

# ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ II

## PROJECT 2014 – 2015

Text-Independent Speaker Recognition using MFCC

**Ακαδημαϊκό Έτος:** 2014 - 2015

.....  
**Ονοματεπώνυμο:** Δημήτριος Ζαρκάδας || **Αριθμός Μητρώου:** 4982  
**Ονοματεπώνυμο:** Χαράλαμπος Κανελλόπουλος || **Αριθμός Μητρώου:** 4994  
.....

## 0. Εισαγωγή

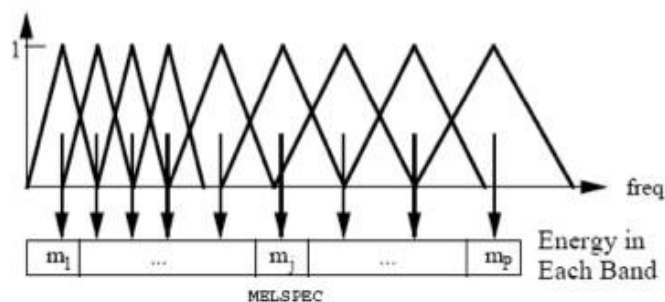
Σκοπός της παρούσας εργασίας είναι η ανάδειξη τεχνικών διαχείρισης δεδομένων χρονοσειρών και η πραγματοποίηση αναζητήσεων βάση του περιεχομένου αυτών ή κάποιου πεδίου των δεδομένων. Ειδικότερα, επιλέχθηκε οι χρονοσειρές αυτές να αποτελούν φωνητικές ηχογραφήσεις. Αυτές αναλύονται, εξάγονται χαρακτηριστικά και δημιουργούνται προφίλ ομιλητών. Με αυτό τον τρόπο μπορούμε είτε να μάθουμε συγκεκριμένα χαρακτηριστικά για ομάδες των ομιλητών που εμείς καθορίζουμε κατά την αναζήτηση μας είτε να δώσουμε ως είσοδο μια χρονοσειρά σαν ερώτημα, και να μας επιστραφεί σε ποιόν χρήστη ανήκει η ηχογράφηση ή σε ποιους μοιάζει περισσότερο. Τέλος, υπάρχει και η δυνατότητα να διαπιστώσουμε το φύλο του ομιλητή που εισάγαμε ως είσοδο. Για την υλοποίηση της χρησιμοποιήθηκε το περιβάλλον την Matlab και το Σύστημα Διαχείρισης Βάσεων Δεδομένων της PostgreSQL. Επιπλέον για τη διαχείριση των χαρακτηριστικών και τις συγκρίσεις των διανυσμάτων έγινε χρήση της PostGIS. Ακολουθεί λεπτομερής ανάλυση των εξής διαδικασιών :

- ✓ Εξαγωγή χαρακτηριστικών.
- ✓ Δεικτοδότηση και Αποθήκευση σε Βάση δεδομένων.
- ✓ Διεξαγωγή ερωτημάτων.
- ✓ Ανάλυση και Αξιολόγηση Υλοποίησης.

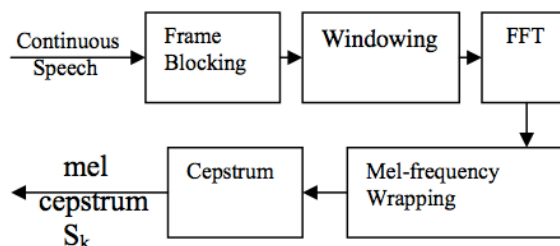
## 1. Εξαγωγή Χαρακτηριστικών

### I. Θεωρητική προσέγγιση :

Για την εξαγωγή χαρακτηριστικών χρησιμοποιήθηκε η διαδικασία επεξεργασίας ήχου MFCC (Mel Frequency Cepstral Coefficients). Πρόκειται για μια αναπαράσταση του σήματος μέσα από ένα διάνυσμα συντελεστών, οι οποίοι υπολογίζονται με βάση τον λογαριθμικό φάσματος ισχύος που προκύπτει από το μέτρο του μετασχηματισμού Fourier (FFT). Συγκεκριμένα, γίνεται δειγματοληψία της ενέργειας του σήματος σε διαφορετικές ζώνες συχνοτήτων και οι συντελεστές ενέργειας υπόκεινται σε μετασχηματισμό DCT (Discrete Cosine Transform) για την εξαγωγή των συντελεστών MFCC.



