



**ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΝΩΝΙΩΝ**

Γεώργιος Κουρουπέτρογλου, Καρέν Ξιπτερίδης και Ευάγγελος Μιτσόπουλος

Τεχνικές πρόσβασης σε υπολογιστικά περιβάλλοντα

**ΑΘΗΝΑ
2001**

Επικοινωνία: Γεώργιος Κουρουπέτρογλου
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών,
Τμήμα Πληροφορικής και Τηλεπικοινωνιών,
Πανεπιστημιόπολη, Ιλίσια,
15784 Αθήνα

Τηλέφωνο: 210 7275305
Fax: 210 6018677

Ηλεκτρονικό Ταχυδρομείο: koupe@di.uoa.gr
Ιστοθέση: www.e-bility.gr

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ.....	5
2. ΠΛΗΚΤΡΟΛΟΓΙΑ ΚΑΙ ΠΛΗΚΤΡΑ	6
2.1. ΕΙΣΑΓΩΓΗ	6
2.2. ΔΙΑΤΑΞΕΙΣ ΠΛΗΚΤΡΟΛΟΓΙΩΝ (KEYBOARD LAYOUTS).....	7
2.2.1. Η Κανονική Διάταξη (QWERTY).....	7
2.2.2. Η Απλοποιημένη διάταξη DVORAK (Dvorak Simplified Keyboard)	8
2.2.3. Αλφαβητικές Διατάξεις (Alphabetical Keyboards)	10
2.2.4. Άλλες διατάξεις πληκτρολογίων	10
2.2.5. Πληκτρολόγια Εισαγωγής Δεδομένων (Data-Entry Keyboards).....	11
2.2.6. Συμπέρασμα.....	12
2.3. ΦΥΣΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΩΝ ΠΛΗΚΤΡΩΝ ΚΑΙ ΤΩΝ ΠΛΗΚΤΡΟΛΟΓΙΩΝ	13
2.3.1. Ύψος και Κλίση των Πληκτρολογίων.....	13
2.3.2. Προφίλ Πληκτρολογίων (Keyboard Profile).....	13
2.3.3. Μορφή και Μέγεθος των Πλήκτρων	15
2.3.4. Χρώμα και Ετικέτες (Color and Labeling).....	16
2.3.5. Πλήκτρα Ειδικού Σκοπού (Special Purpose Keys)	16
2.3.6. Διευθέτηση των Πλήκτρων Ελέγχου του Δρομέα.....	17
2.4. ΕΝΑΛΛΑΚΤΙΚΑ ΠΛΗΚΤΡΟΛΟΓΙΑ (ALTERNATIVE KEYBOARD DESIGNS)	17
2.4.1. Διαχωρισμένα Πληκτρολόγια ή Σπαστά Πληκτρολόγια (Split Keyboards).....	17
2.4.2. Έγχορδα Πληκτρολόγια (Chord Keyboards)	22
2.4.3. Συμπεράσματα	24
3. ΣΥΣΚΕΥΕΣ ΕΝΔΕΙΞΗΣ (POINTING DEVICES).....	25
3.1. ΕΙΣΑΓΩΓΗ	25
3.2. ΣΥΣΚΕΥΕΣ ΕΜΕΣΗΣ ΕΝΔΕΙΞΗΣ (DIRECT POINTING DEVICES)	26
3.2.1. Οθόνες Αφής (Touch Screens).....	26
3.2.2. Φωτογραφίδα (Light Pens).....	29
3.2.3. Ράβδος κεφαλής (Head stick)	31
3.3. ΕΜΜΕΣΕΣ ΣΥΣΚΕΥΕΣ ΕΝΔΕΙΞΗΣ (INDIRECT POINTING DEVICES)	31
3.3.1. Πινακίδες Γραφικών (Graphic Tablets)	31
3.3.2. Ποντίκια (Mice).....	34
3.3.3. Ιχνόσφαιρες (Trackballs).....	36
3.3.4. Χειριστήρια Χειρός (Joysticks)	38
3.4. ΣΥΣΚΕΥΕΣ ΟΡΙΖΟΝΤΑ (DEVICES ON THE HORIZON).....	39
3.4.1. Είσοδος ελεγχόμενη με τα μάτια (Eye-controlled Input).....	39
3.4.2. Έλεγχος Κεφαλής (Head Tracking)	41
3.4.3. Είσοδος με χειρονομίες (Gestural and Spatial Input).....	41
3.5. ΣΥΜΠΕΡΑΣΜΑΤΑ	42
4. ΔΙΑΚΟΠΤΕΣ ΚΑΙ ΤΕΧΝΙΚΕΣ ΣΑΡΩΣΗΣ (SWITCHES AND SCANNING TECHNIQUES).....	44
4.1. ΕΙΣΑΓΩΓΗ	44
4.2. ΣΥΝΔΕΣΗ ΔΙΑΚΟΠΤΩΝ ΜΕ ΤΟΝ ΥΠΟΛΟΓΙΣΤΗ	45
4.3. ΠΑΡΟΥΣΙΑΣΗ ΑΝΤΙΚΕΙΜΕΝΩΝ ΕΠΙΛΟΓΗΣ	45
4.3.1. Σύνολο Επιλογών (Selection Set)	45
4.3.2. Τονισμός (Highlighting)	46
4.3.3. Ανάδραση (Feedback)	46
4.4. ΜΕΘΟΔΟΙ ΣΑΡΩΣΗΣ (SCANNING METHODS)	46
4.4.1. Μικρά σύνολα επιλογών (Small Selection Sets).....	46
4.4.2. Μεγάλα σύνολα επιλογών (Larger Selection Sets).....	47
4.4.3. Πτυσσόμενα μενού.....	47
4.5. ΤΕΧΝΙΚΕΣ ΣΑΡΩΣΗΣ (SCANNING TECHNIQUE)	47
4.5.1. Απλή Σάρωση (Simple Scan)	47
4.5.2. Ομαδική Σάρωση (Group Scan)	47
4.5.3. Κατευθυνόμενη Σάρωση (Directed Scan).....	48

4.5.4.	Αυτόματη Σάρωση (Autoscan).....	49
4.5.5.	Σάρωση Μετάπτωσης (Toggle Scan).....	49
4.5.6.	Σάρωση Μοναδικού Βήματος (Single Step Scan).....	49
4.5.7.	Επανεκκίνηση Σάρωσης (Restarting Scan).....	49
4.6.	ΔΙΑΚΟΠΤΗΣ ΜΕ ΑΓΓΙΓΜΑ ΤΗΣ ΓΛΩΣΣΑΣ (TONGUE TOUCH KEYPAD).....	49
5.	ΤΕΧΝΟΛΟΓΙΕΣ ΟΜΙΛΙΑΣ	51
5.1.	ΕΙΣΑΓΩΓΗ	51
5.2.	ΑΝΑΓΝΩΡΙΣΗ ΟΜΙΛΙΑΣ	51
5.2.1.	Υπαγόρευση	52
5.2.2.	Κατανόηση ομιλίας.....	53
5.2.3.	Αναγνώριση ομιλητή.....	54
5.2.4.	Αναγνώριση γλώσσας.....	56
5.2.5.	Ανθρώπινοι παράγοντες στην αναγνώριση ομιλίας.....	56
5.2.6.	Τύποι εφαρμογών της αναγνώρισης ομιλίας.....	58
5.2.7.	Τύποι λαθών αναγνώρισης.....	59
5.2.8.	Συστήματα αναγνώρισης ελληνικής ομιλίας.....	59
5.3.	ΠΑΡΑΓΩΓΗ ΟΜΙΛΙΑΣ	60
5.3.1.	Κωδικοποίηση κυματομορφής.....	61
5.3.2.	Ανάλυση-σύνθεση.....	61
5.3.3.	Σύνθεση με κανόνες.....	62
5.3.4.	Εμπορικές εφαρμογές παραγωγής ομιλίας.....	62
5.3.5.	Ανθρώπινοι παράγοντες.....	64
5.4.	ΟΛΟΚΛΗΡΩΜΕΝΑ ΔΙΑΛΟΓΙΚΑ ΣΥΣΤΗΜΑΤΑ ΜΕ ΟΜΙΛΙΑ.....	64
5.4.1.	Πολυτροπία.....	65
5.4.2.	Οι προθέσεις των εταιρειών.....	65
5.5.	ΧΡΗΣΤΕΣ ΜΕ ΕΙΔΙΚΕΣ ΑΝΑΓΚΕΣ.....	66
5.6.	ΠΡΩΤΟΚΟΛΛΑ ΚΑΙ ΠΡΟΤΥΠΑ ΕΠΙΚΟΙΝΩΝΙΑΣ	66
6.	ΤΕΧΝΟΛΟΓΙΕΣ ΕΛΕΓΧΟΥ ΜΕ ΕΓΚΕΦΑΛΙΚΑ ΚΥΜΑΤΑ (BRAIN CONTROL TECHNOLOGIES)	69
	ΑΝΑΦΟΡΕΣ.....	71
	ΑΝΑΦΟΡΕΣ ΣΕ ΙΣΤΟΣΕΛΙΔΕΣ	76

1. Εισαγωγή

Οι τεχνικές πρόσβασης σε υπολογιστικά περιβάλλοντα αποτελούν ένα ουσιαστικό μέρος των συστημάτων διαπροσωπικής επικοινωνίας που είναι βασισμένα σε υπολογιστές. Στον παρόντα τόμο παρουσιάζονται οι διαθέσιμες σήμερα τεχνικές πρόσβασης καθώς και οι αντίστοιχες συσκευές εισόδου. Συγκεκριμένα εξετάζονται διάφοροι τύποι πληκτρολογίων, συσκευές ένδειξης, διακόπτες και τεχνικές σάρωσης καθώς και η χρήση των τεχνολογιών ομιλίας. Τέλος παρουσιάζονται τεχνολογίες ελέγχου με εγκεφαλικά κύματα που παρουσιάστηκαν πρόσφατα. Η αρχική μορφή του τόμου αυτού δημιουργήθηκε στο πλαίσιο του ερευνητικού έργου ΑΙΝΕΙΑΣ: ΑΝΑΠΤΥΞΗ ΕΥΕΛΙΚΤΩΝ ΣΥΣΤΗΜΑΤΩΝ ΕΝΑΛΛΑΚΤΙΚΗΣ ΚΑΙ ΕΠΑΓΓΕΛΤΙΚΗΣ ΕΠΙΚΟΙΝΩΝΙΑΣ ΜΕΣΩ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΤΟΥ ΔΙΑΔΙΚΤΥΟΥ, το οποίο χρηματοδοτήθηκε από τη Γενική Γραμματεία Έρευνας και Τεχνολογίας (ΓΓΕΤ) στο πλαίσιο του Επιχειρησιακού Προγράμματος ΕΠΕΤ II με συγχρηματοδότηση από το Ευρωπαϊκό Κοινωνικό Ταμείο της Ευρωπαϊκής Ένωσης.

2. Πληκτρολόγια και Πλήκτρα

2.1. Εισαγωγή

Ο βασικός τρόπος εισαγωγής κειμένου εξακολουθεί να είναι ακόμα το πληκτρολόγιο (keyboard). Αυτή η συσκευή, αν και συχνά δέχεται κριτική, γνώρισε εντυπωσιακή επιτυχία. Εκατομμύρια άνθρωποι χρησιμοποιούν το πληκτρολόγιο με ταχύτητα πάνω από 15 πατήματα πλήκτρων ανά δευτερόλεπτο (περίπου 150 λέξεις ανά λεπτό), αν και ο ρυθμός των αρχάριων είναι 1 πάτημα το δευτερόλεπτο και ο ρυθμός για μέσους χρήστες είναι 5 πατήματα ανά δευτερόλεπτο (περίπου 50 λέξεις ανά λεπτό).

Τα σύγχρονα πληκτρολόγια επιτρέπουν γενικά ένα μόνο πάτημα πλήκτρου τη φορά, αν και το διπλό πάτημα πλήκτρων (SHIFT και γράμμα) χρησιμοποιείται για παραγωγή κεφαλαίων γραμμάτων και ειδικών λειτουργιών (CTRL και γράμμα).

Μπορούμε να πετύχουμε μεγαλύτερη ταχύτητα στην είσοδο δεδομένων χρησιμοποιώντας έγχορδα πληκτρολόγια [3] (chord keyboards, Σχήμα 1) τα οποία επιτρέπουν το ταυτόχρονο πάτημα μερικών πλήκτρων μαζί που αντιστοιχεί σε μερικούς χαρακτήρες ή και λέξεις.

Οι πρακτικογράφοι των δικαστηρίων χρησιμοποιούν συνήθως έγχορδα πληκτρολόγια για την εισαγωγή του πλήρους κειμένου των συζητήσεων, τα οποία φτάνουν σε έναν ρυθμό πάνω από 300 λέξεις ανά λεπτό. Για να μπορέσει κάποιος να συγκρατήσει τους πολύπλοκους τύπους πατημάτων των έγχορδών πληκτρολογίων απαιτείται εκμάθηση μηνών και συχνή χρήση του.



Σχήμα 1. Έγχορδο Πληκτρολόγιο

Το πληκτρολόγιο μορφής πιάνου [3] (piano keyboard, Σχήμα 2) είναι μια άλλη εντυπωσιακή συσκευή εισόδου δεδομένων, που μας δίνει τη δυνατότητα να πατάμε ταυτόχρονα μερικά δάχτυλα για την επιλογή κάποιων χαρακτήρων ή και λέξεων.



Σχήμα 2. Πληκτρολόγιο μορφής πιάνου

Το μέγεθος και η διάταξη των πλήκτρων πάνω στο πληκτρολόγιο επηρεάζουν πάρα πολύ τους χρήστες. Η ικανοποίηση του χρήστη από τη χρήση μιας μορφής πληκτρολογίου είναι ένας σημαντικός παράγοντας σχεδιασμού. Μεγάλα πληκτρολόγια με πολλά πλήκτρα δίνουν την εντύπωση του επαγγελματισμού και της πολυπλοκότητας, αλλά μπορούν να φοβίσουν έναν καινούριο χρήστη. Τα μικρά πληκτρολόγια δίνουν την αίσθηση σε μερικούς χρήστες ότι στερούνται ισχύς, αλλά το συμπαγές μέγεθός τους ελκύει άλλους.

2.2. Διατάξεις Πληκτρολογίων (Keyboard Layouts)

Η θέση των πλήκτρων των γραμμάτων και των αριθμών έχει γίνει θέμα έρευνας, θεωριών, διαμάχης, διεκδίκησης και πρωτότυπων εργασιών, μέχρις ότου εμφανίστηκαν τα πρώτα συμβατά πληκτρολόγια δακτυλογράφησης.

Αν και νωρίτερα είχαν ήδη κατασκευαστεί γραφομηχανές, η πρωτότυπη σχεδίαση το 1868 από τους Sholes, Gidden και Soule [50] ήταν η πρώτη που περιελάμβανε τα χαρακτηριστικά των σύγχρονων γραφομηχανών. Τα γράμματα αρχικά είχαν αλφαβητική διάταξη.

2.2.1. Η Κανονική Διάταξη (QWERTY)

Οι γρήγοροι δακτυλογράφοι αντιμετώπιζαν προβλήματα με τον αρχικό σχεδιασμό του πληκτρολογίου του Sholes, επειδή οι ράβδοι εκτύπωσης των διαδοχικών πλήκτρων ανακατευόταν μεταξύ τους (Σχήμα 3).



