θα πραγματοποιηθεί μόνο μία φορά ανά χρήστη. Χρησιμοποιείται ο JW Player, έκδοση 5.10 [12], ως ένας media player ευρείας αποδοχής από όλους τους browsers. Ως διεπαφή στην χρήση παρουσιάζεται μόνο η επιλογή για την έναρξη της αναπαραγωγής.

Κατά την έναρξη και ολοκλήρωση της αναπαραγωγής, αποθηκεύεται η ενέργεια του χρήστη σε cookie, με αποτέλεσμα να είναι γνωστό, κάθε στιγμή, το στάδιο του πειράματος που έχει τελεστεί. Η επεξεργασία των cookies υλοποιείται με Java Script στον Η/Υ του χρήστη (client-side), ενώ η ανάγνωση αυτών στον εξυπηρετητή (server-side). Αναλόγως τα στοιχεία που υπάρχουν στο cookie, τυπώνονται τα αντίστοιχα μηνύματα στον browser του χρήστη, καθοδηγώντας τον για την συνέχεια του πειράματος.

Την ακρόαση του ηχητικού αρχείου ακολουθεί η διαδικασία επισήμανσης με την χρήση του SAM. Οι διαθέσιμες επιλογές φαίνονται στο Σχήμα 2.1.



Σχήμα 2.1 Οι διαθέσιμες επιλογές για την επισήμανση του ηχητικού αρχείου

Για κάθε ηχητικό αρχείο μπορεί να επιλεγεί μόνο μία επισήμανση για κάθε διαθέσιμη συναισθηματική κατάσταση - συνιστώσα. Οι επιλογές κάθε χρήστη αποθηκεύονται στην βάση δεδομένων της πλατφόρμας για την εξαγωγή τους κατά το πέρας ολόκληρου του πειράματος. Σε κάθε επιλογή του χρήστη, αποστέλλεται στον server η επιλογή, ή οι επιλογές, όπου και αποθηκεύονται στην βάση

δεδομένων.

### 3 Συμπεράσματα

Στην παρούσα εργασία παρουσιάστηκε μία διαδικτυακή πλατφόρμα για την υλοποίηση πειραμάτων συναισθηματικής επισήμανσης ηχητικών δεδομένων. Η πλατφόρμα αυτή προσφέρει την δυνατότητα ελέγχου της ροής του πειράματος, την επισήμανση των ηχητικών δεδομένων με την χρήση του SAM καθώς και την εξαγωγή των αποτελεσμάτων σε εύκολα επεξεργάσιμη μορφή. Βάσει του σχεδιασμού αλλά και της εγγενούς δυσκολίας, από την σκοπιά της οργάνωσης και του χρόνου που απαιτείται για την διεξαγωγή πειραμάτων συναισθηματικής επισήμανσης, η πλατφόρμα αναμένεται να προσφέρει την δυνατότητα για την αυτοματοποιημένη πραγματοποίηση τέτοιων πειραμάτων.

### 4 Αναφορές

- [1] Z. Shiliang, et al., “Affective MTV analysis based on arousal and valence features” IEEE International Conference on Multimedia and Expo, pp. 1369-1372 (2008)
- [2] D. Casacuberta, “DJ el Niño: expressing synthetic emotions with music” AI Soc. **18**, pp. 257-263 (2004)
- [3] K. Drossos et al., “Emergency Voice/Stress-level Combined Recognition for Intelligent House Applications” Audio Engineering Society 132nd Convention (2012)
- [4] M. Murugappan, et al., “Time-Frequency Analysis of EEG Signals for Human Emotion Detection” BIOMED 2008 (2008)
- [5] Y. Cheng, et al., “Study on Music Emotion Cognition Model Based on Applying the Improved Gene Expression Programming” Digital Media and its Application in Museum & Heritages, Second Workshop, pp. 344-351 (2007)
- [6] P. N. Juslin and P. Laukka, “Communication of emotions in vocal expression and music performance: different channels, same code?” Psychological Bulletin **129**, pp. 770-814 (2003)
- [7] M. M. Bradley and P. J. Lang, “Measuring emotion: The self-assessment manikin and the semantic differential” Journal of Behavior Therapy and Experimental Psychiatry **25**, pp. 49-59 (1994)
- [8] M. M. Bradley and P. J. Lang, “Affective reactions to acoustic stimuli” Psychophysiology **37**, pp. 204-215 (2000)
- [9] R. R. Cornelius, “Theoretical approaches to emotion” International Speech Communication Association (ISCA) Tutorial and Research Workshop (I-TRW) on Speech and Emotion (2000)
- [10] Hevner, K., “Experimental studies of the elements of expression in music” American Journal of Psychology **48**, pp. 246-268 (1936)
- [11] C. A. Smith and P. C. Ellsworth, “Patterns of cognitive appraisal in emotion” Journal of Personality & Social Psychology **48**, pp. 813-838 (1985)
- [12] JW Player: Download | LongTail Video | Home of thw JW Player, available at <http://www.longtailvideo.com/players/jw-flv-player/>, accessed at 20/08/2012