Διαγραμματικά η MFCC περιγράφεται από το παρακάτω σχήμα :



### Frame Blocking :

Το σήμα εισόδου διαχωρίζεται σε πλαίσια των 100ms με επικαλύψεις σταθερού μήκους των 30ms.

### Windowing :

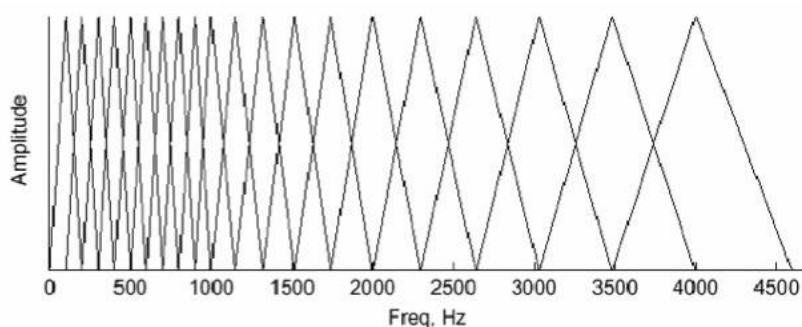
Δημιουργούνται “windows” για όλα τα frames, με στόχο την εξάλειψη ασυνεχειών στη διάρκεια τους. Έτσι λοιπόν, αν  $w(n)$  η συνάρτηση που το πραγματοποιεί για  $N$  μέρη σε κάθε frame, θα έχουμε στην έξοδο  $y(n)=w(n)*x(n)$ . Για το βήμα αυτό χρησιμοποιείται η αναπαράσταση Hamming.

### FFT :

Εκτελείται μετασχηματισμός Fourier για κάθε frame. Αυτό έχει ως αποτέλεσμα τη μεταφορά του προσδιορισμού του σήματος από το πεδίο του χρόνου (sec) στο πεδίο της συχνότητας (Hz).

### Mel Frequency Wrapping :

Το ανθρώπινο αυτί αντιλαμβάνεται τη συχνότητα μη γραμμικά. Έρευνες έχουν δείξει ότι αυτή κλιμακώνεται γραμμικά έως το 1kHz και λογαριθμικά πάνω από αυτό. Έτσι το Mel-Scale (Melody Scale) μιμείται το ανθρώπινο αυτί και με φίλτρα προσδιορίζει το σήμα γραμμικά και λογαριθμικά στις αντίστοιχες περιοχές συχνοτήτων. Για την μεταφορά των συχνοτήτων  $f$  στη κλίμακα Mel γίνεται χρήση του τύπου :  $mel(f) = 2595 * \log_{10}(1 + f/700)$



