



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
"ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΚΑΙ ΤΗ ΒΙΟΛΟΓΙΑ"**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Ανάλυση Πολυμεσικού Περιεχομένου για την
Εκτίμηση Επιπέδου Κατάθλιψης**

Χρήστος Σμαΐλης

**Επιβλέποντες: Περαντώνης Σταύρος, Διευθυντής Ερευνών ΕΚΕΦΕ-
Δημόκριτος**

ΑΘΗΝΑ

ΑΥΓΟΥΣΤΟΣ 2015

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ανάλυση Πολυμεσικού Περιεχομένου για την Εκτίμηση Επιπέδου Κατάθλιψης

Χρήστος Β. Σμαΐλης

A.M.: ΠΙΒ0107

ΕΠΙΒΛΕΠΟΝΤΕΣ: **Περαντώνης Σταύρος**, Διευθυντής Ερευνών ΕΚΕΦΕ-
Δημόκριτος

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ: **Περαντώνης Σταύρος**, Διευθυντής Ερευνών, ΕΚΕΦΕ-
Δημόκριτος

Γιαννακόπουλος Θεόδωρος, Επιστημονικός Συνεργάτης,
ΕΚΕΦΕ-Δημόκριτος

Σπύρου Γεώργιος, Ειδικός Λειτουργικός Επιστήμονας
(βαθμίδα Α'), Ίδρυμα Ιατροβιολογικών Ερευνών της
Ακαδημίας Αθηνών

Αύγουστος 2015

ΠΕΡΙΛΗΨΗ

Ένα σημαντικό ποσοστό του παγκόσμιου πληθυσμού, επηρεάζεται από τη κατάθλιψη. Σύμφωνα με στοιχεία του Παγκόσμιου Οργανισμού Υγείας ως το 2012 υπήρχαν τουλάχιστον 350 εκατομμύρια άνθρωποι που υπέφεραν από κάποια μορφή κατάθλιψης. Η κακή ψυχική υγεία συχνά επηρεάζει άτομα σε ηλικία εργασίας, γεγονός που έχει έντονες κοινωνικές επιπτώσεις. Για την αξιολόγηση του επιπέδου κατάθλιψης ενός ασθενούς οι ψυχολόγοι και οι ψυχίατροι χρησιμοποιούν ερωτηματολόγια (π.χ Beck Depression Index-II - BDI-II) καθώς και στοιχεία των εκφράσεων του προσώπου και της φωνής.

Στα πλαίσια αυτής της πτυχιακής η εκτίμηση του επιπέδου κατάθλιψης αντιμετωπίστηκε ως πρόβλημα πρόβλεψης του βαθμού κατάθλιψης ενός ατόμου στη κλίμακα BDI-II χρησιμοποιώντας τεχνικές μηχανικής μάθησης για την αναγνώριση χαρακτηριστικών του προσώπου και της φωνής.

Πιο συγκεκριμένα, τα video που χρησιμοποιήθηκαν ως δεδομένα εισόδου στο πλαίσιο αυτής της εργασίας προήλθαν από το Depression Recognition Sub-Challenge του 2014 Audio-Visual Emotion Challenge and Workshop (AVEC 2014). Η εκτίμηση του επιπέδου κατάθλιψης έγινε με χρήση μοντέλων παλινδρόμησης SVR (Support Vector Regressor) βασισμένων σε χαρακτηριστικά εικόνας που εξήχθησαν με χρήση του αλγορίθμου AAM(Active Appearance Model), καθώς και ποικιλίας χαρακτηριστικών ήχου.

Τα αποτελέσματα έδειξαν πως ο συνδυασμός των χαρακτηριστικών AAMs και ήχου, οδήγησε σε βελτίωση της ακρίβειας της διάγνωσης της κατάθλιψης συγκριτικά με τις περιπτώσεις όπου χρησιμοποιείται μόνο μία πηγή πληροφορίας. Επιπλέον υπήρξε 6% αύξηση της απόδοσης του συστήματος χρησιμοποιώντας συνδυασμό AAM χαρακτηριστικών με τα Baseline LGBP-TOP χαρακτηριστικά, συγκριτικά με τη περίπτωση αποκλειστικής χρήσης των Baseline LGBP-TOP χαρακτηριστικών του AVEC 2014 Dataset .

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Ανάλυση Πολυμεσικού Περιεχομένου

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Depression, Video Analysis, Audio Analysis

ABSTRACT

A considerable portion of the world population is affected by depression. According to data from the World Health Organization, until 2012 there were at least 350 million people living with some form of depression. Mental illness often affects people of working age, a fact that has strong social implications causing significant losses and burdens to the economic system. For the assessment of Depression level, psychologists and psychiatrists base their evaluation of a patient's condition on questionnaires (e.g. Beck Depression Index-II - BDI-II) as well as expressive facial and vocal cues.

Within the context of this thesis we aim to predict the depression level of an individual in the BDI-II scale, utilizing facial and voice features, as well as, machine learning techniques.

In particular, the videos we used as input data stem from the Depression Recognition Sub-Challenge of the 2014 Audio-Visual Emotion Challenge and Workshop (AVEC 2014). The estimation of depression level was performed, using Support Vector Regressor (SVR) models, based on features extracted from AAMs (Active Appearance Models) as well as a variety of audio features.

Results indicated that the fusion of AAMs and Audio features leads to better performance compared to individual modalities. Moreover the performance was improved by 6% by fusing AAM and LGBP-TOP Features compared to the performance of Baseline LGBP-TOP features of AVEC 2014.

SUBJECT AREA: Multimedia Content Analysis

KEYWORDS: Depression, Video Analysis, Audio Analysis