

ΑΝΑΓΝΩΡΙΣΗ ΝΕΥΜΑΤΟΣ ΜΕ ΤΗ ΧΡΗΣΗ ΑΙΣΘΗΤΗΡΙΟΥ ΑΔΡΑΝΕΙΑΣ ΚΑΙ ΤΕΧΝΙΚΩΝ ΕΚΜΑΘΗΣΗΣ ΜΗΧΑΝΗΣ

Ρήγας Κουσκουρίδας, Μιχάλης Μπακαλούμης και Αντώνιος
Γαστεράτος

Δημοκρίτειο Πανεπιστήμιο Θράκης, Τμήμα Μηχανικών Παραγωγής &
Διοίκησης, Πανεπιστημιούπολη, Κιμμέρια, 671 00 Ξάνθη,
rkouskou@pme.duth.gr

ΠΕΡΙΛΗΨΗ

Στην παρούσα εργασία κατασκευάστηκε ένας αλγόριθμος, ο οποίος, βασισμένος στη θεωρία της ταξινομησίας προτύπων, είναι ικανός να κατατάξει επιτυχώς μια σειρά από κινήσεις, καταγεγραμμένες με τη βοήθεια του αισθητηρίου αδράνειας MT_x . Η κάθε μια από αυτές τις κινήσεις, αναπαριστά ένα νεύμα του χρήστη, και με τη σειρά του, αντιστοιχεί σε μια λέξη. Έτσι, ο απώτερος σκοπός της εργασίας ήταν η δημιουργία ενός μικρού αυτοματοποιημένου λεξιλογίου, το οποίο, μπορεί να χρησιμοποιηθεί από άτομα με αναπηρίες για την απλούστερη εκτέλεση ορισμένων εργασιών. Η εργασία μπορεί να έχει ενδιαφέρουσες εφαρμογές στα πεδία των συστημάτων αυτόματου ελέγχου, της τεχνητής νοημοσύνης κ.ά.

Λέξεις κλειδιά: Εκμάθηση μηχανής, αισθητήριο αδράνειας, ευφυείς ταξινομητές.

1 ΕΙΣΑΓΩΓΗ

Όσο η επιστήμη εξελίσσεται, τόσο αυξάνεται και το πλήθος των μηχανών που χρησιμοποιεί καθημερινά ο άνθρωπος. Ιδιαίτερα τις τελευταίες δεκαετίες έχουν καταγραφεί σημαντικές τεχνολογικές ανακαλύψεις. Αυτές αφορούν, είτε εξεζητημένες εφαρμογές, όπως για παράδειγμα ένας ρομποτικός βραχίονας που χρησιμοποιείται για ιατρικές επεμβάσεις, είτε εφαρμογές στην καθημερινή ζωή του καθενός, όπως η χρήση μιας ηλεκτρικής κουζίνας. Όσο το πλήθος των μηχανών αυξάνεται, τόσο αυξάνεται και η εξάρτηση του ανθρώπου από αυτές. Έτσι, αποκτά όλο και μεγαλύτερη σημασία ο τρόπος με τον οποίο χρησιμοποιείται η μηχανή, και κατά συνέπεια η διαδικασία αλληλεπίδρασης ανθρώπου μηχανής (Human Machine Interface - HMI). Η αλληλεπίδραση αυτή μπορεί να είναι είτε τόσο απλή, όσο η χρήση ενός εργαλείου κηπουρικής από τον άνθρωπο και τόσο πολύπλοκη, όσο η χρήση του πίνακα ελέγχου ενός αεροπλάνου. Γενικότερα, ο σκοπός της HMI είναι να καταστήσει την λειτουργικότητα της τεχνολογίας αυτονόητη, στα πλαίσια του αποδοτικότερου χειρισμού και ελέγχου των μηχανών με στόχο τη «σοφότερη» λήψη αποφάσεων από τον άνθρωπο (Han, 2010). Από την άλλη, η επιστήμη της εκμάθησης μηχανής στοχεύει στη δημιουργία μηχανών ικανών να μαθαίνουν, δηλαδή να βελτιώνουν την απόδοσή τους σε κάποιους τομείς μέσω της αξιοποίησης προηγούμενης γνώσης και εμπειρίας. Ένας αρκετά γενικός ορισμός που θα μπορούσε να δοθεί για μια διαδικασία εκμάθησης μηχανής δίνεται από τον Mitchell (1997): «Ένα πρόγραμμα υπολογιστή λέμε ότι μαθαίνει από την εμπειρία E ως προς κάποια κλάση εργασιών T και μέτρο απόδοσης P , αν η απόδοσή του σε εργασίες από το T , όπως μετριέται από το P , βελτιώνεται μέσω της εμπειρίας E ». Ωστόσο, δεν είναι ακόμη προσιτή η δημιουργία μηχανών που να μαθαίνουν τόσο καλά και τόσο μεγάλη ποικιλία πραγμάτων όσο ο άνθρωπος. Παρ' όλα αυτά, έχουν αναπτυχθεί αλγόριθμοι για συγκεκριμένες περιοχές εκμάθησης, οι οποίοι έχουν επιτρέψει την εμφάνιση εμπορικών εφαρμογών με σημαντική επιτυχία. Από τη μια, για προβλήματα όπως η αναγνώριση φωνής και η εξαγωγή πληροφοριών (data mining) από μεγάλες βάσεις δεδομένων, η χρήση αλγορίθμων αυτού του πεδίου αποτελεί το αποδοτικότερο μέσο. Από την άλλη, όμως, έχουν σχεδιαστεί προγράμματα ικανά από το να μαθαίνουν να παίζουν τάβλι σε επίπεδο ανάλογο με των παγκόσμιων πρωταθλητών (Wiering, 2010) μέχρι να μαθαίνουν να οδηγούν αυτόνομα οχήματα σε δημόσιες λεωφόρους (Borges, 2009).