### Cepstrum :

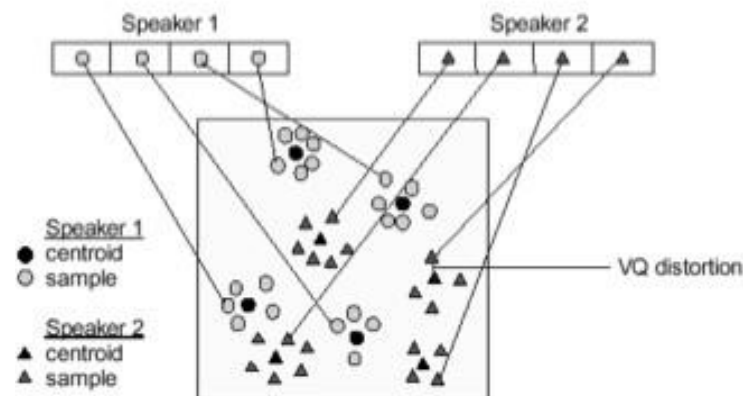
Σε αυτό το βήμα οι mel frequencies μεταφέρονται πίσω στην κλίμακα του χρόνου. Η αναπαράσταση που παράγεται βασίζεται σε ιδιότητες του φωνητικού φάσματος. Λόγω του ότι η mel κλίμακα παράγει πραγματικούς αριθμούς, η μεταφορά στην κλίμακα του χρόνου γίνεται με DCT μετασχηματισμό. Οι συντελεστές αυτού προκύπτουν από τον παρακάτω τύπο, όπου συνήθως  $k=20$  :

$$c_n = \sum_{k=1}^K (\log \tilde{S}_k) \left[ n \left( k - \frac{1}{2} \right) \frac{\pi}{K} \right],$$

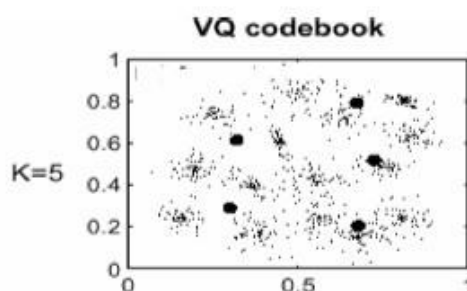
Αυτοί οι συντελεστές που προκύπτουν αποτελούν το ακουστικό διάνυσμα (acoustic vector) και είναι ικανοί να αντιπροσωπεύσουν την φωνή ενός ατόμου. Μετά το πέρας αυτών των διαδικασιών, θα έχουμε στην έξοδο για καθένα frame την αναπαράσταση του με τους  $k$  συντελεστές Cepstrum. Αν θεωρήσουμε ότι έχουμε ένα στιγμιότυπο των 10 δευτερολέπτων και κάθε frame ξεκινά ανά 30ms τότε θα έχουμε στην έξοδο έναν πίνακα (20 x 333). Έτσι σε επόμενο βήμα θα πρέπει να μειώσουμε τη διαστατικότητα του, κάτι που επιτυγχάνεται με χρήση του Vector Quantization (VQ) που αναλύεται παρακάτω.

### Vector Quantization (VQ) :

Στο σημείο αυτό περιορίζουμε την μεγάλη διάσταση του πίνακα μας σε  $n$ . Ουσιαστικά πρόκειται για ομαδοποίηση των χαρακτηριστικών εξαγωγής και την αντιπροσώπευση τους από  $n$  centroids. Για παράδειγμα αν έχουμε 2 ομιλητές και την αναπαράσταση των Cepstrum συντελεστών τους σε δύο διαστάσεις, ο VQ αλγόριθμος μας δίνει :



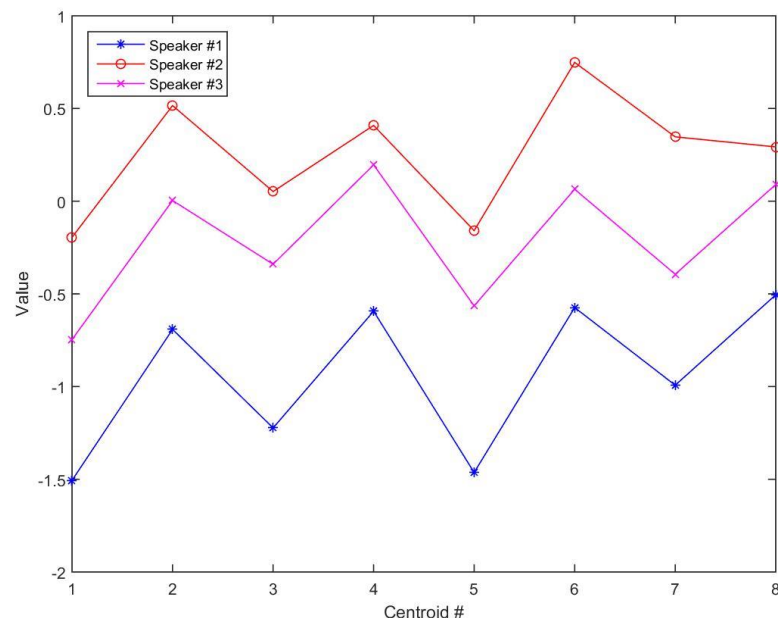
Συγκεκριμένα, θεωρούμε ότι έχουμε 330 frames και αναπαράσταση των cepstrum συντελεστών του στο διδιάστατο χώρο. Αν επιλέξουμε  $k=5$  centroids θα έχουμε τους αντιπροσώπους-frames, δηλαδή 5 ζευγάρια των 20 συντελεστών. Τελικώς στην έξοδο θα έχουμε έναν πίνακα (20 x 5) με εμφανώς μειωμένη διαστατικότητα.