Σχήμα 3: Πληκτρολόγιο Sholes

Οι σύγχρονες QWERTY [50] διατάξεις (που ονομάζονται έτσι συμφωνά με τα 6 πάνω αριστερά πλήκτρα γραμμάτων) αυξάνουν το χώρο μεταξύ των κοινών ζευγαριών γραμμάτων, με αποτέλεσμα να ελαττώνουν τη συχνότητα μπλοκαρίσματος που προκαλείται από τη διαδοχική πίεση των ράβδων τυπώματος.

Η πρωτότυπη διάταξη QWERTY εμφανίστηκε το 1878 (Σχήμα 4). Υπήρξαν επίσης μερικές προσπάθειες βελτίωσης των διατάξεων QWERTY με την ανάπτυξη παρεμφερών διευθετήσεων.



Σχήμα 4. Η πρωτότυπη διάταξη QWERTY

2.2.2. Η Απλοποιημένη διάταξη DVORAK (Dvorak Simplified Keyboard)

Η πιο γνωστή από τις προσπάθειες βελτίωσης της διάταξης DVORAK είναι η απλοποιημένη διάταξη DVORAK (Σχήμα 5) [14] (γνωστή και ως DSK, Dvorak Simplified Keyboard). Ο August Dvorak για τη σχεδίαση αυτή έλαβε το 1936 πατέντα προτυποποίησης στις ΗΠΑ. Ο Dvorak σχεδίασε τη διάταξη αυτή χρησιμοποιώντας τα αποτελέσματα ερευνών που αφορούσαν στο χρόνο και στην κίνηση, καθώς και επιστημονικές μετρήσεις της αποτελεσματικότητας. Ο Dvorak σχεδίασε τη διάταξη για δακτυλογράφηση με δέκα δάκτυλα.

Οι βασικές αρχές της διάταξης DVORAK περιλαμβάνουν υποθέσεις όπως: α) είναι ευκολότερο και γρηγορότερο να μάθουμε να εκτελούμε απλές κινήσεις από ότι πολύπλοκες κινήσεις, β) οι ρυθμικές κινήσεις είναι λιγότερο κουραστικές από ότι οι ακανόνιστες. Με τη διάταξη DSK οι δακτυλογράφοι χρησιμοποιούν το δεξί τους χέρι περισσότερο από το αριστερό. Η τοποθέτηση των φωνηέντων και των συμφώνων, που συχνά βρίσκονται στις δύο αντίθετες άκρες του πληκτρολογίου, αυξάνει τη συχνότητα δακτυλογράφησης με τα δύο χέρια.



Σχήμα 5. Διάταξη DVORAK

Πολλά πειράματα, πολλές δοκιμές και αναλυτικές έρευνες έχουν συγκρίνει το DSK και το QWERTY. Παρακάτω παρουσιάζονται οι κυριότερες

Έρευνες του Strong [52]

Το 1956 ο Strong συνέκρινε δύο ομάδες δακτυλογράφων. Ο Strong εκπαίδευσε δακτυλογράφους που χρησιμοποιούσαν QWERTY επάνω σε διατάξεις DSK έως ότου αυτοί έφτασαν στο επίπεδο απόδοσης που είχαν προηγουμένως σε διατάξεις QWERTY. Αυτό διήρκεσε περίπου 28 μέρες. Στο δεύτερο μέρος του

πειράματος η ομάδα DSK έλαβε επιπλέον μαθήματα για να αυξήσει την ταχύτητα και την ορθότητα της δακτυλογράφησης σε DSK. Η άλλη ομάδα σύγκρισης QWERTY άρχισε το πείραμα στο δεύτερο μισό και έκανε επιπρόσθετα μαθήματα σε διάταξη QWERTY. Μετά το τέλος της εκπαίδευσης η ομάδα QWERTY εκτελούσε καλύτερα τις δακτυλογραφήσεις από την ομάδα DSK (Alden, Daniels και Kanarick 1972, Noyes 1983). Ο Strong κατέληξε στο συμπέρασμα πως δεν υπάρχει πλεονέκτημα στο DSK και πως η επιπλέον εκπαίδευση σε σύστημα QWERTY είναι περισσότερο λειτουργική (Yamada 1980) [58].

Προσομοίωση Kinkeard [29]

Στο πείραμά του ο Kinkeard συνέπτυξε τους χρόνους, που χρειάζονται τα δάχτυλα να τυπώσουν οποιαδήποτε ακολουθία δύο χαρακτήρων (λεγόμενα διγράμματα ή διγραφήματα) σε διάταξη QWERTY. Το δεύτερο μέρος της ανάλυσης ήταν ο υπολογισμός της συχνότητας εμφάνισης των διγραμμάτων αυτών στα Αγγλικά.

Ο Kinkeard έκανε την υπόθεση ότι ο χρόνος κίνησης ενός συγκεκριμένου δακτύλου (χρόνος πατήματος πλήκτρου) είναι ο ίδιος και για το DSK και για το QWERTY. Αυτή η υπόθεση έγινε με την εξής λογική: τα πλήκτρα και οι στήλες και των δύο πληκτρολογίων έχουν την ίδια διεύθυνση και έτσι απαιτούνται οι ίδιες κινήσεις των δακτύλων. Η μόνη διαφορά που υπήρχε μεταξύ των δύο διατάξεων ήταν στην τοποθεσία των ίδιων πλήκτρων και στη συχνότητα κίνησης του ίδιου δακτύλου. Ο Kinkeard χρησιμοποίησε το άθροισμα όλων των "συχνοτήτων διγραμμάτων επί το χρόνο πατήματος πλήκτρων" για να υπολογίσει την ταχύτητα δακτυλογράφησης για κάθε διάταξη.

Ως αποτέλεσμα αυτής της ανάλυσης, διαπίστωσε πως στην καλύτερη περίπτωση η χρήση DSK είναι 2,3% φορές γρηγορότερη από τη χρήση QWERTY.

Προσομοίωση Υπολογιστών (Computer Simulation)

Οι Norman και Fisher το 1982 [43] έκαναν μια ακόμα σύγκριση του DSK και του QWERTY χρησιμοποιώντας προσομοίωση της κίνησης του χεριού και του δακτύλου ενός έμπειρου δακτυλογράφου μέσω υπολογιστή. Στο τέλος υπολόγισαν πως η διάταξη DSK παρουσιάζει πλεονέκτημα της τάξης του 5,4% στην ταχύτητα δακτυλογράφησης από τη διάταξη QWERTY. Η εφαρμογή του μοντέλου παρουσίασε ρυθμό δακτυλογράφησης περίπου 58 λέξεις ανά λεπτό για τη διάταξη DSK και περίπου 56 λέξεις ανά λεπτό για τη διάταξη QWERTY

Συμπέρασμα

Πολλές έρευνες έχουν αποδείξει πως η διάταξη DSK είναι γρηγορότερη από τη QWERTY. Ωστόσο υπάρχουν ασυμφωνίες στο κατά πόσο οι διατάξεις DSK είναι καλύτερες των διατάξεων QWERTY. Προηγούμενες μετρήσεις έχουν διαπιστώσει πως το DSK είναι από 15% έως και 50 % γρηγορότερο από το QWERTY (Yamada 1980). Πιο πρόσφατες μετρήσεις δίνουν πιο μικρούς αριθμούς της τάξεως 2% μέχρι 17% (Norman and Fisher 1982) [43].

Συμφωνά με τα παραπάνω αποτελέσματα η χρήση της DSK δεν θα παρουσιάζει σημαντική βελτίωση στην πράξη. Για παράδειγμα, ο δακτυλογράφος με μέση ταχύτητα 50 λέξεις ανά λεπτό μετά από συνολική εκπαίδευση θα έχει ταχύτητα από 52.5 μέχρι 55 λέξεις ανά λεπτό. Δηλαδή, αν υποθέσουμε ότι μια σελίδα περιλαμβάνει περίπου 800 λέξεις, ο εκπαιδευμένος χρήστης δακτυλογραφώντας 8 ώρες την ημέρα θα αυξήσει την παραγωγή του από 30 σελίδες ανά ημέρα σε 31.5-33 σελίδες ανά ημέρα. Επίσης, ο χρήστης που είναι εκπαιδευμένος σε QWERTY μπορεί να μεταφέρει με ευκολία την εμπειρία του σε άλλες διατάξεις, ενώ ο χρήστης που είναι εκπαιδευμένος σε DSK δεν μπορεί να το κάνει αυτό.

2.2.3. Αλφαβητικές Διατάξεις (Alphabetical Keyboards)

Μια άλλη μέθοδος σχεδίασης πληκτρολογίων είναι η τοποθέτηση των γραμμάτων των πλήκτρων σε αλφαβητική διάταξη. Τέτοιες διατάξεις εμφανίζονται σε μερικά παιχνίδια παιδιών, στα χρηματιστηριακά τερματικά, σε κάποιες φορητές συσκευές δεδομένων και μερικές φορές εμφανίζονται σαν προεπιλεγμένα πληκτρολόγια πάνω στην οθόνη για κάποιες εφαρμογές που απαιτούν οθόνες αφής.

Σύγκριση QWERTY και Αλφαβητικών Πληκτρολογίων

Ο Hirsch το 1970 [26] δοκίμασε μια ομάδα μη επαγγελματιών δακτυλογράφων σε διάταξη QWERTY και μια άλλη ομάδα σε αλφαβητικές διατάξεις. Μετά από επτά ώρες πρακτικής η ομάδα QWERTY βελτίωσε την ταχύτητα δακτυλογράφησης από 1,47 σε 1,99 πατήματα πλήκτρων ανά δευτερόλεπτο. Ωστόσο, η ομάδα που δούλεψε σε αλφαβητική διάταξη δεν έφτασε ούτε τον ρυθμό που είχε η QWERTY πριν το πείραμα (1,47 πατήματα πλήκτρων ανά δευτερόλεπτο σε QWERTY συγκρίθηκαν με 1,11 πατήματα σε αλφαβητικές διατάξεις).

Οι Norman και Fisher το 1982 [43] δοκίμασαν μη δακτυλογράφους σε τέσσερα διαφορετικά πληκτρολόγια: QWERTY, αλφαβητική-οριζόντια διάταξη (τα γράμματα από Α έως το Ζ είναι διευθετημένα από αριστερά προς δεξιά αρχίζοντας από το πάνω αριστερό πλήκτρο), αλφαβητική-διαγώνια (με γράμματα διευθετημένα διαγώνια αρχίζοντας από το πάνω αριστερό πλήκτρο) και τυχαίο πληκτρολόγιο (τα γράμματα είναι διευθετημένα τυχαία). Η δακτυλογράφηση στο QWERTY ήταν περίπου 65% γρηγορότερη από τα άλλα πληκτρολόγια. Η στατιστική έρευνα αποκάλυψε πως τα πρώτα τρία πληκτρολόγια ήταν σημαντικά καλύτερα από την τυχαία διάταξη και το QWERTY ήταν καλύτερο από τις δύο αλφαβητικές διατάξεις (οι οποίες μεταξύ τους δεν είχαν σημαντική διαφορά).

Συμπέρασμα

Τα πληκτρολόγια με αλφαβητικές διατάξεις (που προφανώς αδιαφορούν για ειδικές διατάξεις) δεν παρουσιάζουν πλεονεκτήματα έναντι της διάταξης QWERTY για μη έμπειρους χρήστες που χρησιμοποιούν πληκτρολόγια μικρού μεγέθους. Η απόδοση της διάταξης QWERTY μπορεί να είναι καλύτερη από την αλφαβητική διάταξη, επειδή η διευθέτηση QWERTY δεν είναι τυχαία, με αποτέλεσμα να ελαττώνει της δυσκολίες αναζήτησης. Μια άλλη πιθανή εξήγηση είναι το ότι οι περισσότεροι άνθρωποι, ακόμα και χωρίς εμπειρία στη δακτυλογράφηση, έχουν κάποια πείρα χρήσης της διάταξης QWERTY. Γενικά, όλα τα γεγονότα αποδεικνύουν ότι στις περισσότερες περιπτώσεις οι σχεδιαστές πρέπει να προτιμούν τη διάταξη QWERTY και όχι τις αλφαβητικές διατάξεις.

2.2.4. Άλλες διατάξεις πληκτρολογίων

Μερικοί ερευνητές έχουν κατασκευάσει μη "κανονικές" διατάξεις με μη "κανονικές" διευθετήσεις. Οι Getschow, Rosen και Goodenough-Trepagnier το 1986 [18] σε μια έρευνά τους χρησιμοποίησαν μια συνάρτηση αναζήτησης της τεχνητής νοημοσύνης (τον αλγόριθμο "greedy") για την ανάπτυξη μιας διάταξης, η οποία ελαχιστοποιεί τη μέση απόσταση μεταξύ των αγγλικών διγραμμάτων (με πλήκτρα τοποθετημένα σε διάταξη πίνακα 5x5).

Θεωρητικά, αυτή έπρεπε να είναι η καλύτερη διάταξη για έμπειρο δακτυλογράφο που χρησιμοποιεί στυλό ή δακτυλογραφεί με ένα δάχτυλο, αλλά ο Getschow δεν έκανε πειράματα ελέγχου με χρήστες αναφορικά με τη διάταξη αυτή.

Ο Lewis το 1992 [31] σε ένα πείραμά του χρησιμοποίησε ένα πρόγραμμα ανάλυσης για τη σχεδίαση μιας διάταξης ελάχιστης απόστασης όμοιας με αυτή που είχε κατασκευάσει ο Getschow.

Χρησιμοποιώντας τη συνάρτηση κόστους, βασισμένη στη συχνότητα εμφάνισης των αγγλικών διγραμμάτων, την οποία είχαν υπολογίσει οι Fitts' Law και Lewis, κατασκεύασε τη διάταξη ελάχιστης απόστασης που είναι 27% ταχύτερη από το QWERTY. Χρησιμοποιώντας την ίδια υπόθεση (υψηλή εξάσκηση δακτυλογράφησης) ο Lewis το 1992 [31] υπολόγισε πως η διάταξη που ανέπτυξε ο Getschow θα ήταν περίπου 31% ταχύτερη από τη διάταξη QWERTY.

Ένα μειονέκτημα της διάταξης αυτής είναι το ότι η συνάρτηση κόστους του Lewis προβλέπει μόνο την απόδοση ειδικού και πολύ έμπειρου δακτυλογράφου, αλλά δεν είναι σε θέση να αξιολογήσει τη δακτυλογράφηση του μη ειδικού.

Οι Matias, MacKenzie και Buxton το 1993 [39] κατασκεύασαν το πληκτρολόγιο ενός χεριού ονομαζόμενο μισό-QWERTY, (Σχήμα 6) που σχεδιάστηκε για να βελτιώσει το πλεονέκτημα εμπειρίας που είχαν οι χρήστες με τη διάταξη QWERTY.

	!	@	#	\$	%					
	1	0	2	9	3	8	4	7	5	6
Tab	Q	W	E	R	T					
delete	P	O	I	U	Y					
	A	S	D	F	G					
	:	L	K	J	H					
Shift	Z	X	C	V	B					
return	/	.	,	M	N					
						Flip				

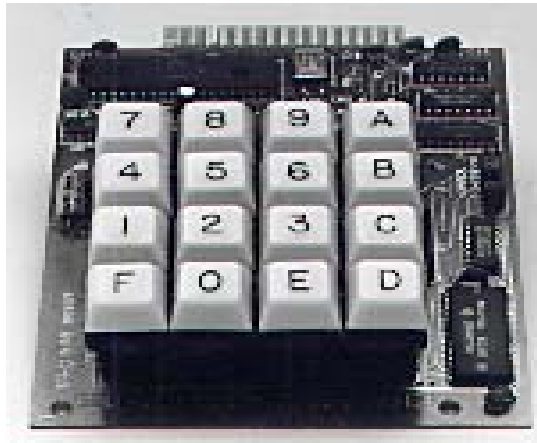
Σχήμα 6

Η διάταξη μισό-QWERTY έχει δύο λειτουργικούς χαρακτήρες σε κάθε πλήκτρο που βρίσκονται στο αριστερό μισό της διάταξης QWERTY, με κατοπτρικό είδωλο του δεξιού μισού τοποθετημένο σε αριστερό μέρος. Για παράδειγμα το πλήκτρο Q είναι επίσης και P, το πλήκτρο T είναι επίσης και πλήκτρο Y. Ο χρήστης πατάει το πλήκτρο κανονικά, για να πάρει το κανονικό γράμμα που είναι συσχετισμένο με το πλήκτρο, και πιέζοντας και κρατώντας το κενό παίρνει τον εναλλακτικό χαρακτήρα.