Οι βασισμένες στα στιγμιότυπα (instance-based, εν συντομία IB) μέθοδοι μάθησης έχουν μια θεμελιώδη διαφορά από τις άλλες μεθόδους μάθησης που έχουν αναπτυχθεί: δεν κατασκευάζουν ένα γενικό και ρητά διατυπωμένο μοντέλο που προσεγγίζει τη συνάρτηση-στόχο καθολικά. Το μόνο που κάνουν στη φάση της μάθησης είναι να αποθηκεύουν τα δεδομένα εκπαίδευσης, γι' αυτό είναι γνωστές

και ως τεχνικές βασισμένες στη μνήμη (memory-based). Η γενίκευση πέρα από τα παρατηρημένα δεδομένα γίνεται κάθε φορά που εμφανίζεται ένα νέο στιγμιότυπο προς κατάταξη - ταξινόμηση. Τότε, ένα σύνολο από σχετιζόμενα με αυτό γνωστά στιγμιότυπα ανακαλείται από τη μνήμη και χρησιμοποιείται για την κατάταξη του νέου στιγμιότυπου. Κατά καιρούς έχουν παρουσιαστεί αρκετοί ταξινομητές ικανοί να κατατάζουν επιτυχώς σειρές στιγμιότυπων λαμβάνοντας υπόψη a priori γνώση. Οι πιο γνωστοί αλγόριθμοι είναι η μέθοδος των ελαχίστων τετραγώνων (Least Squares - LS), η τεχνική των k- κοντινότερων γειτόνων (k-nearest neighbor - kNN), οι μηχανές διανυσμάτων υποστήριξης (Support Vector Machines - SVM) και η μέθοδος της προσαρμοστικής ώθησης (AdaBoost).

1.1 ΟΡΙΣΜΟΣ ΠΡΟΒΛΗΜΑΤΟΣ ΚΑΙ ΣΤΟΧΟΣ ΤΗΣ ΠΑΡΟΥΣΑΣ ΕΡΓΑΣΙΑΣ

Μια από τις πιο ενδιαφέρουσες εφαρμογές στη ρομποτική είναι η αναγνώριση κίνησης, όπως είναι τα νεύματα κεφαλής ανθρώπων με προβλήματα ομιλίας. Πιο συγκεκριμένα, ένα ευφρές σύστημα αισθητηρίων μπορεί να φανεί πολύ χρήσιμο σε έναν άνθρωπο με φωνητικά προβλήματα, σε περίπτωση που μπορεί να αναγνωρίσει είτε τα νεύματα του κεφαλιού του είτε τις χειρονομίες του. Η παρούσα έρευνα εστιάζει στην πρώτη περίπτωση την ίδια ώρα που έχουν παρουσιαστεί αρκετές έρευνες προωθώντας επαρκείς λύσεις στο πρόβλημα με τη χρήση αισθητηρίων όρασης (Karoor, 2001, Kawato, 2000, Peng, 2005, Sidner, 2006). Από την άλλη, τα αισθητήρια αδράνειας έχουν χρησιμοποιηθεί ευρέως σε διάφορες εφαρμογές που στοχεύουν είτε στην εκτίμηση του διανύσματος θέσης στόχων (Kygiakoulis, 2010) είτε στην υποβοηθούμενη φυσική αποκατάσταση ανθρώπων (Jovanov, 2005). Στόχος της παρούσας εργασίας είναι η αναγνώριση νευμάτων του κεφαλιού με τη χρήση του αισθητηρίου αδράνειας MT_x (Xsens, 2010) και ευφρών ταξινομητών. Συγκεκριμένα, τοποθετήσαμε το τελευταίο στο κεφάλι ενός ανθρώπου – στόχου, ενώ στη συνέχεια το υποκείμενο δέχθηκε ερωτήσεις με τυποποιημένες απαντήσεις της μορφής «ΝΑΙ», «ΣΙΓΟΥΡΑ ΝΑΙ», «ΟΧΙ», «ΣΙΓΟΥΡΑ ΟΧΙ », «ΔΕΝ ΞΕΡΩ», «ΑΡΙΣΤΕΡΑ», και «ΔΕΞΙΑ». Μέρος των καταγεγραμμένων κινήσεων του αισθητηρίου χρησιμοποιήθηκαν για την εκπαίδευση του συστήματος (μια μέτρηση για κάθε κίνηση - απάντηση) ενώ οι υπόλοιπες για τον έλεγχο της απόδοσής του. Τα πειραματικά αποτελέσματα ενισχύουν την ορθότητα της επιλογής του αισθητηρίου αδράνειας ως μέσο αναγνώρισης νευμάτων καθώς σε όλες τις περιπτώσεις το προτεινόμενο σύστημα ανιχνεύει επιτυχώς τις προς αναγνώριση κινήσεις. Το υπόλοιπο κείμενο που ακολουθεί είναι διαρθρωμένο ως εξής: Στην Παράγραφο 2 παρουσιάζεται αναλυτικά και λεπτομερώς το αισθητήριο αδράνειας που χρησιμοποιήθηκε ενώ, η Παράγραφος 3 είναι αφιερωμένη στην παράθεση των ευφρών ταξινομητών και των τεχνικών εκτίμησης σφάλματος που έλαβαν χώρα στα πλαίσια της παρούσας έρευνας. Τα πειραματικά αποτελέσματα παρατίθενται στην Ενότητα 4, καθώς τα συμπεράσματα και η μελλοντική εργασία παρουσιάζονται στην Παράγραφο 5.