## II. Υλοποίηση της MFCC σε κώδικα Matlab :

Η εκτέλεση της συνάρτησης `trainVQ.m` υλοποιεί την εξαγωγή χαρακτηριστικών χρησιμοποιώντας και τις συναρτήσεις `mfcc.m`, `vqldbg.m`, `melfb.m`, `disteu.m`. Η είσοδος της `trainVQ` δέχεται ένα πλήθος ( $n \geq 2$ ) αρχείων ήχου μορφής `.wav` ενός ομιλητή. Για κάθε αρχείο εκτελείται ξεχωριστά ανάλυση MFCC κι έπειτα Vector Quantization για  $k=8$  cendroids. Έτσι έχουμε στην έξοδο ένα πίνακα διαστάσεων  $(20 \times 8)$ . Για να επιτύχουμε όμως μειωμένη διαστατικότητα, παίρνουμε το μέσο όρο των 20 συντελεστών για καθένα εκ των 8 cendroids. Εν τέλει στη έξοδο έχουμε αναπαράσταση του αρχείου ήχου με ένα  $(1 \times 8)$  διάνυσμα. Η παραπάνω διαδικασία επαναλαμβάνεται για κάθε αρχείο που εισήχθη στην `trainVQ`. Στο τέλος παίρνοντας το μέσο όρο των  $n$  αρχείων για τις 8 τιμές, λαμβάνουμε το διάνυσμα αναπαράστασης για τον ομιλητή.

Εφαρμόζοντας λοιπόν την παραπάνω διαδικασία για 3 διαφορετικούς ομιλητές παίρνουμε το εξής γράφημα :



Παρατηρούμε ότι έχουμε μία πολύ καλή διαφοροποίηση των ατόμων και συνεπώς μπορούμε να εμπιστευτούμε τη διαδικασία της εξόρυξης.

Εδώ να σημειωθεί πως για την υλοποίηση της εξαγωγής των χαρακτηριστικών έγινε χρήση του paper "*Speaker Identification Using Mel Frequency Cepstral Coefficients* (Md. Rashidul Hasan, Mustafa Jamil, Md. Golam Rabbani Md. Saifur Rahman).

## 2. Δεικτοδότηση & Αποθήκευση Σε Βάση Δεδομένων

---

Για τη δημιουργία και διαχείριση της βάσης δεδομένων έγινε χρήση του open-source λογισμικού PostgreSQL DBMS (Database Management System). Το PostgreSQL χρησιμοποιείται για οντοκεντρικές βάσεις δεδομένων και επιτρέπει διαχείριση γεωμετρικών τύπων δεδομένων. Επιπλέον, έγινε χρήση του extension PostGIS, το οποίο υποστηρίζει τους τύπους δεδομένων geometry καθώς και συναρτήσεις που αφορούν αυτούς, όπως η κλάση ST (spatial type). Όπως θα δούμε παρακάτω κάνουμε χρήση της ST\_distance για τον προσδιορισμό των αποστάσεων μεταξύ των LineString.