Μετά το τέλος του ελέγχου του πειράματος (περίπου σε οκτώ ώρες) η ταχύτητα των δακτυλογράφων είχε φτάσει το 50% της κανονικής τους ταχύτητας σε δακτυλογράφηση με δύο χέρια.

2.2.5. Πληκτρολόγια Εισαγωγής Δεδομένων (Data-Entry Keyboards)

Εκτός από το αλφαριθμητικό τμήμα, τα περισσότερα πληκτρολόγια έχουν ξεχωριστά αριθμητικά πλήκτρα για εισαγωγή δεδομένων. Επίσης αυξάνεται συνεχώς η χρήση τηλεφωνικών πλήκτρων στα τερματικά (Σχήμα 7) για την εισαγωγή δεδομένων στον υπολογιστή (φαινόμενο που αρχικά αναφέρθηκε από τους Bayes και Thompson το 1983 [2]).



Σχήμα 7

Σε αυτή την παράγραφο παρουσιάζουμε τη σχεδίαση των πληκτρολογίων για τηλέφωνα και για άλλες εφαρμογές.

Διάταξη των Αριθμών και των Γραμμάτων

Ο Lutz και ο Charanis το 1955 [35] πραγματοποίησαν μια μελέτη με έξι διαμορφώσεις πλήκτρων, για να προσδιορίσουν πού και σε ποια θέση αναμένουν οι άνθρωποι να εμφανιστεί το κάθε ψηφίο σε ένα πληκτρολόγιο δέκα πλήκτρων που είναι για χρήση από απόσταση. Η διευθέτηση των πλήκτρων ήταν σε δύο οριζόντιες γραμμές πέντε πλήκτρων, δύο κατακόρυφων γραμμών των πέντε πλήκτρων ή σε τρεις γραμμές με τρία πλήκτρα και ένα πλήκτρο τοποθετημένο στην κορυφή, στη βάση, αριστερά ή δεξιά από το μπλοκ των εννέα πλήκτρων.

Γενικά, οι άνθρωποι οργανώνουν τους αριθμούς και τα γράμματα στη διάταξη με την οποία διαβάζουν κείμενα (που είναι από αριστερά προς δεξιά και από πάνω προς τα κάτω). Όταν οι αριθμοί είναι τοποθετημένοι ήδη στο πληκτρολόγιο: α) οι άνθρωποι τοποθετούν τα γράμματα στα πλήκτρα από αριστερά προς δεξιά και από πάνω προς τα κάτω, όταν οι αριθμοί έχουν αυτή τη διευθέτηση και β) αν η διάταξη των αριθμών δεν είναι από αριστερά προς δεξιά και από πάνω προς κάτω περίπου, οι μισοί άνθρωποι τοποθετούν τα γράμματα σύμφωνα με τη διάταξη των αριθμών και οι άλλοι μισοί προτιμούν να τοποθετούν τα γράμματα από αριστερά προς δεξιά και από πάνω προς τα κάτω.

Οι Goodman, Dickinson και Francas το 1983 [20] σε μια σχετική έρευνά τους χρησιμοποίησαν μια μέθοδο προσομοίωσης, για να καθορίσουν την καλύτερη διάταξη των πλήκτρων για τα πληκτρολόγια του Telidon (Canadas Videotex System). Δοκίμασαν την οριζόντια (από 1 έως 0 σε μία μόνο γραμμή) και την τηλεφωνική διάταξη. Η ταχύτητα και η ακρίβεια της τηλεφωνικής διάταξης ήταν λίγο καλύτερη σε εργασίες χρόνο-αντίδρασης. Ωστόσο, οι διαφορές σε πιο ρεαλιστικές εφαρμογές δεν ήταν σημαντικές από στατιστικής πλευράς.

2.2.6. Συμπέρασμα

Η σχεδίαση και η χρήση των πληκτρολογίων εισόδου δεδομένων εξαρτάται σε μεγάλο βαθμό από τις απαιτήσεις της εργασίας. Για αριθμητική ή μικτή είσοδο η τηλεφωνική διάταξη είναι λίγο καλύτερη από διάταξη αριθμομηχανής, ειδικά για ανθρώπους που δεν είναι εξοικειωμένοι με αριθμομηχανές ή με μηχανές πρόσθεσης.

Οι έμπειροι χρήστες που χρειάζονται να εκτελέσουν εκτεταμένες πράξεις με αριθμούς προτιμούν να έχουν ένα ξεχωριστό αριθμητικό πληκτρολόγιο πάνω ή δίπλα στο πλήρες πληκτρολόγιο (που περιλαμβάνει και το αλφάβητο).

2.3. Φυσικά Χαρακτηριστικά των Πλήκτρων και των Πληκτρολογίων

2.3.1. Ύψος και Κλίση των Πληκτρολογίων

Ο Galitz το 1965 [17] δοκίμασε μια σειρά από πληκτρολόγια με κλίσεις 9, 21 και 33 μοιρών. Αν και δεν υπήρχε διαφορά στην απόδοση που οφείλεται στην κλίση, οι δακτυλογράφοι προτίμησαν τη γωνία 21 μοιρών. Ο Galitz πρότεινε τα πληκτρολόγια να είναι προσαρμόσιμα μεταξύ των 10 και 35 μοιρών έτσι ώστε να ικανοποιούν τις ατομικές προτιμήσεις.

Στη μελέτη των Suther και McTyre το 1982, [53] οι έμπειροι δακτυλογράφοι χρησιμοποίησαν πληκτρολόγια λεπτού-προφίλ (30mm) με κλίση 5, 10 και 25 μοιρών και πληκτρολόγια χοντρού-προφίλ με κλίση 15 μοιρών. Το πείραμα έδειξε πως δεν υπάρχουν διαφορές στην απόδοση πληκτρολόγησης και για τις τέσσερις κλίσεις. Κανένας από τους δακτυλογράφους δεν προτίμησε το πληκτρολόγιο με κλίση 5 μοιρών. Ένας προτίμησε την κλίση των 25 μοιρών και οι υπόλοιποι εκτίμησαν το πληκτρολόγιο με κλίση 15 μοιρών. Αυτή η μελέτη επίσης διαπίστωσε κάποιες ιδιότητες σχετικές με το ανάστημα και το μήκος των χεριών. Οι ψηλοί άνθρωποι και εκείνοι με μακριά χέρια τείνουν να επιλέγουν τα πληκτρολόγια με μικρή κλίση, ενώ οι κοντοί άνθρωποι και οι άνθρωποι με κοντά χέρια προτιμούν πληκτρολόγια με μεγάλες κλίσεις. Οι Suther και McTyre πρότειναν τα πληκτρολόγια να είναι προσαρμόσιμα μεταξύ 10 και 25 μοιρών.

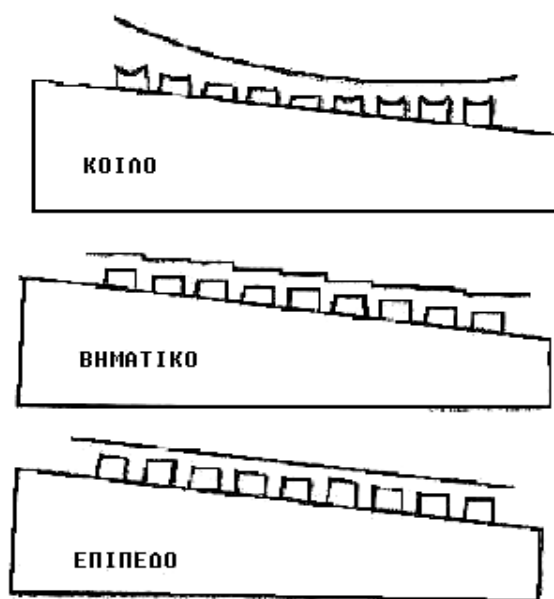
Οι Najjar, Stanton και Bowen το 1988 [41] εξέτασαν την απόδοση και τις ιδιότητες των επαγγελματιών δακτυλογράφων χρησιμοποιώντας τρία πληκτρολόγια με ύψος 74 cm, 99 cm και 125 cm και τρεις γωνίες (κλίσεις) 0, 15 μοίρες αρνητικό και 15 μοίρες θετικό. Στο χαμηλότερο ύψος η απόδοση στις 0 και -15 μοίρες ήταν σημαντικά καλλίτερη από αυτή σε πληκτρολόγιο με κλίση 15 μοιρών. Οι συμμετέχοντες προτίμησαν τη δακτυλογράφηση σε πληκτρολόγια με 0 και με θετική κλίση 15 μοιρών στην περίπτωση πληκτρολογίων με μέτριο και με ψηλό ύψος. Χρησιμοποιώντας πληκτρολόγια με θετικές κλίσεις με χαμηλό ύψος, οι χρήστες παρουσίασαν κόπωση στους καρπούς των χεριών.

Συμπέρασμα

Από τα αποτελέσματα των ερευνών φαίνεται πως οι μεγάλες ποικιλίες, (σε ύψος και κλίση) πληκτρολογίων δεν επηρεάζουν την απόδοση δακτυλογράφησης. Οι χρήστες φαίνεται να προτιμούν κάποια κλίση του πληκτρολογίου. Η γωνία κλίσης πρέπει να προσαρμόζεται τουλάχιστον σε γωνία 15 μοιρών και μπορεί να αποκτήσει και πιο απότομες κλίσεις για την ικανοποίηση ειδικότερων αναγκών. Αυτό συμφωνεί και με τις τυποποιημένες υποδείξεις του ANSI/HFS 100-1988 που προτείνουν κλίσεις μεταξύ 0 και 25 μοιρών.

2.3.2. Προφίλ Πληκτρολογίων (Keyboard Profile)

Οι σχετικές κλίσεις και η θέση των γραμμών των πλήκτρων στο πληκτρολόγιο δημιουργούν το προφίλ του πληκτρολογίου. Τα περισσότερα πληκτρολόγια διακρίνονται σε τρεις κατηγορίες: επίπεδα (οι κορυφές των πλήκτρων είναι παράλληλες με την κλίση του πληκτρολογίου), κοίλα και βηματικά (Σχήμα 8).



Σχήμα 8. Προφίλ πληκτρολογίων.

Οι Pacì και Gabbrielli [46] σε ένα πείραμά τους το 1984, αποτίμησαν την απόδοση τριών δακτυλογράφων που χρησιμοποιούσαν βηματικά και κοίλα πληκτρολόγια. Η γωνία με την οποία τα δάκτυλα του χρήστη ακουμπούσαν τα πλήκτρα ρυθμιζόταν από 2 έως 13 μοίρες για βηματικά πληκτρολόγια και από 8 έως 11 μοίρες για κοίλα. Η απόδοση, σύμφωνα με την αναφορά, ήταν καλύτερη στα κοίλα πληκτρολόγια και οι δακτυλογράφοι εξέφρασαν την προτίμησή τους για αυτά. Οι Pacì και Gabbrielli πρότειναν το κοίλο προφίλ να χρησιμοποιείται για τα αλφαριθμητικά πλήκτρα, το βηματικό προφίλ για τα αριθμητικά και το επίπεδο προφίλ για πλήκτρα λειτουργιών (function keys).

Ο Magyar το 1985 [38] έκανε σύγκριση των αποδόσεων και των ιδιοτήτων δώδεκα δακτυλογράφων χρησιμοποιώντας επίπεδα, βηματικά και κοίλα προφίλ. Κάθε χρήστης εκτέλεσε είκοσι φορές τη δοκιμή δακτυλογράφησης κάθε μέρα με κάθε προφίλ πληκτρολογίου. Οι δακτυλογράφοι λάμβαναν ανάδραση ρυθμού παραγωγής και λαθών μετά από κάθε δοκιμή και συμπλήρωναν το καθημερινό ερωτηματολόγιο μετά από τη χρήση του κάθε προφίλ πληκτρολογίου.

Η απόδοση της παραγωγής εξόδου για επίπεδα πληκτρολόγια ήταν σημαντικά χαμηλότερη είτε επρόκειτο για βηματικά, είτε για κοίλα πληκτρολόγια, τα οποία μεταξύ τους δεν είχαν καμία σημαντική διαφορά. Αν και ο ρυθμός των εντοπισμένων λαθών ήταν συγκρίσιμος μεταξύ και των τριών προφίλ, ο ρυθμός μη εντοπισμένων λαθών στο επίπεδο προφίλ ήταν σημαντικά μεγαλύτερος. Οι δακτυλογράφοι παραπονέθηκαν ότι το μέγεθος και η τοποθεσία των πλήκτρων "backspace", "enter" και του "shift" μπορεί να δυσκολέψει τη δακτυλογράφηση, ανεξάρτητα από τα διαφορετικά προφίλ των πληκτρολογίων. Παρόλα αυτά, σαν συμπέρασμα, το βηματικό και το κοίλο προφίλ είναι καλύτερα από το επίπεδο προφίλ.

Συμπέρασμα

Αν και οι διαφορές μεταξύ των επίπεδων, των βηματικών και των προφίλ πιάτου πληκτρολογίων μπορεί να είναι αρκετά μικρές, έχει γενικά συμφωνηθεί πως τα βηματικά και τα κοίλα πληκτρολόγια είναι περισσότερο αποδεκτά.

2.3.3. Μορφή και Μέγεθος των Πλήκτρων

Με την εξάπλωση των φορητών συσκευών, οι σχεδιαστές των πληκτρολογίων έχουν την τάση να ελαττώνουν το μέγεθος των πληκτρολογίων μικραίνοντας το μέγεθος των πλήκτρων. Παράλληλα με την απαίτηση αυτή, οι σχεδιαστές πρέπει να έχουν υπ' όψη τους την επίδραση της ελάττωσης του μεγέθους των πλήκτρων στην απόδοση δακτυλογράφησης. Και αυτό γιατί οι ακμές των πλήκτρων είναι τα σημεία άμεσης επαφής μεταξύ του χρήστη και του πληκτρολογίου. Το μέγεθος και η μορφή των πλήκτρων μπορεί να έχουν μεγάλη σχέση με την απόδοση της δακτυλογράφησης.

Η πρόταση του Clare

Ο Clare [7] πρότεινε τέσσερις απαιτήσεις για το σχεδιασμό της μορφής των πλήκτρων:

1. Ο χρήστης πρέπει να έχει τη δυνατότητα να βλέπει την ετικέτα των πλήκτρων.
2. Το δάκτυλο πρέπει να μπορεί να τοποθετηθεί στο πλήκτρο χωρίς να χτυπά άλλο πλήκτρο ή δάκτυλο.
3. Η κατανομή της πίεσης του δακτύλου πρέπει να καθορίζεται από την τοποθεσία του δακτύλου πάνω στο πλήκτρο.
4. Η ισχύς της πίεσης των πλήκτρων πρέπει να καταναμηθεί με τον ίδιο τρόπο σε όλα τα δάκτυλα.

Ο Clare πρότεινε οι ακμές των πλήκτρων να είναι σε μορφή τετραγώνου με μήκος ακμής 12.7 mm και τα κέντρα των πλήκτρων να έχουν απόσταση 19 mm μεταξύ τους.