2 ΤΟ ΑΙΣΘΗΤΗΡΙΟ MT_x

Το αισθητήριο MT_x , που παρουσιάζεται στο Σχήμα 1κ), είναι ένας μικρός και ακριβής ανιχνευτής προσανατολισμού. Μπορεί επίσης να περιγραφεί ως μια μικρής κλίμακας, ολοκληρωμένη μονάδα μέτρησης αδρανειακών μεγεθών (inertial measurement unit). Έχει ενσωματωμένα 3D (τριδιάστατα) μαγνητόμετρα, με ενσωματωμένο επεξεργαστή ικανό να υπολογίζει γωνίες roll, pitch και yaw σε πραγματικό χρόνο. Επιπροσθέτως, παρέχει 3D drift-free προσανατολισμό και μετράει 3 φυσικά μεγέθη, με τρεις συνιστώσες σε κάθε ένα από τους 3 άξονες. Τα μεγέθη αυτά είναι η γωνιακή ταχύτητα που μετράται από τα γυροσκόπια, η γραμμική επιτάχυνση από τα επιταχυνσιόμετρα και η ένταση του μαγνητικού πεδίου της γης. Οι μονάδες μέτρησης της γραμμικής επιτάχυνσης είναι σε m/s^2 , της γωνιακής ταχύτητας σε rad/s και της έντασης του μαγνητικού πεδίου σε a.u. Επίσης, χρησιμοποιείται σε διάφορα επιστημονικά πεδία όπως στη βιομηχανική (biomechanics), στον αθλητισμό, στην εικονική πραγματικότητα (virtual reality), στα κινούμενα σχέδια (animation) και στην απαθανάτιση εικόνας (motion capture). Το MT_x υπολογίζει τον προσανατολισμό του συστήματος συντεταγμένων του αισθητηρίου S, σε σχέση με το σύστημα συντεταγμένων της γης G. Το σύστημα συντεταγμένων της γης που χρησιμοποιείται ως αναφορά, καθορίζεται σύμφωνα με τον κανόνα του δεξιού χεριού ως εξής: Ο άξονας X είναι θετικός όταν κοιτάει στον τοπικό μαγνητικό Νότο. Ο άξονας Y καθορίζεται σύμφωνα με το κανόνα του δεξιού χεριού (Δύση). Ο άξονας Z είναι θετικός όταν κοιτάει προς τα πάνω. Ο εξερχόμενος προσανατολισμός ορίζεται ως ο προσανατολισμός του S σε σχέση με το G. Η ανάλυση γωνίας κυμαίνεται στα επίπεδα των 0.05° RMS (τυπική απόκλιση ίση με 1σ για zero mean angular random walk), η στατική ακρίβεια για της γωνίες roll και pitch είναι μικρότερη των 0.5° ενώ εκείνης της γωνίας yaw μικρότερη της 1.0° . Επιπλέον, ο μέγιστος ρυθμός δειγματοληψίας αντιστοιχεί στα 120 Hz, γεγονός που καθιστά το αισθητήριο ικανό για εφαρμογές που απαιτούν εκτέλεση σε πραγματικό χρόνο.

Από την άλλη, το αισθητήριο αποτυγχάνει να αποδώσει πραγματικές μετρήσεις σε περιπτώσεις όπου οι δονήσεις σε κάθε άξονα ξεπερνούν την τιμή των 20000 m/s², η εισερχόμενη τάση δεν είναι μεταξύ των τιμών [-0.3V, 16V] και η θερμοκρασία δεν κυμαίνεται στο εύρος [-5 C°, 60 C°].

3 ΤΑΞΙΝΟΜΗΤΕΣ ΚΑΙ ΤΕΧΝΙΚΕΣ ΕΚΤΙΜΗΣΗΣ ΣΦΑΛΜΑΤΟΣ

Στα πλαίσια της παρούσας έρευνας επιλέχθηκαν τέσσερις ευρέως διαδεδομένοι ταξινομητές (γραμμικοί και μη), που παρουσιάζουν σχετικά μικρή ταχύτητα εκτέλεσης σε συνδυασμό με αξιοσημείωτη απόδοση κατηγοριοποίησης. Επιπρόσθετα, στην ενότητα που ακολουθεί παρουσιάζονται και οι πιο γνωστές τεχνικές εκτίμησης σφάλματος ταξινόμησης.