Το PostGIS υποστηρίζει και διαχειρίζεται τα Geometry αντικείμενα σε ένα SAM (Spatial Access Method). Αυτό είναι το GiST ή Generalized Search Tree, το οποίο αποτελεί μια γενίκευση των B-tree και R-tree, αντικαθιστώντας μάλιστα το τελευταίο στην PostgreSQL από την 9η έκδοση και μετά. Επιπροσθέτως, διευκολύνει την εισαγωγή νέων δεδομένων και βελτιστοποιεί τις αναζητήσεις αναλόγως με τον τύπο αυτών. Τέλος, στις δύο διαστάσεις έχει την συμπεριφορά ενός R-tree.

Στη συγκεκριμένη περίπτωση, διαθέτουμε για κάθε ομιλητή ένα LineString που αντιπροσωπεύει τις χρονοσειρές ομιλίας και είναι της μορφής :

```
LINESTRING (value1 value2 value3 value4 value5 value6 value7 value8)
```

Έχουμε δηλαδή αναπαράσταση των 8 χαρακτηριστικών κάθε ομιλητή μέσω του string. Αυτά χρησιμοποιούνται για τον προσδιορισμό των αποστάσεων και την διεξαγωγή των ερωτημάτων.

Για την αποθήκευση των χαρακτηριστικών του χρησιμοποιείται ένας πίνακας (voicemap) που περιέχει τρεις στήλες γνωρισμάτων :

- **Name** : το όνομα του κάθε ομιλητή.
- **Genre** : το φύλο του κάθε ομιλητή .
- **Geom** : προσδιορισμός του Geometry Line που προκύπτει από την εξόρυξη χαρακτηριστικών.

## 3. Διεξαγωγή Ερωτημάτων

---

Για τη διεξαγωγή ερωτήματος με βάση κάποιο πεδίο των δεδομένων της βάσης, δίνεται η δυνατότητα στο χρήστη να επιλέξει ποια δεδομένα θέλει να εμφανιστούν και ποια κριτηρία αυτά να ικανοποιούν κι επιστρέφονται οι σχετικές χρονοσειρές και τα δεδομένα τους που τα ικανοποιούν. Το ερώτημα προς τη βάση είναι της μορφής :

```
SELECT data (name,genre,geom) FROM audio_recogn.voicemap WHERE criteria
```

Ακόμη, όπως προαναφέρθηκε, για την διεξαγωγή των ερωτημάτων με βάση το περιεχόμενο της χρονοσειράς χρησιμοποιούμε την μέθοδο ST\_distance του PostGIS. Αυτή υπολογίζει αποστάσεις μεταξύ geometry αντικειμένων ίδιου τύπου. Ειδικότερα η μορφή του ερωτήματος είναι :

```
SELECT name,genre,geom <-> ST_GeomFromText(''LINESTRING(x)'') AS distance  
FROM audio_recogn.voicemap  
ORDER BY geom <-> ST_GeomFromText(''LINESTRING(x)'')  
LIMIT y OFFSET 0;
```

Η `GeomFromText` μετατρέπει το string `LINESTRING (X)` σε μορφή `Geometry` αναγνωρίσιμη από το PostgreSQL. Έτσι, εκτελούμε ερωτήματα δίνοντας δύο ή περισσότερα δείγματα ήχου και το σύστημα μας απαντά, συγκρίνοντας το `LineString` που προκύπτει με αυτά του πίνακα `voicemap` της βάσης δεδομένων. Τα αποτελέσματα επιστρέφονται σε φθίνουσα σειρά, βάση της απόστασης, ενώ το πλήθος τους καθορίζεται από το χρήστη κι έχουν μορφή :

Speaker	Genre	Distance
<i>name</i>	<i>male/female</i>	<i>value</i>
...	...	...