Μέγεθος Πλήκτρων της Αριθμομηχανής

Ο Deininger το 1960 [13] έκανε ένα πείραμα μελετώντας το μέγεθος και τη μορφή ενός αριθμητικού πληκτρολογίου με δέκα πλήκτρα. Η ταχύτητα και η ακρίβεια της δακτυλογράφησης βελτιώθηκε αυξάνοντας το μέγεθος των πλήκτρων από 9.5 mm σε 12.7 mm.

Για να προτυποποιήσουν την ιδιότητα αυτή, οι Logicchio και Lewis [34] το 1991 χρησιμοποίησαν 15 άτομα και τρεις εμπορικές αριθμομηχανές με διαφορετικού μεγέθους πλήκτρα και διαφορετικές αποστάσεις μεταξύ των πλήκτρων. Δεν υπήρχε καμία σημαντική διαφορά στην ακρίβεια, αλλά η ταχύτητα βελτιώθηκε με αύξηση του μεγέθους των πλήκτρων από τετράγωνα 10mm σε ορθογώνια πλήκτρα 14x10 mm.

Αλφαριθμητικά Πληκτρολόγια (Alphanumeric Keyboards)

Το περιοδικό PC Magazine (Rosch) [48] ανέφερε τα αποτελέσματα των πειραμάτων δακτυλογράφησης, με σκοπό την αποτίμηση διαφορετικών πληκτρολογίων. Η απόδοση της δακτυλογράφησης ήταν πολύ μικρή με πληκτρολόγια που είχαν μικρά πλήκτρα. Οι Logicchio και ο Kennedy [33] το 1987 ερεύνησαν το αποτέλεσμα μείωσης των κάθετων αποστάσεων μεταξύ των πλήκτρων από 19 σε 15 mm. Βρήκαν πως δεν υπήρχε διαφορά στο ρυθμό δακτυλογράφησης ή στο ρυθμό λαθών μετά από δύομισι ώρες πρακτικής, αλλά υπήρχε σταθερή και δυνατή προτίμηση από τους χρήστες της τυπικής απόστασης πλήκτρων 19 mm. Οι Wiklund, Dumas και Hoffman [57] το 1987 πραγματοποίησαν ένα πείραμα με τέσσερα διαθέσιμα στο εμπόριο πληκτρολόγια με τη χρήση ενός ή δύο χεριών.

Το μέγεθος των πλήκτρων ρυθμίστηκε από 19x19 mm σε 13x12 mm. Αν και ο Wiklund δεν παρουσίασε στατιστικές αναλύσεις από την τελική αναφορά, δεν είναι δύσκολο να καταλάβει κανείς ότι ο ρυθμός

δακτυλογράφησης είναι ταχύτερος για τα μεγάλα πλήκτρα ανεξάρτητα από το αν ο χρήστης χρησιμοποιεί το ένα ή και τα δύο του χέρια.

Μετά από μια σειρά πειράματα μπορούμε με αρκετή σιγουριά να πούμε πως τα τετράγωνα πλήκτρα μπορεί να είναι τα καλύτερα, επειδή παρέχουν μεγαλύτερη επιφάνεια με την ίδια απόσταση μεταξύ των κέντρων των πλήκτρων.

Συμπέρασμα

Η απόσταση μεταξύ των κέντρων δύο γειτονικών πλήκτρων είναι γενικά 19 mm. Η χρήση μικρότερης απόστασης μεταξύ των κέντρων έχει σαν αποτέλεσμα τη μείωση της ταχύτητας της δακτυλογράφησης. Οι σχεδιαστές συνήθως τείνουν να ελαττώνουν το μέγεθος του πληκτρολογίου τους παρά να το μεγαλώνουν (ειδικά στη σχεδίαση των φορητών συσκευών). Τελικά οι σχεδιαστές πρέπει να προσπαθούν να παρουσιάζουν πλήκτρα πλήρους μεγέθους (με απόσταση 19 mm μεταξύ τους) για την κύρια περιοχή δακτυλογράφησης (ειδικά για την αλφαριθμητική περιοχή και, αν είναι δυνατόν, και για το αριθμητικό). Τα αποτελέσματα αυτά ταιριάζουν με τις τυποποιήσεις που προτείνει η ANSI/HFS 100-1988 (Human Factor Society) σύμφωνα με τις οποίες η οριζόντια απόσταση μεταξύ των γειτονικών πλήκτρων είναι μεταξύ 18 και 19 mm και η κάθετη απόσταση μεταξύ 18 και 21 mm.

Η μορφή των πλήκτρων πρέπει να ταιριάζει κατάλληλα με την άκρη του δακτύλου, ώστε να διευκολύνεται η τοποθέτηση του δακτύλου. Η τυποποίηση ANSI/HFS 100-1988 δεν προτείνει καμία τυποποίηση σε σχέση με τη μορφή του κοιλώματος της επιφάνειας και αποδέχεται οποιαδήποτε μορφή επιφάνειας πλήκτρων (τετράγωνο, κύκλο ή ορθογώνιο) που συμμορφώνεται με τις αποστάσεις που προτάθηκαν.

2.3.4. Χρώμα και Ετικέτες (Color and Labeling)

Οι υποδείξεις για τα χρώματα και για τις ετικέτες των πλήκτρων και των πληκτρολογίων κυρίως βασίζονται στην ορατότητα και στους σκοπούς κωδικοποίησης, παρά σε πειραματικές έρευνες. Ούτε τα πλήκτρα αλλά ούτε και τα πληκτρολόγια, δεν πρέπει να λάμπουν για να μη δημιουργούν πηγή αντανάκλασης του φωτός. Τα ονόματα των πλήκτρων πρέπει να είναι ευανάγνωστα και κατανοητά στους χρήστες. Η ομαδοποίηση και ο χρωματισμός πλήκτρων με παρόμοιες λειτουργίες εξυπηρετεί την εικονική αναζήτηση.

2.3.5. Πλήκτρα Ειδικού Σκοπού (Special Purpose Keys)

Τα πλήκτρα ειδικού σκοπού είναι εκείνα τα πλήκτρα τα οποία δεν είναι μέρος της τυπικής αλφαριθμητικής περιοχής ή του αριθμητικού πληκτρολογίου. Τέτοια πλήκτρα είναι το "Backspace", το "Enter", τα πλήκτρα ελέγχου του δρομέα και τα πλήκτρα λειτουργιών (Function keys).

Στα μοντέρνα πληκτρολόγια το πλήκτρο "Enter" υιοθετεί μια πληθώρα από σχήματα, συμπεριλαμβανομένων του κάθετου πλήκτρου "Enter" (μακρύ και λεπτό), οριζόντιου πλήκτρου "Enter" (πλατύ και λεπτό) και του πλήκτρου "Enter" της μορφής ανάποδου-L. Μετά από μια σειρά μελέτες οι επιστήμονες έχουν καταλήξει στο συμπέρασμα πως το πλήκτρο μορφής ανάποδο-L προτιμάται πιο πολύ από τους χρήστες παρόλο που δεν υπάρχει σημαντική διαφορά στην ταχύτητα.

Το πλήκτρο "Backspace" κατέχει συνήθως το χώρο ενός ή δύο κανονικών πλήκτρων και συνήθως τοποθετείται στην πάνω δεξιά γωνία της τυποποιημένης περιοχής πληκτρολόγησης. Ο Magyar [37] σε μελέτες του προσπάθησε να υπολογίσει την απόδοση των χρηστών σε δύο τύπους πλήκτρων "Backspace" και ανακάλυψε πως οι χρήστες προτιμούν και παρουσιάζουν καλύτερη απόδοση με το διπλό πλήκτρο του "Backspace".

Ο ρυθμός λαθών για το πλήκτρο "Enter" οριζόντιας μορφής και μορφής ανάποδο-L είναι σχεδόν ίσος, ενώ η αναλογία του ρυθμού λαθών του μονού πλήκτρου "Backspace" σε σύγκριση με το διπλό πλήκτρο "Backspace" είναι 9 για απλό κείμενο (δακτυλογράφηση με πληκτρολόγιο QWERTY) και 24 για πολύπλοκα κείμενα.

2.3.6. Διευθέτηση των Πλήκτρων Ελέγχου του Δρομέα

Τα πλήκτρα ελέγχου του δρομέα είναι τα πλήκτρα-βέλη που χρησιμοποιούνται για τη μετακίνηση του δρομέα πάνω στην οθόνη. Οι διαφορετικές διατάξεις για αυτά τα τέσσερα πλήκτρα εμφανίζονται σε διάφορες σχεδιάσεις πληκτρολογίων. Οι περισσότερο κοινές διευθετήσεις είναι το κουτί, ο σταυρός και το ανάποδο-T. Σε διευθέτηση κουτιού τα πλήκτρα "αριστερά" και "δεξιά" βρίσκονται ακριβώς κάτω από τα πλήκτρα "πάνω" και "κάτω" σε μορφή πίνακα 2x2. Έτσι η διευθέτηση κουτιού έχει κατάλληλη ικανότητα ελέγχου-απόκρισης για τα πλήκτρα "αριστερά", "δεξιά" και όχι για τα πλήκτρα "πάνω", "κάτω".

Η διευθέτηση σταυρού έχει τα τέσσερα πλήκτρα τοποθετημένα γύρω από ένα κενό. Αυτή η διευθέτηση έχει σωστή ικανότητα ελέγχου-απόκρισης για όλα τα πλήκτρα, αλλά είναι κάπως αδέξια για τον χρήστη. Η διευθέτηση ανάποδο-T είναι κάτι παρόμοιο με τη διευθέτηση σταυρού, αλλά το κάτω βελάκι τοποθετείται στον ελεύθερο χώρο της διευθέτησης σταυρού. Αυτό έχει σαν αποτέλεσμα στο σχήμα αντίστροφου-T τα βελάκια "αριστερά" και "δεξιά" να τοποθετούνται αριστερά και δεξιά του πλήκτρου "κάτω" αντίστοιχα, και το προς τα "πάνω" βελάκι να βρίσκεται ακριβώς πάνω από το πλήκτρο του βέλους "κάτω".

Ο Emmons το 1984 [15] σε μια μελέτη του παρουσίασε την ανωτερότητα της διάταξης σταυρού έναντι της διάταξης κουτιού ειδικά για μη έμπειρους χρήστες. Στις μετέπειτα έρευνες ο Emmons διαπίστωσε πως δεν υπάρχει διαφορά στην απόδοση μεταξύ των διευθετήσεων σταυρού και ανάποδο-T, αλλά ανέφερε πως οι 13 από τους 18 συμμετέχοντες προτίμησαν το ανάποδο-T. Παρόμοια αποτελέσματα ανέφερε και ο Magyar σε μια ερευνά του το 1986 [37].

2.4. Εναλλακτικά Πληκτρολόγια (Alternative Keyboard Designs)

Μεγάλος αριθμός ερευνητών έχουν μελετήσει την αναδόμηση των πληκτρολογίων και την τροποποίηση εννοιών της λειτουργίας του πληκτρολογίου. Οι δύο κυριότερες περιοχές έρευνας στο σχεδιασμό εναλλακτικών πληκτρολογίων είναι τα σπαστά πληκτρολόγια και τα εναρμονίζοντα πληκτρολόγια.

2.4.1. Διαχωρισμένα Πληκτρολόγια ή Σπαστά Πληκτρολόγια (Split Keyboards)

Οι μελέτες για το πως οι άνθρωποι κρατάνε τα χέρια και το βραχίονά τους σε σχέση με τη θέση τους αναφορικά με συνηθισμένους δακτυλογράφους έχουν καθοδηγήσει τους ερευνητές να υποθέσουν πως τα συμβατικά πληκτρολόγια μπορεί να προκαλούν πόνο και κούραση στους χρήστες.

Σύμφωνα με αυτές τις παρατηρήσεις μερικοί κατασκευαστές και εφευρέτες έχουν αναπτύξει ή έχουν συνεισφέρει στην ανάπτυξη διαχωρισμένων πληκτρολογίων. Αυτά τα πληκτρολόγια έχουν τουλάχιστον δύο τμήματα, τα οποία παρουσιάζουν μεταξύ τους μια γωνία. Με αυτόν τον τρόπο δημιουργούνται μη-παράλληλες στήλες πλήκτρων. Μερικά διαχωρισμένα πληκτρολόγια έχουν καθορισμένη γωνία μεταξύ των τμημάτων του πληκτρολογίου, ενώ μερικά άλλα επιτρέπουν τη ρύθμιση της γωνίας (Σχήμα 9).



Σχήμα 9: Διαχωρισμένο πληκτρολόγιο.

Το Πληκτρολόγιο K

Ο Kroemer [30] το 1972 κατασκεύασε ένα διαχωρισμένο πληκτρολόγιο και το ονόμασε Πληκτρολόγιο K προς τιμήν του παλιού υπερασπιστή των διαχωρισμένων πληκτρολογίων, του Klockenberg. Τα Πληκτρολόγια K έχουν 30 πλήκτρα, τα οποία είναι τοποθετημένα σε ευθείες στήλες, σε καμπυλωμένες γραμμές και έχουν σα στόχο να συνδυάζουν διαφορετικά μήκη δαχτύλων. Οι ράβδοι κενού (space bars), ένας για κάθε δάχτυλο, κυρτώνονται για να εφάπτονται στον αντίχειρα και το πληκτρολόγιο παρέχει πληθωρική υποστήριξη της παλάμης.

Το πείραμα με το Πληκτρολόγιο K [30] υποδεικνύει πως η γωνία της πλευρικής κλίσης των δύο τμημάτων δεν επηρεάζει την ταχύτητα και το ρυθμό δακτυλογράφησης. Στις δοκιμές διαφορετικών γωνιών (0, 30, 60 και 90 μοίρες) τα δεδομένα έδειξαν πως οι χρήστες προτιμούν την κλίση των 60 μοιρών. Σε ένα άλλο πείραμα [30] οι δακτυλογράφοι έκαναν σε σημαντικό βαθμό περισσότερα λάθη (περίπου 5% παραπάνω) σε κανονικά πληκτρολόγια από ότι σε Πληκτρολόγια K με 45 μοίρες κλίσης. Οι χρήστες των κανονικών πληκτρολογίων αρκετά συχνά παραπονιούνται για πόνο, ενώ οι δακτυλογράφοι σε Πληκτρολόγια K είχαν αναφέρει ότι η δακτυλογράφηση ήταν περισσότερο ευχάριστη.

Το Πληκτρολόγιο STR

Ο Nakaseko [42] με μερικούς συναδέλφους του (Nakaseko, Grandjean, Hunting και Gierer το 1985) πραγματοποίησαν μια σειρά από πειράματα δοκιμάζοντας διαφορετικές γωνίες μεταξύ των διαχωρισμένων τμημάτων, δοκιμάζοντας επίσης διαφορετικές γωνίες κλίσης και διαφορετικά μεγέθη πλήκτρων πάνω σε πειραματικό πληκτρολόγιο. Μεταγενέστερες έρευνες πάνω στο πληκτρολόγιο αυτό οδήγησαν σε ένα τελικό προϊόν το Standard Telephone and Radio (STR) [60], το οποίο κέρδισε το πρώτο βραβείο σχεδίασης στο Ergodesign το 1984.