3.1 ΕΥΦΥΕΙΣ ΤΑΞΙΝΟΜΗΤΕΣ

Η μέθοδος των ελαχίστων τετραγώνων (Least Squares - LS): Αποτελεί αδιαμφισβήτητη την πλέον χρησιμοποιούμενη τεχνική ταξινόμησης καθώς συνδυάζει χαμηλό υπολογιστικό κόστος και σχετικά ικανοποιητική απόδοση. Η τελευταία αυξάνεται κατακόρυφα σε περιπτώσεις όπου τα στιγμιότυπα μπορούν να κατηγοριοποιηθούν με τη χρήση γραμμικών ταξινομητών.

Η μέθοδος των k -κοντινότερων γειτόνων (k -Nearest Neighbor - k NN): Αποτελεί την πιο βασική IB μέθοδο μάθησης καθώς επίσης, μπορεί να χαρακτηριστεί ως μια ευρέως χρησιμοποιούμενη τεχνική κατηγοριοποίησης που βασίζεται στη χρήση μέτρων βασισμένων στην απόσταση. Η κεντρική ιδέα είναι πως η τιμή της συνάρτησης-στόχου για ένα νέο στιγμιότυπο βασίζεται αποκλειστικά και μόνο στις αντίστοιχες τιμές των k πιο “κοντινών” του στιγμιότυπων εκπαίδευσης, τα οποία αποτελούν τους “γειτόνες” του. Τρία ζητήματα πρέπει να αποφασιστούν προκειμένου να καθοριστεί πλήρως ο αλγόριθμος: Ο ορισμός της απόστασης μεταξύ δύο στιγμιότυπων, δηλαδή μιας μετρικής πάνω στο χώρο των στιγμιότυπων (instance space), που θα εκφράζει την εγγύτητα, ή αλλιώς την “ομοιότητα” μεταξύ των στιγμιότυπων, ο τρόπος συνδυασμού των τιμών των k κοντινότερων γειτόνων και τέλος, η τιμή του k .

Μηχανές διανυσμάτων υποστήριξης (Support Vector Machines - SVM): Χαρακτηρίζεται ως ένα είδος συγκερασμού γραμμικών μοντέλων και IB μάθησης. Ο όρος των γραμμικών μοντέλων αναφέρεται στην κατηγορία μοντέλων, των οποίων η συνάρτηση στόχου αποτελεί ένα γραμμικό συνδυασμό των χαρακτηριστικών του προβλήματος με ένα σύνολο βαρών προς εκμάθηση. Κυριότερος εκπρόσωπος της κατηγορίας αυτής είναι τα μοντέλα γραμμικής παλινδρόμησης (linear regression models). Στόχος του αλγορίθμου είναι η επιλογή ενός μικρού αριθμού στιγμιότυπων εκπαίδευσης από κάθε κλάση, των διανυσμάτων υποστήριξης (support vectors), που συνορεύουν στο χώρο του προβλήματος με στιγμιότυπα άλλων κλάσεων. Τα επιλεγμένα στιγμιότυπα χρησιμοποιούνται για την κατασκευή μιας γραμμικής συνάρτησης διάκρισης (discriminant function), ικανής να τα διαχωρίσει όσο το δυνατόν καλύτερα. Τα συστήματα ταξινόμησης που βασίζονται στον αλγόριθμο αυτό αποτελούν σήμερα μια από τις δημοφιλέστερες προσεγγίσεις στο χώρο της κατηγοριοποίησης κειμένου. Αυτό οφείλεται στην ευρωστία, στην αποτελεσματικότητα και στην ταχύτητα που επιδεικνύουν, αλλά και στη ικανότητά τους να παράγουν μη γραμμικές επιφάνειες απόφασης. Καθιστούν, έτσι, υπολογιστικά εφικτή την επίλυση ενός μεγάλου αριθμού πρακτικών προβλημάτων μάθησης που δεν μπορούν να αντιμετωπιστούν από άλλα γραμμικά μοντέλα. Υπάρχουν διάφορες παραλλαγές των SVM που καθορίζονται κυρίως από τον «πυρήνα» (kernel) που θα χρησιμοποιηθεί για την εκπαίδευση.

Η μέθοδος της προσαρμοστικής ώθησης (AdaBoost): Αποτελεί μία γενικευμένη τεχνική βελτίωσης της αποτελεσματικότητας οποιουδήποτε αλγορίθμου εκπαίδευσης. Θεωρητικά, το boosting μπορεί να μειώσει κατά πολύ το σφάλμα που κάνουν αδύναμες μηχανές ταξινόμησης (weak learners). Δηλαδή, μηχανές που έχουν επίδοση λίγο καλύτερη από αυτήν της τυχαίας επιλογής (σε πρόβλημα 2 τάξεων η τυχαία επιλογή αντιστοιχεί σε επίδοση 1/2, σε πρόβλημα 3 τάξεων σε επίδοση 1/3 κτλ.). Υπάρχουν διάφοροι τύποι μεθόδων boosting. Όλες συνίστανται στην σειριακή εκπαίδευση μίας ομάδας ταξινομητών. Το σύνολο εκπαίδευσης για κάθε μηχανή προκύπτει από το αρχικό ανάλογα με την επίδοση της προηγούμενης στη σειρά μηχανής. Στο τέλος, τα αποτελέσματα των μηχανών συνδυάζονται με κάποιο τρόπο ώστε να προκύψει το τελικό αποτέλεσμα της μεθόδου. Υπάρχουν τέσσερις εκδοχές του αλγορίθμου, ο διακριτός, ο πραγματικός, ο ήπιος και ο Logistic AdaBoost.

