

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Τεχνολογία Πολυμέσων

Άνοιξη 2018

Εργασία 7η

Η εργασία έχει στόχο την τμηματοποίηση σημάτων ήχου και την ταξινόμηση στη συνέχεια των τμημάτων που προκύπτουν. Για την επίτευξη αυτών των στόχων θα χρησιμοποιηθεί μόνο το πλάτος του σήματος. Συγκεκριμένα θα υπολογισθεί το μέσο πλάτος σε μη επικαλυπτόμενα πλαίσια διάρκειας 20 ms.

Αρχικά οι τιμές της ισχύος του σήματος x θα εξομαλυνθούν με χρήση ενός γραμμικού συστήματος υλοποιημένου με τη συνάρτηση Matlab

```
filtfilt(b, [1 -b], x.^2)
```

και τιμή της παραμέτρου $b = 0.05$. Εκτιμάται στη συνέχεια το στιγμιαίο πλάτος λαμβάνοντας την τετραγωνική ρίζα της εξομαλυμένης ισχύος και υπολογίζεται το μέσο πλάτος σε κάθε πλαίσιο. Ζητείται να παρασταθεί γραφικά το μέσο πλάτος με οριζόντιο άξονα το χρόνο. Το μέσο πλάτος εξομαλύνεται επίσης με το ανωτέρω φίλτρο, και λαμβάνονται οι τιμές $\{A(n), n = 1, \dots, N_B\}$, όπου N_B είναι το πλήθος των πλαισίων. Υπολογίζονται ακολούθως οι μη προσημασμένες διαφορές $D(n) = |A(n+1) - A(n-1)|$. Ζητείται να παρασταθούν γραφικά με οριζόντιο άξονα το χρόνο. Να εντοπισθούν οι σημαντικές διαφορές στο μέσο πλάτος ως τα τοπικά μέγιστα της διαφοράς $D(n)$ που είναι πάνω από το διπλάσιο της μέσης τιμής της διαφοράς και απέχουν τουλάχιστον 0.5 sec. Με τον τρόπο αυτό κατακερματίζεται το σήμα του ήχου. Οι στιγμές των μεγίστων, όπου η διαφορά είναι σημαντική, είναι πιθανές στιγμές αλλαγής από φωνή σε μουσική ή αντίστροφα.

Για να ληφθεί απόφαση ταξινόμησης για τα τμήματα που προέκυψαν από την ανίχνευση των τοπικών μεγίστων, θα μετρηθούν για κάθε τμήμα η μέση τιμή και η τυπική απόκλιση για το μέσο πλάτος. Η απόφαση λαμβάνεται με βάση το λόγο της μέσης τιμής προς την τυπική απόκλιση. Αν ο λόγος είναι μικρότερος από 2.25, λαμβάνεται η απόφαση ότι πρόκειται για σήμα φωνής, διαφορετικά μουσικής. Να διαχωρισθούν τα τμήματα μουσικής και φωνής και να περιγραφούν με το λόγο της μέσης τιμής προς την τυπική απόκλιση του μέσου πλάτους και με τον ίδιο λόγο για το μέσο ρυθμό εναλλαγής προσήμου υπολογισμένο στα ίδια πλαίσια. Ο μέσος ρυθμός εναλλαγής προσήμου ορίζεται ως

$$z = \frac{1}{2(N-1)} \sum_{n=1}^{N-1} |\text{sign}(x(n+1)) - \text{sign}(x(n))|.$$

Τα αποτελέσματα θα δοθούν για το σήμα *TestSegment.wav*. Να εξαχθούν συμπεράσματα και να δοθεί η αναφορά της εργασίας σε HTML.

Χρήσιμες συναρτήσεις : `audioread`, `filtfilt`, `findpeaks`, `mean`, `std`.