Στο αρχικό πείραμα ο Nakaseko ανέφερε πως οι χρήστες βρήκαν το διαχωρισμένο πληκτρολόγιο ευχάριστο και ενόησαν τη σχεδίαση με κλίση 10 μοιρών και γωνία ανοίγματος 25 μοίρες. Στο δεύτερο πείραμα, 30 συμμετέχοντες δακτυλογράφησαν για 30 λεπτά σε κάθε ένα από τα τρία πληκτρολόγια. Το ένα από τα πληκτρολόγια ήταν ένα παραδοσιακό πληκτρολόγιο με άνετη υποστήριξη αγκώνα-καρπού, το άλλο ήταν ένα πειραματικό πληκτρολόγιο (που προτιμήθηκε από τους χρήστες στο προηγούμενο πείραμα) με άνετη επίσης υποστήριξη αγκώνα-καρπού και το τελευταίο ήταν το ίδιο πειραματικό πληκτρολόγιο με μικρή υποστήριξη αγκώνα-καρπού. Οι συμμετέχοντες χρησιμοποίησαν τα πληκτρολόγια με τυχαία σειρά και μετά το τέλος της

χρήσης συμπλήρωναν το ερωτηματολόγιο, το οποίο έδειχνε αν ο χρήστης ένοιωθε πόνο στα άνω άκρα ή στους ώμους. Το πείραμα κατέγραφε την ποσότητα πίεσης που ασκούσαν οι συμμετέχοντες και έκανε πολλές μετρήσεις πάνω στα διαφορετικά μέρη του σώματος. Οι χρήστες ανέφεραν (βαθμολογούσαν) τις αισθήσεις τους μετά από κάθε δοκιμή χρησιμοποιώντας μια κλίμακα επτά βαθμών δηλώνοντας με το επτά πολύ κούραση και με το ένα πολύ ξεκούραση και αξιολογούσαν τα πληκτρολόγια ανάλογα με την απόδοσή τους.

Η τελική αναφορά του Nakaseko βασιζόταν πάνω στην αξιολόγηση της απόδοσης που συμπλήρωσαν οι συμμετέχοντες στο πείραμα και διαπίστωσε ότι λιγότερο από το 30% των συμμετεχόντων προτίμησαν τα παραδοσιακά πληκτρολόγια, αφού τα δύο τρίτα των χρηστών προτίμησαν ένα από τα δύο πειραματικά πληκτρολόγια.

Οι Brigham και Clark [6] το 1986 έκαναν μια μελέτη συγκρίνοντας το Πληκτρολόγιο STR με το κανονικό (τυπικό) πληκτρολόγιο. Είκοσι πεπειραμένοι δακτυλογράφοι εξασκήθηκαν για είκοσι λεπτά με τα πληκτρολόγια και μετά πληκτρολόγησαν επτά φορές από είκοσι λεπτά με πέντε λεπτά διάλειμμα μεταξύ τους. Ο Brigham και ο Clark ανέφεραν πως η απόδοση του κανονικού πληκτρολογίου ήταν ανώτερη από το Πληκτρολόγιο STR σε όλες τις περιπτώσεις. Οι συμμετέχοντές του πειράματος έκριναν επίσης πως το κανονικό πληκτρολόγιο ήταν πιο άνετο στη χρήση και για αυτό το λόγο προτιμήθηκε από αυτούς.

Το Πληκτρολόγιο Kinesis

Το πληκτρολόγιο Kinesis είναι ένα ανάγλυφο, μη-ρυθμιζόμενο διαχωρισμένο πληκτρολόγιο με υποστήριξη αγκώνα-καρπού, το οποίο εξέχει από τη πρώτη γραμμή των πλήκτρων κατά 14 cm. Οι Jahns, Litewka, Farrand και Hargreaves [28] το 1991 ανέφεραν ένα πιλοτικό πείραμα, στο οποίο οκτώ υγιείς δακτυλογράφοι χρησιμοποίησαν το πληκτρολόγιο Kinesis περίπου για οκτώ ώρες σε τρεις φάσεις. Οι έξι από τους οκτώ συμμετέχοντες ανέφεραν πως γενικά προτίμησαν το πληκτρολόγιο Kinesis. Ωστόσο, μετά από οκτώ ώρες πρακτικής πάνω στα πληκτρολόγια Kinesis οι συμμετέχοντες στο πείραμα ακόμη δακτυλογραφούσανε, κατά μέσο όρο, πιο αργά από ότι σε κανονικό πληκτρολόγιο (περίπου 95% της βασικής τους απόδοσης) και με περισσότερα λάθη.

Οι Smith και Cronin [51] το 1993 πραγματοποίησαν ένα πείραμα στο οποίο 25 συμμετέχοντες χρησιμοποίησαν πληκτρολόγια και των δύο ειδών, το κανονικό και το πληκτρολόγιο Kinesis, για τη δακτυλογράφηση κανονικού και τυχαίου κειμένου. Αντίθετα από το πληκτρολόγιο Kinesis, το κανονικό πληκτρολόγιο σε αυτή την περίπτωση δεν έχει ενσωματωμένη την υποστήριξη αγκώνα-καρπού. Κατά τη διάρκεια του τμήματος του πειράματος όπου γινόταν η σύγκριση, οι μισοί χρησιμοποιούσαν το κανονικό πληκτρολόγιο στην αρχή και οι άλλοι μισοί χρησιμοποιούσαν το Kinesis στην αρχή. Οι ερευνητές επίσης συγκέντρωναν και ηλεκτρομυϊκές (EMG) μετρήσεις από 11 χρήστες. Τα αποτελέσματα των EMG έδειξαν την επιβάρυνση των μυών των χεριών και του αγκώνα. Το αποτέλεσμα έδειξε πως οι χρήστες δακτυλογραφούσαν σημαντικά γρηγορότερα με το κανονικό πληκτρολόγιο, αλλά δεν υπήρχε σημαντική διαφορά στα λάθη. Οι συμμετέχοντες για λόγους άνεσης προτίμησαν το πληκτρολόγιο Kinesis, αλλά διάλεξαν το κανονικό πληκτρολόγιο για λόγους απόδοσης.

Ασύρματα Πληκτρολόγια

Υπάρχουν δυο ξεχωριστές τεχνολογίες ασύρματων πληκτρολογίων βασισμένες σε υπέρυθρες (IR) και σε ραδιοσυχνότητες (RF) (Σχήμα 10).

Το υπέρυθρο φως δεν έχει καμία διαφορά από οποιαδήποτε άλλη πηγή φωτός και υποτάσσεται στους ίδιους νόμους της φυσικής, με μόνη διαφορά ότι δουλεύει κάτω από τη συχνότητα του κόκκινου φωτός και για αυτό δεν είναι ορατό από το ανθρώπινο μάτι.



Σχήμα 10: Ασύρματο Πληκτρολόγιο

Τα υπέρυθρα σήματα παράγονται κυρίως από ένα φωτοδίοδο (Light Emitting Diode). Τα σήματα αυτά έχουν ένα περιορισμό που συνήθως αναφέρεται σαν “οπτική ευθεία”, όπου ο αποδέκτης και ο αποστολέας πρέπει να βρίσκονται σε οπτική ευθεία ο ένας από τον άλλον. Οι τοίχοι και άλλα φυσικά εμπόδια εμποδίζουν τα σήματα αυτά, περιορίζοντας έτσι την επικοινωνία σε ένα δωμάτιο. Τα υπέρυθρα σήματα έχουν και περιορισμό απόστασης που είναι γύρω στα 30 μέτρα.

Οι ραδιοσυχνότητες διαφέρουν από τα υπέρυθρα στο ότι δεν περιορίζονται στο ένα δωμάτιο. Οι συνηθισμένοι τοίχοι δεν είναι ικανοί να εμποδίσουν τις ραδιοσυχνότητες οπότε δεν υπάρχει ο περιορισμός της “οπτικής ευθείας”. Επίσης είναι δυνατή η μετάδοση του σήματος πάνω από 30 μέτρα.

Πληκτρολόγια χωρίς πλήκτρα

Τα πληκτρολόγια αυτά αποτελούνται από δυο τρούλους πάνω στους οποίους άνετα ξεκουράζονται τα χέρια. Ο κάθε τρούλος κινείται, “ολισθαίνει” σε μία από τις οκτώ κατευθύνσεις σε σχέση με την αρχική θέση ξεκούρασης. Μπορούμε να θεωρήσουμε τις κινήσεις αυτές πάνω σε μια διάταξη πυξίδας: Β, ΒΔ, Δ, ΝΔ, Ν, ΝΑ, Α, ΝΑ. Κάθε τρούλος είναι σε θέση να ολισθαίνει και στις οκτώ κατευθύνσεις (Σχήμα 11).

Οι δακτυλογράφοι δημιουργούν πατήματα πλήκτρων απλώς ολισθαίνοντας τους δυο τρούλους σε μια από τις οκτώ σχετικές κατευθύνσεις. Αυτό μπορεί να ηχεί λίγο περίπλοκο, αλλά είναι εύκολο στη χρήση. Οι δακτυλογράφοι μπορούν να φτάσουν σε αξιόλογες ταχύτητες δακτυλογράφησης μόνο με λίγη εκμάθηση.



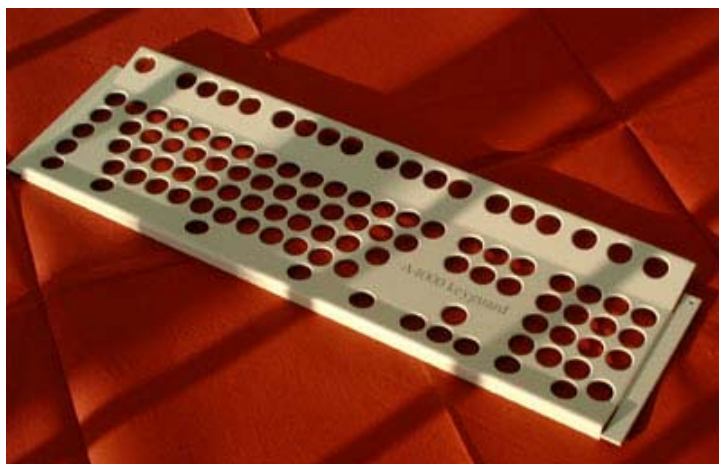
Σχήμα 11: Πληκτρολόγιο χωρίς Πλήκτρα

Σταθεροποιητής για Πληκτρολόγιο (Key guards)

Ο Σταθεροποιητής για το πληκτρολόγιο είναι μια πλάκα η οποία εφαρμόζεται πάνω στο πληκτρολόγιο και έχει τρύπες πάνω από ορισμένα ή και όλα τα πλήκτρα του πληκτρολογίου. Ο σταθεροποιητής για πληκτρολόγιο ξεκουράζει τα χέρια των χρηστών κατά τη δακτυλογράφηση αλλά κυρίως έχει σχεδιαστεί για να εμποδίζει τα απρόσεκτα πατήματα των πλήκτρων που μπορεί να προκαλέσει ένας χρήστης με κινητικές δυσκολίες (Σχήμα 12, 13). Με τη χρήση αυτών των συσκευών μπορούν να αποφευχθούν οι ταυτόχρονες ενέργειες που απαιτούνται για κάποιες λειτουργίες όπως εκείνες που απαιτούν πάτημα μερικών πλήκτρων ταυτόχρονα.



Σχήμα 12: Σταθεροποιητής για Πληκτρολόγιο



Σχήμα 13: Σταθεροποιητής για Πληκτρολόγιο

Τα άτομα με ειδικές ανάγκες που σχετίζονται με αυτά τα προϊόντα παρουσιάζουν:

- Τρεμούλιασμα χεριών
- Μειωμένο συντονισμό κινήσεων.

Εικονικά Πληκτρολόγια

Το εικονικό πληκτρολόγιο (Σχήμα 14) είναι μια εφαρμογή που προσομοιώνει το κανονικό πληκτρολόγιο πάνω στην οθόνη του υπολογιστή. Οι χρήστες έχουν τη δυνατότητα να “πατάνε” τα πλήκτρα του πληκτρολογίου χρησιμοποιώντας ποντίκια, ιχνόσφαιρες ή κάποια από τις συσκευές ένδειξης, οι οποίες παρουσιάζονται στα επόμενα κεφάλαια.

Τα πλήκτρα που πατάει ο χρήστης στέλνουν το αντίστοιχο σύμβολο ή την αντίστοιχη ενέργεια σε ένα άλλο πρόγραμμα που έχει επιλέξει ο χρήστης και που συνήθως είναι ένας επεξεργαστής κειμένου ή ένα πρόγραμμα διαχείρισης ηλεκτρονικού ταχυδρομείου.

Με τη βοήθεια του εικονικού πληκτρολογίου έχουμε τη δυνατότητα να ελέγχουμε ολόκληρο τον υπολογιστή μόνο με τη βοήθεια μίας συσκευής ένδειξης. Τα πληκτρολόγια αυτά προτείνονται κυρίως στους χρήστες με κινητικές αναπηρίες ? και στα άτομα με δυσκινησία. Τα εικονικά πληκτρολόγια χρησιμοποιούνται επίσης σε συνδυασμό με μια οθόνη αφής εξαλείφοντας την ανάγκη του πληκτρολογίου.



Σχήμα 14: Εικονικό Πληκτρολόγιο

Συμπέρασμα

Η βιβλιογραφία δίνει δύο αρκετά καθαρά αποτελέσματα της αλληλοεπίδρασης των χρηστών με τα διαχωρισμένα πληκτρολόγια. Πρώτον, η ταχύτητα δακτυλογράφησης είναι γενικά μικρότερη από ότι στο κανονικό πληκτρολόγιο. Δεύτερον, οι μετρήσεις EMG δείχνουν μειωμένη κούραση μυών για πληκτρολόγια με υποστήριξη αγκώνα-καρπού. Οι μετά-αναλύσεις, που έκανε ο Lewis [32] το 1995 πάνω στα δεδομένα που συνδυάστηκαν από όλα τα πειράματα που έγιναν το διάστημα από 1972 έως και το 1993, έδειξαν πως γενικά οι χρήστες ανεξάρτητα του μορφωτικού τους επιπέδου προτιμούν τα κανονικά πληκτρολόγια.

2.4.2. Έγχορδα Πληκτρολόγια (Chord Keyboards)

Με τα παραδοσιακά πληκτρολόγια οι δακτυλογράφοι πατούσαν ένα πλήκτρο κάθε φορά για να γράψουν ένα χαρακτήρα. Τα έγχορδα πληκτρολόγια (Σχήμα 15) επιτρέπουν να πατάμε αρκετά πλήκτρα ταυτόχρονα για την είσοδο δεδομένων (όπως το χτύπημα των πλήκτρων του πιάνου).



Σχήμα 15: Έγχορδο Πληκτρολόγιο

Επειδή ο συνδυασμός των πλήκτρων καθορίζει την είσοδο, τα έγχορδα πληκτρολόγια απαιτούν λιγότερα πλήκτρα από ότι τα κανονικά πληκτρολόγια. Για παράδειγμα ένα έγχορδο πληκτρολόγιο πέντε πλήκτρων με δυαδικούς διακόπτες μπορεί να παράγει μέχρι και 31 συνδυασμούς πλήκτρων. Το έγχορδο πληκτρολόγιο πέντε πλήκτρων με τριαδικούς διακόπτες μπορεί να παράγει μέχρι και 242 συνδυασμούς. Τα έγχορδα πληκτρολόγια δύο χεριών επιτρέπουν ακόμα περισσότερους συνδυασμούς. Μερικά έγχορδα πληκτρολόγια μεταφράζουν τις χορδές σε φωνήματα και συλλαβές, ενώ άλλα μεταφράζουν τις χορδές σε χαρακτήρες και αριθμούς.

Τα πιο συνηθισμένα έγχορδα πληκτρολόγια είναι το Stenograph (1930) (Σχήμα 16) και το Palantype (1941). Οι δακτυλογράφοι που χρησιμοποιούν αυτά τα πληκτρολόγια μαθαίνουν να συνδυάζουν χορδές με φωνήματα για να παράγουν συλλαβές.