3.2 ΤΕΧΝΙΚΕΣ ΕΚΤΙΜΗΣΗΣ ΣΦΑΛΜΑΤΟΣ ΤΑΞΙΝΟΜΗΣΗΣ

Από τη στιγμή που η αξιολόγηση ενός ταξινομητή θα πρέπει να γίνεται με βάση την επίδοση του στην κατηγοριοποίηση άγνωστων δεδομένων, η στρατηγική που χρησιμοποιείται για την εκτίμηση του σφάλματος ταξινόμησης κάποιας μεθόδου διαιρεί το σύνολο των διαθέσιμων προτύπων σε δύο τμήματα: στο σύνολο εκπαίδευσης (training set) που χρησιμοποιείται για την κατασκευή του ταξινομητή και στο σύνολο ελέγχου (testing set) που χρησιμοποιείται για τον υπολογισμό του σφάλματος γενίκευσης. Οι τεχνικές εκτίμησης σφάλματος διαφέρουν μεταξύ τους κυρίως στον τρόπο που γίνεται η διάσπαση των δεδομένων στα δύο σύνολα.

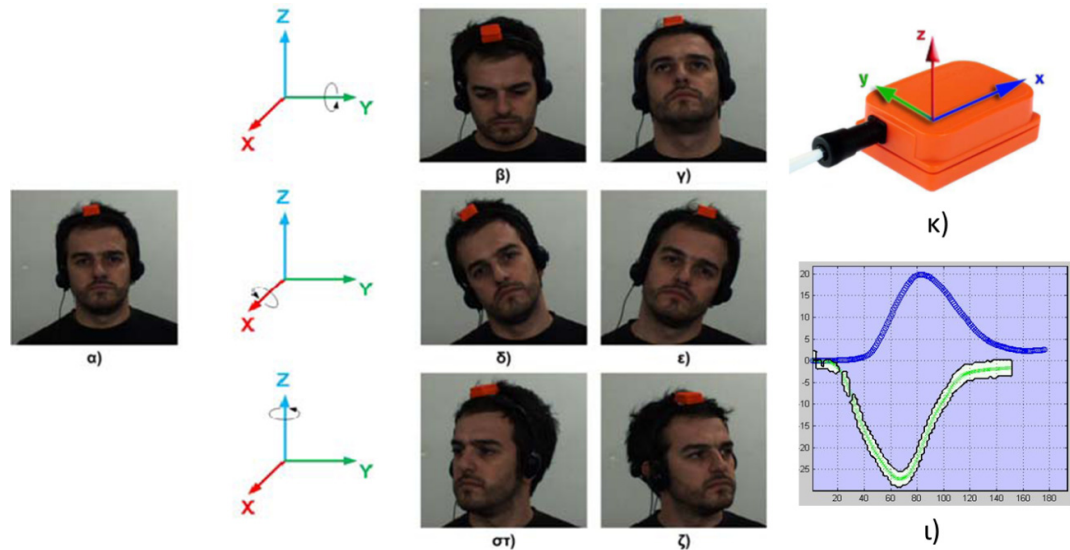
Η τεχνική εκτίμησης σφάλματος Holdout: Το σύνολο των διαθέσιμων δεδομένων χωρίζεται με τυχαίο τρόπο σε δύο ανεξάρτητα σύνολα δεδομένων. Συνήθως ο διαχωρισμός πραγματοποιείται με τέτοιο τρόπο ώστε να χρησιμοποιούμε το 1/3 των δεδομένων για την εκτίμηση τους (testing set) και τα 2/3 για εκπαίδευση (training set). Με άλλα λόγια, η holdout μέθοδος διατηρεί ένα ορισμένο ποσό δεδομένων για δοκιμή και χρησιμοποιεί το υπόλοιπο για εκπαίδευση. Το φυσικό μέτρο απόδοσης για τα προβλήματα ταξινόμησης είναι το success rate (ακρίβεια), το οποίο αντιπροσωπεύει το ποσοστό των σωστά ταξινομημένων περιπτώσεων σε ολόκληρο το σύνολο περιπτώσεων και το error rate το οποίο αντιπροσωπεύει το αντίστοιχο ποσοστό των λανθασμένων εκτιμήσεων.

Η τεχνική εκτίμησης σφάλματος Resubstitution: Με τη μέθοδο αυτή ελέγχεται ο αλγόριθμος ταξινόμησης, χρησιμοποιώντας ολόκληρο το σύνολο δεδομένων, το οποίο έχει προηγουμένως χρησιμοποιηθεί και ως training set. Στην ουσία δε χρησιμοποιείται testing set και τα αποτελέσματα του training set μπορούν να ερμηνευτούν απευθείας, παρουσιάζοντας, κατά συνέπεια, σχεδόν μηδενικό υπολογιστικό κόστος. Σύμφωνα με τα παραπάνω, η resubstitution αντιτίθεται στις μεθόδους της cross-validation (ή των bootstrap μεθόδων), καθώς η δεύτερη συχνά απαιτεί αρκετά μεγάλο διάστημα χρόνου για τις μετατροπές των συνόλων εκπαίδευσης και ελέγχου. Το error rate της τεχνικής αυτής λαμβάνεται από τα δεδομένα εκπαίδευσης (training set). Δεν είναι πάντα 0%, αλλά συνήθως πολύ χαμηλό, διότι οι αλγόριθμοι ταξινόμησης περιέχουν διάφορες, συχνά στατιστικές, παραμέτρους οι οποίες οδηγούν σε αβεβαιότητα ως προς το αποτέλεσμα.