## 4. Ανάλυση & Αξιολόγηση Υλοποίησης

---

### I. Σύστημα & Κώδικες Υλοποίησης :

Η εργασία υλοποιήθηκε στο παρακάτω περιβάλλον :

- Windows 7 Ultimate, Service Pack 1 (x64)
- Matlab R2015a (8.5.0.197613) | 64-bit | Feb. 12, 2015
- PostgreSQL 9.4.4
- PostGIS 2.1.8

Επίσης χρησιμοποιήθηκαν freeware Speech Corpus από την ιστοσελίδα <http://www.voxforge.org> που περιέχουν Wave (.wav) αρχεία των 16bit\_16KHz, τα οποία είναι απαραίτητα για την υλοποίηση μας.

Οι απαραίτητες συναρτήσεις που δημιουργήθηκαν σε κώδικα Matlab παρατίθενται παρακάτω ανά λειτουργία της συνολικής διαδικασίας που υλοποιούν :

Για την εξαγωγή χαρακτηριστικών :

- `trainVQ.m`
- `mfcc.m`
- `melfb.m`
- `vqldb.m`
- `disteu.m`
- `mfcc_training.m`

Για τη δημιουργία νέου profile νέου ομιλητή :

- `new_profile.m`

Για τη διεξαγωγή των ερωτημάτων προς τη βάση :

- `query0.m`
- `query.m`

Για τον έλεγχο ύπαρξης κάποιου χρήστη στη βάση δεδομένων :

- `loginq.m`

Για την επικοινωνία του Matlab με την βάση δεδομένων μας στο PostgreSQL :

- `postgre_configuration.m`

Για τη δημιουργία Linestring για κάθε χρονοσειρά που εισάγεται :

- `make_linestring.m`

Για το γραφικό περιβάλλον που παρέχεται στο χρήστη :

- `first.m / .fig` (Αρχική οθόνη εκκίνησης της εφαρμογής)
- `login.m / .fig` (Οθόνη εισαγωγής στοιχείων για τη σύνδεση στη βάση δεδομένων)
- `main_menu.m / .fig` (Βασικό μενού επιλογής λειτουργίας προς εκτέλεση)
- `profile.m / .fig` (Περιβάλλον καταχώρησης νέου profile ομιλητή)
- `query_menu0.m / .fig` (Περιβάλλον επιλογής ερωτήματος προς τη βάση #1)
- `query_menu.m / .fig` (Περιβάλλον επιλογής ερωτήματος προς τη βάση #2)

## II. Οδηγίες Εκτέλεσης :

Αφού γίνει επιτυχώς η εγκατάσταση του PostgreSQL για τη διαχείριση της βάσης μας και του PostGIS για τη γεωμετρική αναπαράσταση των χρονοσειρών, δημιουργούμε μια νέα βάση δεδομένων (έστω DB2Project) και εντός αυτής πρέπει να προστεθεί ένα schema (έστω `audio_recogn`) και 2 πίνακες για την αποθήκευση των επιτρεπόμενων χρηστών της βάσης (table : `users` | fields : `name, pass`) και των δεδομένων κάθε εισαγόμενης χρονοσειράς (table : `voicemap` | fields: `name, genre, geom`). Έτσι προκύπτουν τα παρακάτω τμήματα κώδικα MySQL :

```
CREATE TABLE audio_recogn.users
(
    name character varying,
    pass character varying
)
WITH (
    OIDS=FALSE
);
ALTER TABLE audio_recogn.users
    OWNER TO postgres;
CREATE TABLE audio_recogn.voicemap
(
    name character varying,
    genre character varying,
    geom geometry
)
WITH (
    OIDS=FALSE
);
ALTER TABLE audio_recogn.voicemap
    OWNER TO postgres;
```

Για διευκόλυνση κατά την δοκιμαστική εκτέλεση της υλοποίησης, παρέχονται έτοιμα προφίλ ομιλητών και λίστα χρηστών στο φάκελο Tables σε μορφή .csv αρχείων τα οποία μπορούν να εισαχθούν απευθείας ως δεδομένα στους αντίστοιχους πίνακες της βάσης.