Αυτή η μέθοδος δακτυλογράφησης επιτρέπει στους στενογράφους να αναπτύσσουν πιο γρήγορους ρυθμούς εισόδου με υψηλή ταχύτητα που κυμαίνεται περίπου στις 250 με 300 λέξεις ανά λεπτό και η οποία είναι ταχύτερη από αυτή και των πιο γρήγορων δακτυλογράφων κατά 100 λέξεις ανά λεπτό. Αυτό όμως απαιτεί τρία περίπου χρόνια απόκτησης εμπειρίας [3]. Οι Beddoes και Hu [3] σε μια εργασία τους ανέπτυξαν ένα καινούργιο έγχορδο πληκτρολόγιο για στενογράφους (το ελάχιστο έγχορδο πληκτρολόγιο ή το MCS), που σχεδιάστηκε για να είναι ευκολότερο στην εκμάθηση. Στο MCS οι χορδές αναπαριστούν ζευγάρια φωνημάτων. Πέντε άνθρωποι που έμαθαν MCS δακτυλογράφηση, μετά από 50 ώρες πρακτικής, αύξησαν την ταχύτητα δακτυλογράφησης από 38 σε 52 λέξεις ανά λεπτό.



Σχήμα 16: Έγχορδο Πληκτρολόγιο

Ο Gopher [21] με τους συνεργάτες του είχαν αναφέρει ένα πείραμα που χρησιμοποιούσε ένα έγχορδο πληκτρολόγιο δύο χεριών που σχεδιάστηκε για εισαγωγή κειμένου [21]. Οι χορδές που χρησιμοποιήθηκαν για τη δημιουργία των γραμμάτων ήταν οι ίδιες σε κάθε μισό και αντιστοιχούσαν στα ίδια δάχτυλα του κάθε χεριού. Οι Gopher και Raij διεξήγαγαν ένα πείραμα με άτομα, τα οποία τύπωναν Εβραϊκό κείμενο σε πληκτρολόγιο QWERTY (τέσσερις συμμετέχοντες), σε έγχορδο πληκτρολόγιο με ένα χέρι (πέντε συμμετέχοντες) και σε έγχορδο πληκτρολόγιο με δύο χέρια (έξι συμμετέχοντες). Ο Gopher ανέφερε πως οι συμμετέχοντες στο πείραμα απομνημόνευσαν τους 23 κωδικούς χαρακτήρων του έγχορδου πληκτρολόγιου περίπου σε 30-45 λεπτά. Αρχικός ρυθμός εισόδου για όλες τις περιπτώσεις ήταν περίπου 7 με 8 λέξεις ανά λεπτό. Μετά από 25 ώρες εκμάθησης ο ρυθμός δακτυλογράφησης για έγχορδα πληκτρολόγια ήταν 32 λέξεις ανά λεπτό, ενώ για το QWERTY ήταν περίπου 20 λέξεις ανά λεπτό. Με παρατεταμένη εκμάθηση (περισσότερο από 25 ώρες) οι συμμετέχοντες χρησιμοποιώντας το έγχορδο πληκτρολόγιο με δύο χέρια, παράγαν ταχύτερη δακτυλογράφηση από τους άλλους συμμετέχοντες (περίπου 59 λέξεις ανά λεπτό).

2.4.3. Συμπεράσματα

Αν και τα σύγχρονα τυπικά πληκτρολόγια αντανακλούν ορισμένες αποφάσεις σχεδιασμού που ελήφθησαν πριν από 100 χρόνια, ενσωματώνονται εδώ και σχεδόν έναν αιώνα στις μεταγενέστερες έρευνες που σχετίζονται με τη δακτυλογράφηση και με τη συμπεριφορά πλήκτρων. Τα αντιπροσωπευτικά δείγματα εναλλακτικών πληκτρολογίων, όπως η διάταξη Dvorak, τα διαχωρισμένα πληκτρολόγια και τα έγχορδα πληκτρολόγια, έχουν γενικά αποτύχει να αποδείξουν στην πράξη τη δυνατότητα ολοκληρωτικής αντικατάστασης των τυπικών πληκτρολογίων, αν και μπορεί να χρησιμοποιηθούν σε ειδικές εφαρμογές και σε ειδικά περιβάλλοντα. Το καλά σχεδιασμένο τυπικό πληκτρολόγιο είναι εξαιρετικά αποτελεσματική συσκευή εισόδου δεδομένων και πιθανότατα θα παραμείνει η σημαντικότερη συνιστώσα στην αλληλεπίδραση ανθρώπου μηχανής στο προβλεπόμενο μέλλον.

3. Συσκευές Ένδειξης (Pointing Devices)

3.1. Εισαγωγή

Όταν η οθόνη χρησιμοποιείται για την παρουσίαση πληροφοριών, όπως οι πληροφορίες εναέριας κυκλοφορίας, σύνταξης κειμένου και Σχεδίασης Υποβοηθούμενης από Υπολογιστή CAD (Computer Aided Design), συχνά είναι δόκιμο, μόλις εμφανιστεί το αντικείμενο, να μπορούμε να το επιλέγουμε άμεσα. Αυτή η προσέγγιση, που παρέχει δυνατότητα άμεσης διαχείρισης, είναι ελκυστική επειδή δεν απαιτεί την εκμάθηση ειδικών εντολών από τους χρήστες, ενώ παράλληλα ελαττώνει τις περιπτώσεις τοπογραφικών λαθών στο πληκτρολόγιο και επικεντρώνει την προσοχή τους στην οθόνη.

Οι συσκευές ένδειξης καθορίζουν την τοποθεσία και την κίνηση στο δισδιάστατο επίπεδο και περιλαμβάνουν: τις οθόνες αφής (touch screens), τις φωτογραφίδες (light pens), τις πινακίδες γραφικών (graphic tablets), τα ποντίκια (mices), τις ιχνόσφαιρες (trackballs) και τα χειριστήρια (joysticks). Το αποτέλεσμα χρησιμοποίησης αυτών των συσκευών αυξάνει συνήθως την απόδοση, παρουσιάζει λιγότερα λάθη, προσφέρει ευκολότερη μάθηση και μεγαλύτερη ικανοποίηση. Οι συσκευές ένδειξης κατατάσσονται σε έξι διαφορετικές κατηγορίες αλληλεπίδρασης:

1. **Επιλογή (Select):** Ο χρήστης επιλέγει από ένα σύνολο αντικειμένων. Αυτή η τεχνική χρησιμοποιείται για παραδοσιακές επιλογές από μενού, για εντοπισμό ενός αρχείου στον κατάλογο αρχείων ή για σημάδεμα ενός μέρους της οθόνης όπως στο σχεδιασμό αυτοκινήτων.
2. **Τοποθεσία (Position):** Ο χρήστης επιλέγει ένα σημείο στο μονοδιάστατο, δισδιάστατο, τρισδιάστατο ή πολυδιάστατο χώρο. Η τοποθεσία μπορεί να χρησιμοποιηθεί στη δημιουργία σχεδίων, στην δημιουργία καινούργιου παραθύρου ή στη μετακίνηση ενός μπλοκ κειμένου μέσα στην εικόνα.
3. **Προσανατολισμός (Orient):** Ο χρήστης επιλέγει την κατεύθυνση στο μονοδιάστατο, δισδιάστατο, τρισδιάστατο ή πολυδιάστατο χώρο. Ο προσανατολισμός μπορεί απλά να στρέψει την εικόνα στην οθόνη, να υποδείξει την κατεύθυνση της κίνησης στην οθόνη ή να ελέγχει τη λειτουργία του χειριού ενός ρομπότ.
4. **Ποσοτικοποίηση (Quantify):** Ο χρήστης καθορίζει αριθμητικές μεταβλητές. Η ενέργεια της ποσοτικοποίησης είναι συνήθως η μονοδιάστατη επιλογή ακέραιων ή πραγματικών αριθμών από ένα σύνολο παραμέτρων όπως οι αριθμοί σελίδων στα έγγραφα, ταχύτητας του πλοίου ή πλάτους του ήχου.
5. **Κείμενο (Text):** Ο χρήστης εισάγει κείμενα στο δισδιάστατο χώρο. Η συσκευή ένδειξης καθορίζει την τοποθεσία της εισαγωγής, διαγραφής ή αλλαγής. Έκτος από τα απλές λειτουργίες πάνω στα κείμενα, υπάρχουν και πιο πολύπλοκες λειτουργίες, όπως το κεντράρισμα, η συγχώνευση, το μέγεθος γραμματοσειράς, ο τονισμός, τα έντονα ή πλάγια γράμματα και η σελιδοποίηση.

Όλες οι προαναφερόμενες λειτουργίες είναι δυνατόν να πραγματοποιηθούν με χρήση ενός πληκτρολογίου, πληκτρολογώντας απλώς αριθμούς ή γράμματα για επιλογή, ακέραιες συντεταγμένες για την τοποθεσία, αριθμό που αντιπροσωπεύει τη γωνία του αντικειμένου, αριθμούς ποσοτικοποίησης, εντολές μετακίνησης του κέρσορα στην οθόνη (βελάκια). Ενώ στο παρελθόν όλες αυτές οι λειτουργίες πραγματοποιούνταν με τη βοήθεια του πληκτρολογίου, τώρα έχουν κατασκευασθεί καινούργιες συσκευές οι οποίες επιτρέπουν εκτέλεση αυτών των λειτουργιών ταχύτερα και με λιγότερα λάθη.

Οι προαναφερόμενες συσκευές μπορούν να ενταχθούν σε δύο κατηγορίες: α) συσκευές άμεσου έλεγχου της επιφάνειας της οθόνης όπως είναι η φωτογραφίδα, η οθόνη αφής και β) συσκευές έμμεσου έλεγχου της οθόνης όπως είναι οι επιφάνειες αφής, το ποντίκι, η ιχνόσφαιρα, το χειριστήριο, τα πλακίδια γραφικών.

3.2. Συσκευές Άμεσης Ένδειξης (Direct Pointing Devices)

3.2.1. Οθόνες Αφής (Touch Screens)

Τεχνολογίες

Οι οθόνες αφής (Σχήμα 17) παράγουν σήματα εισόδου ως ανταπόκριση στο άγγιγμα ή την κίνηση του δαχτύλου πάνω στην οθόνη. Υπάρχουν δύο θεμελιώδεις αρχές λειτουργίας αυτών των συσκευών: είτε επικοινωνεί το επικάλυμμα, είτε τα παραγόμενα σήματα πάνω στην οθόνη επικοινωνούν με διακοπές.



Σχήμα 17: Οθόνη Αφής

Στην πρώτη κατηγορία περιλαμβάνονται οι αντιστατικές, οι χωρητικές, οι πιεζοηλεκτρικές και οι συσκευές διασταυρωμένων συρμάτων (cross-wire). Οι αντιστατικές οθόνες αφής έχουν δύο επίπεδα, ένα αντιστατικό και ένα αγώγιμο. Όταν εφαρμόζεται η πίεση στη οθόνη αφής, οι δύο επιφάνειες εφάπτονται και κλίνει το κύκλωμα. Η χωρητική οθόνη αφής έχει μία αγώγιμη επίστρωση πάνω στην επιφάνεια του γυαλιού της οθόνης. Η χωρητικότητα του σώματος είναι η αιτία που γεννιάται ηλεκτρικό σήμα, όταν ο χρήστης ακουμπάει το επικάλυμμα της οθόνης.

Η πιεζοηλεκτρική οθόνη αφής έχει επίστρωση γυαλιού με διάταξη μετατροπής της ισχύος σε κάθε γωνία του επιστρώματος. Όταν πιεστεί ένα συγκεκριμένο σημείο στην επίστρωση του γυαλιού, η πίεση μεταδίδεται στη διάταξη μετατροπής, η οποία και καθορίζει την αρχική τοποθεσία. Η κίνηση του δαχτύλου που ακολουθεί το πάτημα δεν γίνεται αντιληπτή με αυτή την τεχνική. Η συσκευή διασταυρωμένων συρμάτων χρησιμοποιεί ένα πλέγμα οριζόντιων και κάθετων συρμάτων σε διαφανές φύλλο τοποθετημένο στην οθόνη. Η τάση εμφανίζεται είτε στα οριζόντια είτε στα κάθετα σύρματα και το σήμα δημιουργείται, όταν η ισχύς του πατήματος συμπίπτει με τη διασταύρωση του οριζοντίου και του κάθετου σύρματος.

Οι οθόνες αφής κυμάτων και οι οθόνες αφής υπέρυθρων ενεργοποιούνται, όταν το δάχτυλο προκαλεί διακοπή του σήματος. Στις οθόνες ηχητικών κυμάτων η γυάλινη πλάκα τοποθετείται πάνω στην οθόνη και τα υπερηχητικά κύματα εμφανίζονται στο γυαλί μέσω διάταξης μετατροπής που είναι τοποθετημένη στις άκρες του. Όταν γίνεται διακοπή της δέσμης ακουστικού κύματος οι αντανακλώμενες οριζόντιες και κάθετες κυμάνσεις γίνονται αντιληπτές από τη διάταξη μετατροπής.

Στις οθόνες αφής υπέρυθρων οι φωτοδιόδοι εκπομπής συνδυάζονται με ανιχνευτές φωτός και τοποθετούνται στις αντίθετες άκρες της οθόνης. Όταν ο χρήστης ακουμπήσει το δάχτυλο στην οθόνη γίνεται διακοπή δέσμης φωτός.

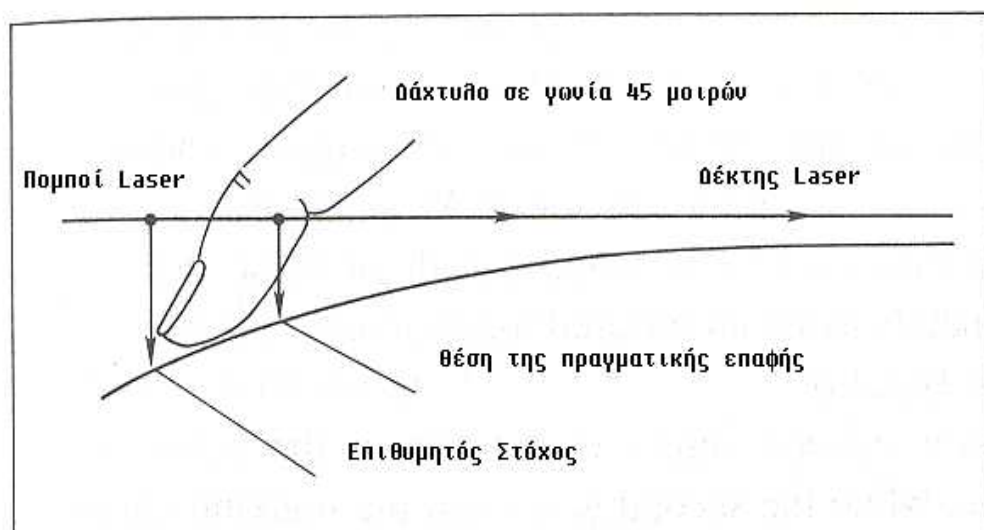
Ανάλυση αφής

Οι οθόνες αντίστασης τείνουν να έχουν τη μεγαλύτερη ανάλυση αφής, περίπου από 1000x1000 μέχρι 4000x4000 ξεχωριστά σημεία. Η ανάλυση των χωρητικών οθονών είναι σχετικά μεγάλη, ενώ οι ακουστικές επιφάνειες έχουν γενικά χαμηλή ανάλυση. Οι υπέρυθρες οθόνες έχουν το πολύ 25x40 σημεία αφής. Για εφαρμογές που απαιτούν μεγάλη ανάλυση αφής οι υπέρυθρες συσκευές είναι γενικώς ακατάλληλες.