4 ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Αρχικά τοποθετήθηκε το αισθητήριο MT_x στο κεφάλι ενός ανθρώπου – στόχου με σκοπό την καταγραφή ορισμένων κινήσεων κάθε μια εκ των οποίων αντιστοιχεί σε ένα νεύμα του ανθρώπινου κεφαλιού, τα οποία με τη σειρά τους αντιπροσωπεύουν συγκεκριμένες λέξεις. Επιλέχθηκαν επτά λέξεις: «ΔΕΞΙΑ», «ΑΡΙΣΤΕΡΑ», «ΝΑΙ», «ΟΧΙ», «ΣΙΓΟΥΡΑ ΝΑΙ», «ΣΙΓΟΥΡΑ ΟΧΙ» και «ΔΕΝ ΞΕΡΩ». Η κίνηση «ΑΡΙΣΤΕΡΑ» ορίστηκε ως η περιστροφή γύρω από τον άξονα X κατά την αρνητική φορά (αριστερόστροφα), δηλαδή η μεταβολή της γωνίας Roll. Η κίνηση «ΔΕΞΙΑ», αντιπροσωπεύει την περιστροφή, επίσης, γύρω από τον άξονα X, αλλά κατά την θετική φορά (δεξιόστροφα). Η λέξη «ΝΑΙ», χαρακτηρίζεται ως η περιστροφή γύρω από τον άξονα Y κατά την θετική φορά, που αντιστοιχεί σε μεταβολές της γωνία Pitch. Από την άλλη, η λέξη «ΟΧΙ», αναπαριστάται από μετρήσεις γύρω από τον άξονα Y, αλλά κατά την αντίθετη πλέον φορά (αρνητική). Επιπροσθέτως, για τις λέξεις «ΣΙΓΟΥΡΑ ΝΑΙ» και «ΣΙΓΟΥΡΑ ΟΧΙ», ισχύει ότι και για τις «ΝΑΙ» και «ΟΧΙ» αντίστοιχα, με μια διαφορά. Αυτή, έχει να κάνει με τη συχνότητα της κίνησης και τη διάρκεια της. Πρακτικά, οι «ΣΙΓΟΥΡΑ ΝΑΙ» και «ΣΙΓΟΥΡΑ ΟΧΙ», μπορούν να περιγραφούν ως τρεις κινήσεις «ΝΑΙ» και «ΟΧΙ», αντίστοιχα, στη σειρά. Η ταχύτητα της κίνησης που αντιστοιχεί στις πρώτες, είναι μεγαλύτερη. Τέλος, η κίνηση «ΔΕΝ ΞΕΡΩ», ορίστηκε ως η περιστροφή γύρω από τον άξονα Z, που αντιστοιχούν σε μεταβολές της γωνίας Yaw. Αυτή τη φορά, υπάρχει, κίνηση και προς τις δύο κατευθύνσεις. Αρχικά η κίνηση είναι αριστερόστροφη, κατά την αρνητική φορά, και στη συνέχεια δεξιόστροφη, κατά την θετική φορά. Στο Σχήμα 1 δίνονται οι οπτικές αναπαραστάσεις των κινήσεων, για την καλύτερη κατανόηση τους. Δεν αναπαριστούνται οι λέξεις «ΣΙΓΟΥΡΑ ΝΑΙ» και «ΣΙΓΟΥΡΑ ΟΧΙ», μιας και ανήκουν στην ίδια κατηγορία με τις «ΝΑΙ» και «ΟΧΙ». Για κάθε κίνηση, ορίστηκε ως θεμελιώδης μεταβολή εκείνη που πραγματοποιείται σε μια μόνο γωνία Euler. Σε αυτή δηλαδή που παρουσιάζεται η μεγαλύτερη διακύμανση. Εν τούτοις, υπήρξαν, μεταβολές και των άλλων δύο γωνιών. Στην περίπτωση, για παράδειγμα, μεταβολής της γωνίας Roll, υπήρξαν μεταβολές και στις γωνίες Pitch και Yaw. Δεν λήφθηκαν υπόψη, όμως, μιας και η κίνηση στις δευτερευουσες γωνίες ήταν μικρή και όχι ικανή να επηρεάσει το τελικό αποτέλεσμα. Για κάθε λέξη, αποθηκεύτηκαν (10) μετρήσεις, δημιουργώντας κατά συνέπεια ένα σύνολο (75) μετρήσεων μέσα στις οποίες συμπεριλαμβάνονται κάποιες κινήσεις του στόχου που δεν μπορούν να τυποποιηθούν. Για την

εκπαίδευση του συστήματος χρησιμοποιήθηκε μια κίνηση από κάθε λέξη ενώ οι υπόλοιπες μετρήσεις λαμβάνονται υπόψη κατά τη διαδικασία του ελέγχου του συστήματος.