Κατόπιν, πρέπει να γίνει εγκατάσταση του κατάλληλου JDBC Driver για να είναι εφικτή η επικοινωνία της Matlab με το DBMS, ενώ η παραμετροποίηση της για τη σύνδεση στη βάση γίνεται μέσω της συνάρτησης `postgre_configuration.m`. Αναλόγως με το όνομα της βάσης και τα στοιχεία σύνδεσης στο PostgreSQL που θέτουμε κατά τη δημιουργία της, τροποποιούμε και τα ορίσματα που δίνονται κατά την κλήση της συνάρτησης (αρχείο `login.m / line 144`).



Πλέον, όλα είναι έτοιμα κι εκτελούμε βηματικά :

1. Εκτελούμε το αρχείο `first.fig` και πατάμε στο κουμπί “Click to login” για να μεταφερθούμε στο επόμενο παράθυρο.
2. Στο `login.fig` που ανοίγει, βάζουμε τα στοιχεία σύνδεσης στη βάση δεδομένων (default : root, root) και πατάμε “Connect”.
3. Αφού έχουμε μεταφερθεί στο κεντρικό μενού επιλογών (`main_menu.fig`) επιλέγουμε την επιθυμητή λειτουργία από τις :
  - a) Δημιουργία νέου προφίλ ομιλητή (πατάμε “Create new voice profile”).
  - b) Πραγματοποίηση ερωτήματος προς τη βάση σχετικά με κάποιο πεδίο των δεδομένων. (Πατάμε “Query #1”).
  - c) Πραγματοποίηση ερωτήματος με βάση το περιεχόμενο κάποιας χρονοσειράς (Πατάμε “Query #2”).

Αν επιλέξουμε a) Δημιουργία νέου προφίλ ομιλητή :

4. Εισάγουμε το όνομα και το φύλο του νέου ομιλητή στα αντίστοιχα πεδία του `profile.fig` και πατάμε “Next”.
5. Στο αναδυόμενο παράθυρο περιηγούμαστε και επιλέγουμε τα αρχεία ήχου τύπου .wav (τουλάχιστον 2) που αντιστοιχούν στον νέο ομιλητή και πατάμε “Open”.
6. Το νέο προφίλ ομιλητή έχει δημιουργηθεί. Πατάμε “Back” και επιστρέφουμε στο κύριο μενού για να επιλέξουμε την επόμενη προς εκτέλεση λειτουργία ή επαναλαμβάνουμε την προηγούμενη διαδικασία για να προσθέσουμε κι άλλα προφίλ ομιλητών.

Αν επιλέξουμε b) Πραγματοποίηση ερωτήματος προς τη βάση σχετικά με κάποιο πεδίο των δεδομένων :

4. Επιλέγουμε (στο `query_menu0.fig`) ποια από τα μετα-δεδομένα των ομιλητών θέλουμε να μας επιστραφούν (πρέπει αρχικά να επιλεγούν όλα και στη συνέχεια να απο-επιλεγούν όσα δεν θέλουμε) και εισάγουμε τα κριτήρια που επιθυμούμε για την αναζήτηση των σχετικών εγγραφών της βάσης. Πατάμε “Execute”.
5. Το σύστημα μας έχει επιστρέψει τα σχετικά αποτελέσματα που ζητήσαμε. Μπορούμε να πραγματοποιήσουμε εκ νέου μια άλλη αναζήτηση ή να πατήσουμε “Back” για επιστροφή στο κύριο μενού κι επιλογή άλλης λειτουργίας.