Παράλλαξη

Παράλλαξη εμφανίζεται, όταν η επιφάνεια αφής ή οι ανιχνευτές αφής είναι διαχωρισμένοι (βρίσκονται σε απόσταση) από το στόχο. Σε όλες τις τεχνολογίες η επιφάνεια αφής ακουμπάει τουλάχιστον ελαφρά την επιφάνεια της οθόνης. Αν το επικάλυμμα είναι διαχωρισμένο από την οθόνη, παρουσιάζονται προβλήματα παράλλαξης. Αυτό το πρόβλημα μπορεί να μειωθεί, αν αναγκάσουμε το χρήστη να είναι ακριβώς απέναντι από το στόχο και να ακουμπάει κάθετα την οθόνη. Δυστυχώς αυτή η απαίτηση μειώνει το πλεονέκτημα της χρήσης αυτής της συσκευής, ειδικά όταν οι στόχοι βρίσκονται στις άκρες της οθόνης.

Όταν οι επιφάνειες αφής υπέρυθρων εφαρμόζονται στις κοινές οθόνες καθοδικού



Σχήμα 18 : Παράλλαξη οφειλόμενη στη μετακίνηση της ακτίνας φωτός από την επιφάνεια της οθόνης.

σωλήνα (CRT), το πρόβλημα της παράλλαξης είναι πιο σοβαρό. Επειδή η δέσμη φωτός ταξιδεύει μόνο ευθεία, αν η οθόνη και η επιφάνεια είναι διαχωρισμένες, η ενέργεια της επιλογής παρουσιάζει λάθη (Σχήμα 18).

Αντοχή

Οι οθόνες αφής παρουσιάζουν κυρίως πρόβλημα, όταν λερώνεται η επιφάνειά τους μετά από συνεχή λειτουργία. Οι αντιστατικές οθόνες χρησιμοποιούν ένα πλαστικό πλέγμα το οποίο μπορεί να γρατσουνιστεί από τη συνεχή χρήση. Οι οθόνες ακουστικής αφής παρουσιάζουν τα ίδια προβλήματα με τις αντιστατικές οθόνες. Οι χωρητικές, πιεζοηλεκτρικές και υπέρυθρες οθόνες τείνουν να αντιστέκονται στις καταστροφές.

Περιορισμοί Περιβάλλοντος

Ο στατικός ηλεκτρισμός είναι δυνατόν να προκαλέσει προβλήματα σε ορισμένες κατηγορίες οθονών αφής. Οι συσκευές ακουστικών κυμάτων μπορεί να ενεργοποιούνται από τη σκόνη και τα γρατσουνίσματα που είναι πάνω στην οθόνη. Επίσης οι οθόνες υπέρυθρων επίσης μπορεί να χαλάσουν από τη σκόνη και τον καπνό.

Ευκολία Χρήσης

Ωστόσο, ένα από τα πιο σημαντικά χαρακτηριστικά των οθονών αφής, το οποίο δεν έχει αναφερθεί πιο πριν, είναι η ευκολία εκμάθησης της λειτουργίας τους. Έτσι, επειδή οι οθόνες αφής απαιτούν φυσικές χειρονομίες ένδειξης, που χρειάζονται ελάχιστη εκμάθηση των βασικών αρχών, μπορούν να ανταποκρίνονται με διαφορετικούς τρόπους στις ενέργειες των χρηστών. Για παράδειγμα, επειδή οι δεσμοί LED για οθόνες αφής υπέρυθρων βρίσκονται πιο πάνω από την οθόνη, ο χρήστης δεν χρειάζεται να ακουμπά την οθόνη για να ενεργοποιήσει τη συσκευή. Αν ο χρήστης δεν το αντιλαμβάνεται, είναι δυνατόν να παρουσιαστεί πρόβλημα με απρόσεκτες ενέργειες και επιλογές μη σωστών αντικειμένων που οφείλονται στην παράλλαξη.

Ένας άλλος παράγοντας είναι ότι χρειάζεται πίεση για την ενεργοποίηση του επιστρώματος. Αν απαιτείται μεγάλη πίεση, η ενεργοποίηση μπορεί να πάρει αρκετά μεγάλο χρονικό διάστημα, πράγμα που μπορεί να ενοχλήσει τον χρήστη. Τέλος, οι οθόνες με μικρή ανάλυση είναι πιθανόν να οδηγήσουν σε απογοητευτικά αποτελέσματα, αν τα σημεία αφής δεν είναι κεντραρισμένα πάνω στους στόχους.

Εμπειρικά Αποτελέσματα

Οι Schulze και Snyder [49] το 1983 συνέκριναν πέντε διαφορετικές οθόνες αφής: ακουστικές, χωρητικές, αντιστασιακές, διασταυρωμένων συρμάτων (cross-wire) και υπέρυθρων. Οι συσκευές με υπέρυθρα επιτρέπουν τη μεγαλύτερη ανάλυση, επειδή δεν έχουν επίστρωμα. Οι χωρητικές συσκευές χρησιμοποιούνται στις οθόνες με το μικρότερο θόρυβο και με μικρή μεταβολή φωτεινότητας. Επομένως, καταλήγοντας συμπεραίνουμε ότι οι συσκευές υπέρυθρων και οι συσκευές διασταυρωμένων συρμάτων παρουσιάζουν τη καλύτερη απόδοση.

Μέγεθος Πλήκτρου

Οι Hall, Cunningham, Roache και Cox [22] το 1988 ερεύνησαν τις επιπτώσεις διαφορετικών παραγόντων στην ακρίβεια, με την οποία ο χρήστης επιλέγει τον στόχο στις πιεζοηλεκτρικές οθόνες αφής. Οι χρήστες χρησιμοποιούν την οθόνη του υπολογιστή και όρθιοι και καθισμένοι. Για καθισμένους χρήστες η έρευνα έδειξε πώς δεν παρουσιάζεται μεγαλύτερο όφελος στην ακρίβεια επιλογής από την αύξηση του μεγέθους του στόχου, του ύψους πέρα από το 22 mm (πιθανότητα σωστής επιλογής 99,5%) και το πλάτος πέρα από το 28mm (πιθανότητα σωστής επιλογής 99,6%). Για όρθιους χρήστες χρειάζεται μεγαλύτερο μέγεθος στόχου για να επιτευχθεί η ίδια ακρίβεια με την προηγούμενη περίπτωση. Δεν υπάρχει αύξηση οφέλους όταν το ύψος φτάσει τα 30 mm (πιθανότητα σωστής επιλογής 99,7%) και το πλάτος φτάσει τα 32 mm (πιθανότητα σωστής επιλογής 99,5%).

Πλεονεκτήματα και Μειονεκτήματα των Οθονών Αφής

Το πιο προφανές πλεονέκτημα των οθονών αφής είναι το ότι η συσκευή εισόδου είναι ταυτόχρονα και συσκευή εξόδου. Για αυτό το λόγο δεν χρειάζεται επιπλέον χώρος και είναι δυνατός ο άμεσος συντονισμός ματιού και δαχτύλου και η άμεση συνάφεια μεταξύ εισόδου του χρήστη και εξόδου της οθόνης. Το δεύτερο πλεονέκτημα είναι το ότι όλες οι εισοδοί βρίσκονται πάνω στην οθόνη και δεν χρειάζεται να απομνημονεύουν διαφορετικές εντολές. Επειδή η μόνη απαίτηση των οθονών αφής είναι η φυσική αφή, χρειάζεται ελάχιστη εκμάθηση ως προς την επιλογή των αντικειμένων. Ο χρήστης μπορεί να ειδικευτεί αρκετά στην επιλογή στόχων σε μικρό χρονικό διάστημα. Τα ειδικά πλεονεκτήματα των χωρητικών, πιεζοηλεκτρικών και οθονών αφής υπέρυθρων είναι το ότι δεν έχουν κινητικά μέρη.

Οι οθόνες αφής έχουν και μερικά μειονεκτήματα. Επειδή η επιφάνεια εξόδου είναι ταυτόχρονα και μέσο πρόσβασης (είσοδος), μπορεί να παρουσιαστούν προβλήματα παράλλαξης. Επίσης, ο χρήστης πρέπει να κάθεται με τα χέρια πλησίον του υπολογιστή. Αυτά τα χαρακτηριστικά των οθονών αφής είναι δυνατόν να

περιορίζουν και τη σχεδίαση και την ευκινησία. Η ανάγκη της επαναλαμβανόμενης ανύψωσης του χεριού προς την οθόνη μπορεί επίσης να προκαλέσει κούραση του χεριού. Ένα άλλο μειονέκτημα είναι το ότι το μέγεθος δακτύλου του χρήστη τείνει να περιορίσει την ανάλυση του στόχου.

Εφαρμογές

Οι οθόνες αφής αν χρησιμοποιούνται σε σχέση με τα δεδομένα που βρίσκονται πάνω στην οθόνη, παρουσιάζουν μέγιστο όφελος. Οι συσκευές αυτές είναι αναποτελεσματικές για την είσοδο νέων γραφικών πληροφοριών ή για ζωγράφισμα με το χέρι. Οι οθόνες αφής είναι αρκετά πρακτικές στις εφαρμογές επιλογής από μενού. Επίσης, οι οθόνες αφής μπορούν να συνεργαστούν με εικονικά πληκτρολόγια για εισαγωγή δεδομένων, όταν η χρήση των συμβατικών πληκτρολογίων δεν είναι δυνατή ή πρακτική. Οι οθόνες αφής είναι πρακτικές στις εφαρμογές κατά τις οποίες είναι επικίνδυνη η εκτροπή της προσοχής από την οθόνη, όπως στις εφαρμογές ελέγχου εναέριας κυκλοφορίας. Επίσης η χρήση τους είναι ωφέλιμη σε εφαρμογές, όπου η πιθανή είσοδος είναι περιορισμένη και καλά καθορισμένη, όπως στους θαλάμους διακυβέρνησης των αεροπλάνων με σκοπό την εύρεση της πορείας. Τελικά, οι οθόνες αφής είναι αποτελεσματικές στις περιπτώσεις που δεν είναι δυνατόν να επιτευχθεί εκμάθηση της χρήσης του συστήματος. Οι οθόνες αφής έχουν χρησιμοποιηθεί με αποτελεσματικότητα, για παράδειγμα στην παρουσίαση πληροφοριών στα εμπορικά κέντρα, στις τράπεζες και στα ξενοδοχεία.

3.2.2. Φωτογραφίδα (Light Pens)

Τεχνολογίες

Η φωτογραφίδα (Σχήμα 19) είναι μια βελόνα που δημιουργεί πληροφορίες θέσης, όταν έλθει σε επαφή με την επιφάνεια της οθόνης. Το στυλό φωτός περιέχει ανιχνευτή φωτός ή φωτοκύτταρο. Όπου δείχνει το στυλό φωτός, οι δέσμες ηλεκτρονίων της οθόνης ανανεώνουν την κηλίδα του φωσφόρου. Η ενέργεια αυτή γίνεται αντιληπτή από τον ανιχνευτή σαν αύξηση της φωτεινότητας και έτσι παράγεται ένα σήμα. Βασισμένες στο συγχρονισμό του σήματος αυτού, υπολογίζονται οι συντεταγμένες της κηλίδας. Η φωτογραφίδα περιλαμβάνει συνήθως και έναν διακόπτη που εμποδίζει τις άσκοπες ενεργοποιήσεις.



Σχήμα 19. Φωτογραφίδα

Τρόποι Λειτουργίας

Η φωτογραφίδα μπορεί να χρησιμοποιηθεί με δύο τρόπους: με ενδείξεις, κατά την οποία ο χαρακτήρας ή η εικόνα μπορούν να επιλέγονται ενδεικνύοντας το σημείο στην οθόνη και ενεργοποιώντας τη φωτογραφίδα. Ο δεύτερος τρόπος λειτουργίας της φωτογραφίδας είναι ο τρόπος του ιχνηλάτη. Ο χρήστης σκοπεύει τη

φωτογραφίδα στο δρομέα πάνω στην οθόνη και το μετακινεί. Όσο ο δρομέας μένει στην περιοχή επίβλεψης της φωτογραφίδας, η γραμμή θα ακολουθεί την κίνησή του. Η φωτογραφίδα πρέπει να μετακινείται με σταθερό ρυθμό, διαφορετικά θα διακοπεί η ιχνηλασία.

Μέγεθος Στόχου

Το κατάλληλο μέγεθος και η οριακή απόσταση μεταξύ των αντικειμένων πάνω στην οθόνη εξαρτώνται από το μέγεθος της περιοχής που «βλέπει» η φωτογραφίδα. Η μεγάλη περιοχή μπορεί να διευκολύνει περισσότερο την επιλογή μικρών στόχων, απαιτεί όμως η στόχοι να έχουν αραιή διάταξη μεταξύ τους έτσι, ώστε να διακρίνονται μεταξύ τους οι γειτονικοί στόχοι. Η περιοχή βλέψης που είναι κατάλληλη για φωτογραφίδα είναι περίπου 5 εικονοστοιχεία (pixels), ωστόσο, ο Hatamian και ο Brown [23] (1985) ανέπτυξαν φωτογραφίδα με ακρίβεια στο $\frac{1}{4}$ του εικονοστοιχείου πάνω στην οθόνη με ανάλυση 1000x1000. Αυτή η ευριστική ικανότητα επιτρέπει στη φωτογραφίδα να επιλέγει ένα αντικείμενο από μια ομάδα μικρών στόχων που βρίσκονται αρκετά κοντά ο ένας στον άλλο.

Καθώς η ικανότητα ανάλυσης του φωτεινού στυλό αυξάνεται συνεχώς, οι ελάχιστοι περιορισμοί μεγέθους στόχου καθορίζονται από τις ανάγκες του χρήστη και όχι από τις ανάγκες της τεχνολογίας. Για παράδειγμα οι Avons, Beveridge, Hickman και Hitch [1] (1983) δήλωσαν πως τα παιδιά που χρησιμοποιούν φωτογραφίδα για επιλογή στόχου παρουσιάζουν σφάλμα διακύμανσης της θέσης από 1.7 έως 3 mm.

Πλεονεκτήματα και Μειονεκτήματα

Όπως και οι οθόνες αφής, η φωτογραφίδα χρησιμοποιεί την οθόνη και ως συσκευή εξόδου και ως συσκευή εισόδου. Αυτό παρέχει άμεση συνάφεια μεταξύ της εισόδου και της εξόδου και επιτρέπει τη χρησιμοποίηση της φυσικής ικανότητας ένδειξης και σκισσογραφήματος για εισαγωγή δεδομένων. Η χρήση της φωτογραφίδας απαιτεί μια επιπλέον ενέργεια ανύψωσης και κατεβάσματος της συσκευής εισόδου. Επειδή η είσοδος πραγματοποιείται πάνω στην επιφάνεια της οθόνης, ο χρήστης της φωτογραφίδα πρέπει να κάθεται με τα χέρια κοντά στην οθόνη. Επιπλέον, το συνεχές κράτημα του στυλό πάνω στην οθόνη μπορεί να είναι κουραστικό. Ο Avons [1] και οι συνεργάτες του (1983) ανέφεραν πως τα μικρά παιδιά δεν μπορούσαν να κρατήσουν τον αγκώνα τους ακίνητο, με αποτέλεσμα να κουράζονται περισσότερο από τα μεγαλύτερα.