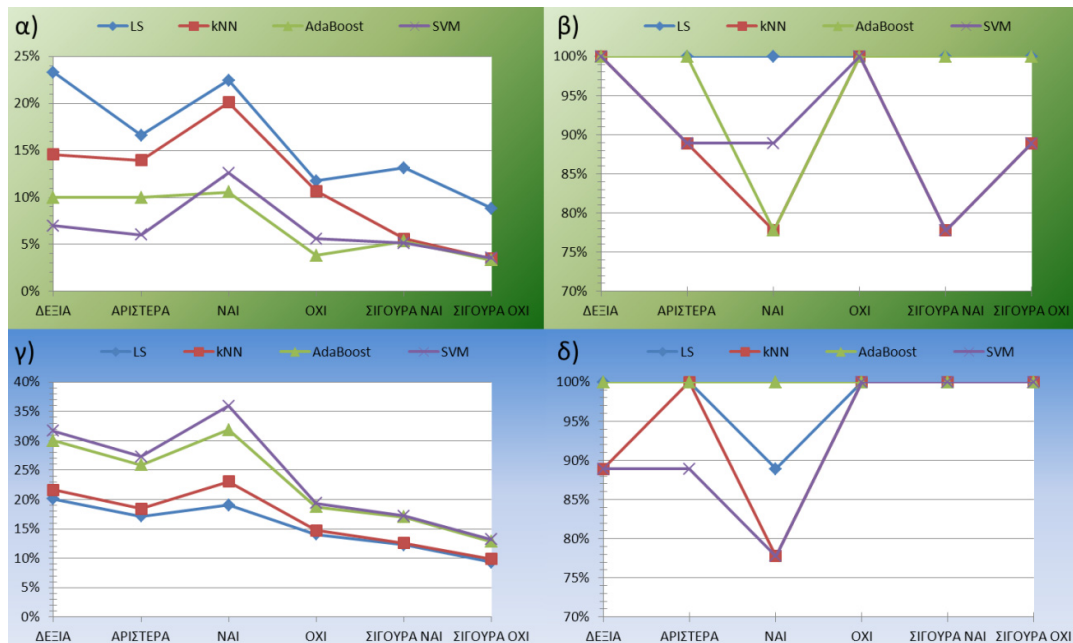


Σχήμα 1 Η εικόνα (α) δείχνει την αρχική θέση του υποκειμένου. Στις (β) και (γ) αναπαρίστανται οι κινήσεις των λέξεων «ΝΑΙ» και «ΟΧΙ» αντίστοιχα. Στα αριστερά τους φαίνεται η περιστροφή στον άξονα Y. Στις (δ) και (ε), οι κινήσεις «ΔΕΞΙΑ» και «ΑΡΙΣΤΕΡΑ» αντίστοιχα ενώ φαίνεται η περιστροφή γύρω από τον άξονα X. Επιπροσθέτως, στις (στ) και (ζ) παρουσιάζεται η κίνηση «ΔΕ ΞΕΡΩ», καθώς επίσης και η περιστροφή στον αντίστοιχο άξονα Z του αισθητηρίου. Η εικόνα (ι) αντιπροσωπεύει το γραφικό αποτέλεσμα μιας ταξινόμησης με τη χρήση των SVM, ενώ το αισθητήριο MT_x που χρησιμοποιήθηκε παρουσιάζεται στο σχήμα (κ).

Είναι προφανές ότι το προτεινόμενο σύστημα μπορεί να διαιρεθεί σε δύο διακριτά στάδια, με το πρώτο να αντιστοιχεί στην φάση εκπαίδευσης, ενώ το δεύτερο στην φάση εκτέλεσης και ελέγχου. Η φάση εκτέλεσης, με τη σειρά της περιλαμβάνει τη δημιουργία ενός φίλτρου ικανού να κατατάσσει αρχικά τις κινήσεις ανάλογα με τον άξονα που παρουσιάζει τις μεγαλύτερες μεταβολές. Με αυτόν τον τρόπο, είμαστε σε θέση να καταλάβουμε αμέσως ότι πιθανή μεταβολή των τιμών του άξονα Y αντιστοιχεί στις λέξεις είτε «ΝΑΙ», είτε «ΣΙΓΟΥΡΑ ΝΑΙ», είτε «ΟΧΙ» είτε «ΣΙΓΟΥΡΑ ΟΧΙ». Σαν επόμενο βήμα, χρησιμοποιώντας τους ταξινομητές που παρουσιάστηκαν στην Ενότητα 3, κατηγοριοποιούμε με αξιοσημείωτη ακρίβεια οποιαδήποτε νέα μέτρηση. Το Σχήμα 2 αναπαριστά ενδεικτικά το μέσο εκτιμώμενο σφάλμα ταξινόμησης και για τις δυο τεχνικές εκτίμησης σφάλματος. Ο ταξινομητής που παρουσίασε την καλύτερη απόδοση ταξινόμησης ήταν ο LS, διότι η φύση των στιγμιότυπων ευνοούσε τη χρήση τυπικών γραμμικών ταξινομητών. Οι αλγόριθμοι SVM, AdaBoost και kNN αν και παρουσιάζουν σχετικά μικρό σφάλμα ταξινόμησης, πλην ελαχίστων εξαιρέσεων, είναι πολύ πιο χρήσιμοι σε περιπτώσεις ταξινόμησης μη γραμμικών μοντέλων. Επιπρόσθετα, πρέπει να σημειωθεί ότι η ταξινόμηση έχει νόημα μόνο σε περιπτώσεις που παρουσιάζουν μεταβολές στις γωνίες Roll και Pitch (λέξεις «ΔΕΞΙΑ», «ΑΡΙΣΤΕΡΑ», «ΝΑΙ», «ΟΧΙ», «ΣΙΓΟΥΡΑ ΝΑΙ», και «ΣΙΓΟΥΡΑ ΟΧΙ»), και όχι σε εκείνες της Yaw («ΔΕΝ ΞΕΡΩ»).