Αν επιλέξουμε c) Πραγματοποίηση ερωτήματος με βάση το περιεχόμενο κάποιας χρονοσειράς :

4. Επιλέγουμε (στο `query_menu.fig`) αν θέλουμε να ελέγξουμε ομοιότητα του ομιλητή σε σχέση με το φύλο ή με τους ομιλητές της βάσης και εισάγουμε τον πλήθος των πιο όμοιων ομιλητών που θέλουμε να μας επιστραφεί. Πατάμε “Execute”.
5. Στο αναδυόμενο παράθυρο περιηγούμαστε και επιλέγουμε τα αρχεία ήχου τύπου .wav (τουλάχιστον 2) που αντιστοιχούν στον ομιλητή που θέλουμε να συγκρίνουμε με τους υπόλοιπους και πατάμε “Open”.
6. Το σύστημα μας έχει επιστρέψει τα σχετικά αποτελέσματα που ζητήσαμε. Μπορούμε να πραγματοποιήσουμε εκ νέου μια άλλη αναζήτηση ή να πατήσουμε “Back” για επιστροφή στο κύριο μενού κι επιλογή άλλης λειτουργίας.

### III. Πειραματικά Αποτελέσματα :

Εκτελέσαμε τη δημιουργία profile και δημιουργήσαμε 10 εγγραφές στον πίνακα `voicemap` (βλ. `voicemap.csv`). Τα αρχεία ήχου βρίσκονται στο φάκελο `WAV_Files` και κάθε υποφάκελος αναφέρεται σε έναν χρήστη και έχει Training Set και Test Set. Τα προφίλ δημιουργήθηκαν χρησιμοποιώντας τα training set, ενώ τα ερωτήματα που κάναμε στη συνέχεια, χρησιμοποιώντας τα Test Set. Για τους διάφορους τύπους ερωτημάτων έχουμε :

- ❖ **Πραγματοποίηση ερωτήματος προς τη βάση σχετικά με κάποιο πεδίο των δεδομένων :**

Ζητήσαμε να μας επιστραφούν τα ονόματα (name) όλων των καταχωρημένων ομιλητών που είναι άνδρες (genre = 'male' ) και λάβαμε τα εξής :

Name
Dimitris
Mitsos
Nikos
Pantelis
Pap
Petros
Zaq

Παρατηρούμε πως πράγματι, τα ονόματα είναι σωστά καθώς 7/10 εγγραφές που έχουμε κάνει στη βάση είναι άνδρες.

- ❖ **Πραγματοποίηση ερωτήματος με βάση το περιεχόμενο κάποιας χρονοσειράς :**

Τα αποτελέσματα που λάβαμε για ομοιότητα ως προς τους ομιλητές της βάσης και για  $k = 3$  όμοιους ομιλητές είναι (όνομα και απόσταση από τον προς εξέταση ομιλητή) :

Training Name	Result 1	Result 2	Result 3
Dimitris	Dimitris 0.037	Pantelis 0.28	Kiki 0.211
Flower	Flower 0.016	Roula 0.212	Mitsos 0.278
Kiki	Kiki 0.015	Pap 0.060	Pantelis 0.075
Mitsos	Mitsos 0.028	Pap 0.046	Kiki 0.012
Nikos	Petros 0.032	Nikos 0.036	Zaq 0.117
Pantelis	Kiki 0.020	Pantelis 0.039	Pap 0.096
Pap	Kiki 0.037	Pap 0.038	Mitsos 0.056
Petros	Petros 0.029	Nikos 0.041	Zaq 0.121
Roula	Roula 0.049	Flower 0.28	Mitsos 0.54
Zaq	Zaq 0.0152	Nikos 0.095	Petros 0.162

Ομοίως, για  $k = 1$  και αναζήτηση ομοιότητας ως προς φύλο ομιλητή με μη καταχωρημένους ομιλητές :

Training Name	Result
Elvis	Male 0.079
Mariana	Female 0.087

Παρατηρούμε πως 9/12 από τα αποτελέσματα μας είναι σωστά ενώ στα 3/12 λανθασμένα, και υπάρχουν πολύ μικρές αποστάσεις του False\_Positive σε σχέση με το False\_Negative που ήταν πάντα δεύτερο.