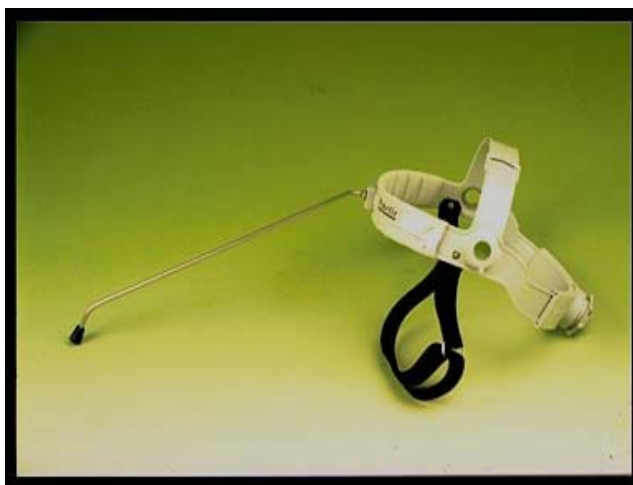
Οι φωτογραφίδες στερούνται συνήθως την δυνατότητα υψηλής ανάλυσης. Ένα πρόβλημα που μπορεί να προκύψει είναι η παράλλαξη, ειδικά όταν το αντικείμενο ένδειξης βρίσκεται στην άκρη της οθόνης. Χρησιμοποιώντας μόνο το κέντρο της οθόνης είναι δυνατόν να λυθεί αυτό το πρόβλημα, ωστόσο η λύση αυτή περιορίζει τη χρησιμότητα της συσκευής. Η πιο ικανοποιητική λύση είναι τα αντικείμενα, που βρίσκονται στις άκρες της οθόνης, να εμφανίζονται αρκετά μεγάλα, με σκοπό να ελαττωθούν στο ελάχιστο (δραστικά) οι επιδράσεις των λανθασμένων τοποθετήσεων. Επίσης η εκμάθηση του χρήστη είναι δυνατόν να συντελέσει, ώστε να ελαχιστοποιηθεί αυτό το είδος λαθών.

Εφαρμογές

Επειδή η φωτογραφίδα χρησιμοποιείται με φυσική χειρονομία ένδειξης είναι κατάλληλη για επιλογές από το μενού. Ενώ η φωτογραφίδα μπορεί να χρησιμοποιηθεί για ζωγραφική, δεν είναι κατάλληλη για ακριβείς σκιαγραφήσεις. Η δυσκολία σταθερού ρυθμού κίνησης που απαιτείται για την ιχνηλασία, προκαλεί πρόσθετα προβλήματα στις εφαρμογές ζωγραφικής. Οι Parrish, Gates, Munger, Grimma και ο Smith (1982) [44] μετά από αρκετά πειράματα κατέληξαν στο συμπέρασμα ότι η φωτογραφίδα είναι πιο χρήσιμη για τον εντοπισμό και την κίνηση ενός συμβόλου στην οθόνη.

3.2.3. Ράβδος κεφαλής (Head stick)

Ο ράβδος κεφαλής είναι μια συσκευή ένδειξης (Σχήμα 20) που προσαρμόζεται πάνω σε κράνος (headpiece) και εκτείνεται από το κέντρο του μετώπου με γωνία προς τα κάτω. Η ράβδος κεφαλής είναι μια συσκευή άμεσης ένδειξης. Χρησιμοποιείται για άμεση επιλογή αντικειμένων, όπως τα πλήκτρα του πληκτρολογίου, ή συμβόλων και λέξεων από ένα ειδικό πληκτρολόγιο. Σε συνδυασμό με οθόνες αφής ο χρήστης έχει τη δυνατότητα να επιλέγει οποιοδήποτε αντικείμενο πάνω στην οθόνη του υπολογιστή. Είναι χρήσιμη στα άτομα που δεν μπορούν να χρησιμοποιήσουν τα χέρια τους ή κάποιο άλλο άκρο του σώματός τους, αλλά διαθέτουν πολύ καλό έλεγχο κίνησης της κεφαλής. Αν ο δείκτης εκτείνεται από το πηγούνι, αναφέρεται σαν ράβδος πηγουνιού (chinstick).



Σχήμα 20. Ράβδος Κεφαλής.

3.3. Έμμεσες Συσκευές Ένδειξης (Indirect Pointing Devices)

3.3.1. Πινακίδες Γραφικών (Graphic Tablets)

Τεχνολογίες

Οι πινακίδες γραφικών αποτελούνται από ένα επίπεδο πλαίσιο τοποθετημένο απέναντι ή δίπλα στην οθόνη (Σχήμα 21). Η επιφάνεια της πινακίδας αναπαριστά την οθόνη και η μετακίνηση του δακτύλου ή του στυλό πάνω στη πινακίδα παρέχει την πληροφορία της θέσης του δρομέα. Η εικόνα παρουσιάζει μία πινακίδα γραφικών που ενεργοποιείται με στυλό.



Σχήμα 21 : Πινακίδα γραφικών.

Στις πινακίδες κωδικοποιημένου πίνακα ένα ειδικό στυλό ανιχνεύει ηλεκτρικά ή μαγνητικά σήματα τα οποία δημιουργούνται από πλέγμα αγωγών στα πλακίδια. Οι πινακίδες βαθμωτής τάσεως έχουν σαν επιφάνεια ένα αγωγίμο πέπλο και με τον υπολογισμό της αύξησης του δυναμικού που προκαλείται από το στυλό είναι δυνατόν να καθορίζεται η θέση του δρομέα πάνω στην οθόνη.

Οι ακουστικές πινακίδες έχουν στυλό που προκαλούν σπίθισμα στην άκρη. Ο ήχος ανιχνεύεται από το μικρόφωνο που βρίσκεται στο γειτονικό άκρο του πλακιδίου. Εναλλακτικά των ακουστικών πλακιδίων είναι τα ηλεκτροακουστικά, τα οποία προκαλούν ηλεκτρικούς παλμούς πάνω στο πλακίδιο που ανιχνεύει το στυλό. Οι ηλεκτροακουστικές πινακίδες είναι λιγότερο ευαίσθητες στο θόρυβο του περιβάλλοντος από τα ακουστικά.

Οι πινακίδες αφής λειτουργούν χωρίς κανένα ειδικό στυλό. Στην τεχνολογία των ακουστικών πινακίδων αφής κύματα υψηλής συχνότητας μεταδίδονται δια μέσου της επιφάνειας. Όταν τα κύματα διακόπτονται από το δάχτυλο αντανακλώνται πίσω στην άκρη της πινακίδας. Η καθυστέρηση μεταξύ της δημιουργίας και της λήψης χρησιμοποιείται για τον υπολογισμό των συντεταγμένων. Οι ωμικές πινακίδες αφής χρησιμοποιούν επίπεδα αγωγίμου και ωμικού υλικού στην επιφάνεια της πινακίδας. Όταν τα επίπεδα πιέζονται από το δάχτυλο δημιουργείται ηλεκτρικό δυναμικό. Οι χωρητικές πινακίδες αφής χρησιμοποιούν το δυναμικό του ανθρώπινου σώματος, για να παράγουν ηλεκτρικά σήματα. Τα σήματα αυτά δημιουργούνται όταν το δάχτυλο του χρήστη εφάπτεται στην επιφάνεια της πινακίδας. Το μειονέκτημα των πινακίδων αφής είναι το ότι μια ακούσια επαφή μπορεί να τις ενεργοποιήσει.

Μέθοδοι Ελέγχου του Δρομέα

Όταν ο χρήστης πατήσει με το δάχτυλο τη πινακίδα, ο δρομέας πάνω στην οθόνη μπορεί να προγραμματιστεί, ώστε να μετακινηθεί από την τρέχουσα τοποθεσία και να εμφανιστεί στην αντίστοιχη θέση του δακτύλου στη πινακίδα. Η κίνηση του δακτύλου πάνω στη πινακίδα γραφικών παράγει νέα θέση που αναφέρεται πάντα στις τρέχουσες συντεταγμένες του δακτύλου πάνω στη πινακίδα γραφικών. Αυτή η μέθοδος ονομάζεται απόλυτος τρόπος ελέγχου του δρομέα. Η δεύτερη προσέγγιση διατηρεί το δρομέα που είναι πάνω στην οθόνη στη θέση του, όταν το δάχτυλο ακουμπήσει αρχικά τη πινακίδα γραφικών. Η κίνηση του δακτύλου προκαλεί αντίστοιχη κίνηση του δρομέα που έχει πάντοτε σχέση με την αρχική θέση του. Αυτή η μέθοδος αναφέρεται σαν σχετικός τρόπος ελέγχου του δρομέα. Στον απόλυτο τρόπο η αρχική θέση του δρομέα ορίζεται κάθε στιγμή, ενώ στο σχετικό τρόπο η αρχική θέση είναι καθορισμένη.

Η επιλογή του τρόπου ελέγχου καθορίζεται από τη φύση της εφαρμογής. Όταν οι περιστάσεις επιβάλλουν το μέγεθος του πλακιδίου να είναι αρκετά μικρότερο από το μέγεθος της οθόνης, ο απόλυτος τρόπος για τη μετακίνηση του δρομέα σε μεγάλες περιοχές πάνω στην οθόνη απαιτεί μικρές κινήσεις στην πινακίδα

Διαμόρφωση του Πλακιδίου

Επειδή η πινακίδα γραφικών δεν εφαρμόζεται πάνω στην οθόνη, άλλα είναι ξεχωριστή συσκευή εισόδου, είναι πιο ευέλικτη σε σχέση με τις οθόνες αφής. Για το λόγο αυτό το μέγεθος της πινακίδας μπορεί να μεταβάλλεται από πολύ μικρό μέχρι αρκετά μεγάλο. Το χαμηλό προφίλ της πινακίδας γραφικών σε σχέση με το χειριστήριο ή την ιχνόσφαιρα, στα οποία θα αναφερθούμε παρακάτω, συντελεί ώστε οι ανεπιθύμητες ενέργειες να είναι λιγότερο πιθανές. Ένα επιπλέον πλεονέκτημα είναι ότι η επιφάνεια της πινακίδας είναι δυνατόν να διαμορφωθεί με πολλούς τρόπους.

Για εφαρμογές που απαιτούν επιλογές από το μενού ένα περίγραμμα μπορεί να τοποθετηθεί πάνω στην πινακίδα έτσι, ώστε να συμφωνεί με την τοποθεσία πάνω στην οθόνη. Εναλλακτικά το περίγραμμα έχει τη δυνατότητα να υποδεικνύει λειτουργίες που δεν είναι άμεσα αναφερόμενες πάνω στην οθόνη. Για παράδειγμα, ένα πακέτο λογισμικού που είναι σχεδιασμένο για τη διδασκαλία μουσικής, μπορεί να περιέχει ένα περίγραμμα με τα πλήκτρα του πιάνου.

Οι Brown, Buxton και Murtagh [8] το 1985 σε μια εργασία τους είχαν προτείνει να χωρίζεται η επιφάνεια της πινακίδας σε διαφορετικές περιοχές σε σχέση με το παράθυρο της οθόνης, έτσι ώστε η κάθε περιοχή να είναι διαμορφωμένη σαν μια διαφορετική συσκευή εισόδου. Για παράδειγμα, μια περιοχή είναι δυνατόν να συμπεριφερθεί σαν τα πλήκτρα, ενώ άλλη να βρίσκεται σε σχετικό τρόπο λειτουργίας και να χρησιμοποιηθεί σαν ιχνόσφαιρα. Αυτή η δυνατότητα μπορεί να είναι χρήσιμη για προτυποποιήσεις συστημάτων.

Τρόποι Εισόδου

Οι πινακίδες αφής είναι ικανές να μεταφέρουν πληροφορίες σε περισσότερες από δύο διαστάσεις. Για παράδειγμα, οι πινακίδες που είναι ευαίσθητες στην πίεση, έχουν αναπτυχθεί για να ανταποκρίνονται διαφορετικά σε διαφορετικές πιέσεις. Έτσι, η ελαφριά πίεση μπορεί να χρησιμοποιηθεί για τον καθορισμό της τοποθεσίας, ενώ η δυνατή πίεση μπορεί να σημαίνει την επιλογή του στόχου. Για εφαρμογές γραφικών διαφορετικοί βαθμοί πίεσης είναι δυνατόν να χρησιμοποιηθούν για να παράγουν γραμμές διαφορετικού πάχους.

Πλεονεκτήματα και Μειονεκτήματα

Η κίνηση που απαιτείται για έλεγχο πινακίδας γραφικών είναι φυσική για πολλούς χρήστες. Επειδή η επιφάνεια της πινακίδας είναι μια ενιαία κατασκευή χωρίς κινητά μέρη, το πλακίδιο είναι δυνατόν να εφαρμοστεί κατάλληλα σε “εχθρικά” περιβάλλοντα στα οποία οι άλλες συσκευές μπορεί και να καταστραφούν. Επίσης, οι πινακίδες έχουν την δυνατότητα να καθαριστούν ευκολότερα από τις οθόνες αφής.

Ο Whitfield το 1983 σε μια εργασία [56] όρισε τέσσερα πλεονεκτήματα των πινακίδων σε σχέση με τις οθόνες αφής. Πρώτον, η οθόνη και η πινακίδα μπορούν να τοποθετηθούν ξεχωριστά σύμφωνα με τις απαιτήσεις του χρήστη. Δεύτερον, τα χέρια του χρήστη δεν καλύπτουν κανένα μέρος της οθόνης. Τρίτον, δεν υπάρχουν προβλήματα παράλλαξης. Τέταρτον, η σκόνη πάνω στην οθόνη δεν επηρεάζει την είσοδο. Επίσης, τα χέρια του χρήστη δεν κουράζονται τόσο πολύ πράγμα που σχετιζόταν με τη διαρκή ανύψωση χεριών προς την οθόνη.

Οι πινακίδες γραφικών παρουσιάζουν και μερικά μειονεκτήματα. Παρόλο που οι πινακίδες κωδικοποιημένων πινάκων και βαθμίδων τάσης τυπικά εξασφαλίζουν υψηλές αναλύσεις, ωστόσο μερικές τεχνολογίες μπορεί να μην εξασφαλίζουν υψηλή ακρίβεια θέσης. Οι πινακίδες γραφικών δεν επιτρέπουν άμεσο συντονισμό ματιού και χεριού, αφού είναι μακριά από την οθόνη. Επειδή οι πινακίδες λειτουργούν με την πίεση του δακτύλου, η διαρκής χρήση είναι δυνατόν να προκαλεί πόνο και κούραση.

Εφαρμογές

Οι πινακίδες γραφικών είναι οι πιο κατάλληλες συσκευές για σκιτσογραφήματα. Ο Parrish το 1982 [44] προσπάθησε να τυποποιήσει τη χρήση του πλακιδίου για όλες τις εφαρμογές σχεδίασης. Οι πινακίδες γραφικών είναι επίσης κατάλληλες για επιλογές αντικειμένων από πίνακες ή μενού, ειδικά όταν χρησιμοποιούνται περιγράμματα. Οι πινακίδες γραφικών έχουν εφαρμογή και στους φορητούς υπολογιστές.

3.3.2. Ποντίκια (Mice)

Τεχνολογίες

Το ποντίκι είναι ένα μικρό κουτί που τοποθετείται κάτω από την παλάμη και τα άκρα των δακτύλων (Σχήμα 22). Η κίνηση του ποντικιού σε επίπεδη επιφάνεια παράγει την κίνηση του δρομέα. Το ποντίκι συνήθως έχει ένα, δύο ή και περισσότερα πλήκτρα, που όταν πιέζονται είναι δυνατόν να εκτελούνται διαφορετικές λειτουργίες, όπως αλλαγές μενού, σχεδίαση γραμμών ή επιβεβαίωση εισόδου. Η κίνηση του ποντικιού ανιχνεύεται είτε μηχανικά είτε οπτικά. Το μηχανικό ποντίκι έχει εφαρμοσμένο στη βάση του ένα σφαιρίδιο. