



## ΟΜΑΔΑ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ

Υπεύθυνος: Καθηγητής Νίκος Φακωτάκης  
Επικ. Καθηγητής Κυριάκος Σγάρμπας

### ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ 2009-2010

#### 1. Εντοπισμός Ομιλίας (Speech Activity Detection)

**Περιγραφή:** Σε ένα σύστημα αναγνώρισης ομιλίας ο δέκτης (μικρόφωνο) μπορεί να καταγράψει και άλλους ήχους εκτός από ομιλία, όπως σιγή (silence), θορύβους (noises), μουσική (music), κλπ. Στόχος της διπλωματικής εργασίας είναι η υλοποίηση ενός αλγορίθμου εντοπισμού της ομιλίας, για περαιτέρω επεξεργασία της από σύστημα αναγνώρισης γλώσσας.

**Επιθυμητές Γνώσεις:** Ψηφιακή Επεξεργασία Σημάτων, Matlab ή C.

**Επιβλέπων Ερευνητής:** Δρ. Ιωσήφ Μπόρας

#### 2. Αναγνώριση Γλώσσας από κείμενο (Language Identification from Text)

**Περιγραφή:** Σε εφαρμογές αυτόματης μετάφρασης κειμένου γίνεται μετάφραση από μια γλώσσα σε άλλη, οικεία προς τον χρήστη. Αυτό προϋποθέτει την εκ των προτέρων γνώση της αρχικής γλώσσας. Εφόσον όμως η γλώσσα δεν μπορεί να αναγνωριστεί από τον χρήστη δεν μπορεί να προχωρήσει στην μετάφρασή της. Σκοπός της παρούσας διπλωματικής είναι η κατασκευή μοντέλων που θα αναγνωρίζουν την γλώσσα στην οποία ανήκει ένα κείμενο.

**Επιθυμητές Γνώσεις:** Πιθανότητες - Στατιστική, Matlab ή C.

**Επιβλέποντες Ερευνητές:** Δρ. Ιωσήφ Μπόρας

#### 3. Πολυγλωσσική Αναγνώριση Ομιλίας (Multilingual Speech Recognition)

**Περιγραφή:** Κάθε γλώσσα αποτελείται από ένα ξεχωριστό σύνολο φωνημάτων που την αποτελούν. Υπάρχουν γλώσσες που μοιράζονται κοινά φωνήματα σε μεγαλύτερο ή μικρότερο ποσοστό. Η δημιουργία επομένως μιας βάσης μοντέλων φωνημάτων ανεξαρτήτου γλώσσας μπορεί να χρησιμοποιηθεί για την δημιουργία συστήματος αναγνώρισης ομιλίας, ακόμη και σε γλώσσες όπου δεν υπάρχουν αρκετά δεδομένα.

**Επιθυμητές Γνώσεις:** Ψηφιακή Επεξεργασία Σημάτων, Matlab ή C.

**Επιβλέποντες Ερευνητές:** Δρ. Ιωσήφ Μπόρας

#### 4. Αυτόματη Αναγνώριση Ομιλίας (Automatic Speech

## Recognition)

**Περιγραφή:** Σε ένα σύστημα αναγνώρισης ομιλίας η απόδοση εξαρτάται από πολλούς παράγοντες, όπως ο όγκος των διαθέσιμων δεδομένων εκπαίδευσης, η τεχνική παραμετροποίησης που χρησιμοποιείται κλπ. Είναι πολύ κρίσιμο για την απόδοση του συστήματος να γίνει βελτιστοποίηση των παραμέτρων του. Η παρούσα διπλωματική στοχεύει στην μέτρηση της επίδρασης των παραμέτρων του συστήματος σε διαφορετικές συνθήκες.

**Επιθυμητές Γνώσεις:** Ψηφιακή Επεξεργασία Σημάτων, Matlab ή C

**Επιβλέπων Ερευνητής:** Δρ. Ιωσήφ Μπόρας

## 5. Αυτόματος Υπολογισμός Μήκους φωνητικού καναλιού από ομιλία

**Περιγραφή:** Η διπλωματική εργασία στοχεύει στην μελέτη της καταλληλότητας φασματικών παραμέτρων ομιλίας για τον υπολογισμό του μήκους της φωνητικής οδού ενός ομιλητή. Θα μελετηθούν διάφορες υπάρχουσες μέθοδοι και θα αξιολογηθεί η χρησιμότητα τους στην αυτόματη αναγνώριση ομιλίας με Vocal Tract Length Normalization.