5 ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην παρούσα εργασία παρουσιάστηκε μια νέα μέθοδος αναγνώρισης νευμάτων κεφαλιού με τη χρήση ενός αισθητηρίου αδράνειας. Οι κινήσεις του κεφαλιού τυποποιήθηκαν με στόχο τη λήψη αποφάσεων σε πιθανές ερωτήσεις που δέχονται ως απαντήσεις τις λέξεις «ΔΕΞΙΑ», «ΑΡΙΣΤΕΡΑ», «ΝΑΙ», «ΟΧΙ», «ΣΙΓΟΥΡΑ ΝΑΙ», «ΣΙΓΟΥΡΑ ΟΧΙ» και «ΔΕΝ ΞΕΡΩ». Επιπροσθέτως, στα πλαίσια της παρούσας έρευνας πραγματοποιήθηκε μια συγκριτική μελέτη διαφόρων ευφώνων ταξινομητών με στόχο την υιοθέτηση εκείνου που παρουσιάζει το μικρότερο σφάλμα ταξινόμησης. Η φύση των στιγμιότυπων καθόρισε σαν κυρίαρχη μέθοδο εκείνη των ελαχίστων τετραγώνων ενώ οι εναπομείναντες τεχνικές (SVM, AdaBoost και kNN) παρουσιάζουν μικρότερο σφάλμα ταξινόμησης σε διαδικασίες κατηγοριοποίησης που απαιτούν μη-γραμμικούς ταξινομητές. Μελλοντικοί στόχοι της έρευνας αποτελεί τόσο ο εμπλουτισμός του λεξιλογίου με περισσότερες λέξεις όσο και η δημιουργία μιας διαπαφής χρήστη μηχανής ικανής να ανταποκρίνεται στις απαιτήσεις του τελευταίου σε συνθήκες πραγματικού χρόνου.



Σχήμα 2 Μέσο εκτιμώμενο σφάλμα ταξινόμησης και ποσοστό επιτυχίας αναγνώρισης νεύματος για τους τέσσερις αλγόριθμους που χρησιμοποιήθηκαν. Τα σχήματα (α) και (γ) παρουσιάζουν την απόδοση των αλγορίθμων όταν για τεχνική εκτίμησης σφάλματος χρησιμοποιηθεί η holdout (για ποσοστά εκπαίδευσης 80% επί του συνόλου) και η resubstitution αντίστοιχα. Τα σχήματα (β) και (δ) αναπαριστούν το ποσοστό επιτυχημένων αναγνωρίσεων νεύματος κεφαλής που σχετίζονται με τα διαγράμματα (α) και (γ).

6 ΒΙΒΛΙΟΓΡΑΦΙΑ

- Borges A. P., Ribeiro R., Avila B. C., Enembreck F., Scalabrin E. E. (2009), "A learning agent to help drive vehicles", *International Conference on Computer Supported Cooperative Work in Design*, pp.282-287.
- Han S., Hong J., Jeong S., Hahn M. (2010), "Robust GSC-based speech enhancement for human machine interface", *IEEE Transactions on Consumer Electronics*, vol.56, no.2, pp.965-970
- Jovanov E, Milenkovic A, Otto C, de Groen P.C. (2005), "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation", *International Journal of Neuroengineering and Rehabilitation*, vol.2, no.1, pp.1-6
- Kapoor A., Picard, R. W. (2001), "A real-time head nod and shake detector", *International Workshop on Perceptive User interfaces*, pp1-5.
- Kawato S., Ohya J. (2000), "Real-time detection of nodding and head-shaking by directly detecting and tracking the between-eyes", *IEEE International Conference on Automatic Face and Gesture Recognition*, pp53-59.
- Kyriakoulis N., Gasteratos A. (2010), "Color based monocular visuo-inertial 3D pose estimation of a volant robot", *IEEE Transactions on Instrumentation and Measurement*, to appear.
- Mitchell T.M. (1997), "Machine learning", McGraw-Hill International Editions
- Peng L., Mandun Z., Xinshan Z., Yangsheng W. (2005), "Head nod and shake recognition based on multi-view model and hidden markov model", *International Conference on Computer Graphics, Imaging and Visualization*, pp.61-64
- Sidner C. L., Lee C., Morency L. P., Forlines C. (2006), "The effect of head-nod recognition in human-robot conversation", *IEEE International Conference on Human Robot Interaction*, pp.290-296.
- Wiering M. A. (2010), "Self-play and using an expert to learn to play backgammon with temporal difference learning", *International Journal of Intelligent Learning Systems & Applications*, vol. 2, no.2, pp.57-68.
- Xsens Technologies B. V. (2010), "<http://www.xsens.com/en/general/mtx>"