**Επιθυμητές Γνώσεις:** Ψηφιακή Επεξεργασία Σημάτων, Matlab ή C.

**Επιβλέπων Ερευνητής:** Δρ. Ιωσήφ Μπόρας

## 6. Υπολογισμός φυσικών χαρακτηριστικών ομιλητή από την ομιλία του

**Περιγραφή:** Το σήμα ομιλίας εκτός από το γλωσσολογικό περιεχόμενο περιέχει και πληροφορίες που αφορούν την ταυτότητα του ομιλητή και άλλα βιομετρικά χαρακτηριστικά του. Στόχος της εργασίας είναι η μελέτη της καταλληλότητας διαφόρων φασματικών παραμέτρων ομιλίας για τον υπολογισμό διαφόρων φυσικών χαρακτηριστικών του ομιλητή, όπως π.χ. το ύψος του.

**Επιθυμητές Γνώσεις:** Ψηφιακή Επεξεργασία Σημάτων, Matlab ή C

**Επιβλέπων Ερευνητής:** Δρ. Ιωσήφ Μπόρας

## 7. Σύστημα μετατροπής χαρακτηριστικών ομιλίας από έναν ομιλητή σε άλλον (voice conversion) και πειραματική εφαρμογή με κώδικα.

**Περιγραφή :** Voice Conversion είναι η διαδικασία μετατροπής χαρακτηριστικών ομιλίας ενός συγκεκριμένου ομιλητή, ώστε να ακούγεται σαν να έχει ειπωθεί από έναν άλλον συγκεκριμένο ομιλητή. Τα σύγχρονα συστήματα μετατροπής ομιλίας βασίζονται τόσο στην μετατροπή της περιβάλλουσας του σήματος (spectral envelope), όσο και στην μετατροπή των λεπτομερειών (spectral details). Η παρούσα διπλωματική εργασία έχει σκοπό την συλλογή πληροφορίας και περιγραφή του state-of-the-art των συστημάτων αυτών καθώς και την υλοποίηση.

**Επιθυμητές Γνώσεις:** Matlab ή C

**Επιβλέπων Ερευνητής:** Λαζαρίδης Αλέξανδρος

## 8. Δημιουργία βάσης ομιλίας και manual annotation

**Περιγραφή :** Η παρούσα διπλωματική εργασία έχει σκοπό την δημιουργία βάσης ομιλίας από διαφορετικούς ομιλητές. Το έργο θα αποτελείται από ηχογράφιση και manual annotation των διαφορετικών ομιλητών για την δημιουργία της βάσης.

**Επιβλέπων Ερευνητής:** Λαζαρίδης Αλέξανδρος

## 9. Αυτόματη μετατροπή κειμένου σε ομιλία (Text-to-Speech Synthesis) και manual annotation βάσης ομιλίας.

**Περιγραφή:** Σύνθεση ομιλίας από κείμενο είναι η διαδικασία μετατροπής κειμένου σε ομιλία από τον υπολογιστή. Η παρούσα διπλωματική εργασία έχει σκοπό την συλλογή πληροφορίας και περιγραφή του state-of-the-art των συστημάτων αυτών, το manual annotation βάσης ομιλίας ενός ομιλητή καθώς και την δημιουργία συστήματος μετατροπής κειμένου σε ομιλία βασισμένο στο framework του συστήματος FESTIVAL.

**Επιβλέπων Ερευνητής:** Λαζαρίδης Αλέξανδρος

## 10. Age recognition from speech

**Περιγραφή:** In the modern speech interaction systems, for achieving a better quality of service, it is desirable in unobtrusive manner to acquire some personal traits of the user from his/her voice. For instance, such traits are the user's gender, age, educational background, emotional state, etc.

The present diploma thesis will focus on age recognition from speech. The age-specific features of the human vocal tract will be exploited in a text-independent scenario. General age categories as minor, adult, elder person, or range-specific age categories such as: younger than twenty, 20-35, 36-50, 51-65, older than 66 are considered. The specifics of the vocal tract that are related to the aging process will be exploited to perform the classification. The practical usefulness of various acoustic features will be evaluated. A score level fusion of state-of-the-art classification approaches, such as Gaussian Mixture Models (GMM) and Support Vector Machines (SVM) will be performed. All experimentations will be based on the SpeechDat(II)-Greek corpora, which contains about 5000 speakers.

**Επιθυμητές γνώσεις:** MATLAB or C programming languages

**Επιβλέπωντας Ερευνητής:** Todor Ganchev

## 11. Gender recognition from speech

**Περιγραφή:** In the modern speech interaction systems it is often desirable to detect the gender of the users. This information can be used in various ways: appropriate adaptation of the dialogue; gender-specific acoustic models for speech, speaker, or language recognition; etc.

The present diploma thesis will focus on gender recognition from speech. The gender-related characteristics of the human vocal tract will be exploited in a text-independent scenario. Various acoustic and prosodic parameters, related to the specifics in the anatomy of the vocal tract and the speaking style of the genders will be used. A score-level fusion of state-of-the-art classification approaches, such as Gaussian Mixture Models (GMM) and Support Vector Machines (SVM) will be performed. The experimentations will be based on the TIMIT and SpeechDat(II)-Greek corpora.

**Επιθυμητές γνώσεις:** MATLAB or C programming languages.

**Επιβλέπωντας Ερευνητής:** Todor Ganchev

## 12. Transmission channel/handset recognition

**Περιγραφή:** Transmission channel/handset recognition is mandatory component in the modern speech interaction systems. The information about the transmission channel and handset type is usually utilized in a channel normalization procedure that precedes the classification stage of many telephone-driven speech applications: speech, speaker, language, or emotion recognition etc.

The present diploma thesis will focus on transmission channel/handset recognition. The transmission channel characteristics will be judged from the alteration they cause to the speech signal. General categories of interest will be: (1) land-based line/carbon-button microphone; (2) land-based line/electret microphone; (3) mobile phone GSM ; (4) Cellular phone, but also other potential mobile standards: TDMA, CDMA, etc.

A Gaussian Mixture Models (GMM)-based classifier will be used. The experimentations will be based on the Switchboard-2 and NIST'01÷04 speech corpora.

**Επιθυμητές γνώσεις:** MATLAB or C programming languages.

**Επιβλέπωντας Ερευνητής:** Todor Ganchev

## 13. Αναλυτής (parser) Νομικών Κειμένων της Ελληνική Γλώσσα.

**Περιγραφή:** Για την αναζήτηση, επεξεργασία και σωστή ταξινόμηση νομικών κειμένων, ένα καλό εργαλείο μπορεί να αναλάβει το ρόλο της προεπεξεργασίας και μορφοποιημένης αποθήκευσης. Συγκεκριμένα θα γίνει προσπάθεια δημιουργίας αυτόματου εργαλείου το οποίο θα έχει ως είσοδο νομοθετικές διατάξεις του κράτους, σε απλό αμορφοποίητο κείμενο "plain text", και σκοπός θα είναι ως έξοδος να βγαίνει το ίδιο κείμενο αλλά με σημειογραφία που θα δηλώνει σε ποια σημεία π.χ. ξεκινάει και τελειώνει ένας νόμος, ένα άρθρο, μία παράγραφος - όπως αυτά ορίζονται μέσα στις νομοθετικές διατάξεις με χρήσεις αριθμών και τυποποιημένης νομικής γλώσσας.

**Επιθυμητές Γνώσεις:** Γνώσεις Προγραμματισμού και Τεχνολογίας Φυσικής Γλώσσας.

**Επιβλέπωντας Ερευνητής:** Χαράλαμπος Τσιμπούρης

## 14. Δημιουργία διαδικτυακής πλατφόρμας εκπαιδευτικών παιχνιδιών.

**Περιγραφή:** Σκοπός της διπλωματικής εργασίας θα είναι η δημιουργία μίας διαδικτυακής πλατφόρμας η οποία θα έχει εκπαιδευτικό χαρακτήρα στον προγραμματισμό. Με την υλοποίηση της πλατφόρμας, ο εκπαιδευόμενος φοιτητής θα μπορεί να "ανεβάσει" την υλοποίηση του σε συγκεκριμένο παιχνίδι (για παράδειγμα την τρίλιζα ή το σκάκι) και θα παίζει ως αντίπαλος με υλοποίηση του ίδιου προβλήματος από άλλον φοιτητή. Η πλατφόρμα θα μπορεί να βγάλει στατιστικά σε ποιους φοιτητές η υλοποίηση κέρδισε τις περισσότερες φορές, αλλά και στατιστικά για καλύτερη και χειρότερη επεξεργαστική ισχύ που καταναλώθηκε.

Ακόμα, θα μπορεί να δώσει και αναπαράσταση του παιχνιδιού δημιουργώντας τις κατάλληλες εικόνες, ώστε ο χαμένος να μπορεί να δει τους λόγους για τους οποίους έχασε. Η πλατφόρμα θα κατασκευαστεί σε γλώσσα PHP και οι εσωτερικές πληροφορίες θα είναι αποθηκευμένες σε βάση της MySQL. Αν και η γνώση των γλωσσών είναι απαραίτητη για τη σωστή διεκπεραίωση, δεν είναι αναγκαία για την έναρξη εφόσον ο φοιτητής που θα δηλώσει την εργασία είναι διατεθειμένος και έχει όρεξη για το τελικό αποτέλεσμα.

**Επιθυμητές/Προαιρετικές Γνώσεις:** Γνώσεις Προγραμματισμού (PHP / C) και Βάσεων Δεδομένων (MySQL)

**Επιβλέπωντας Ερευνητής:** Χαράλαμπος Τσιμπούρης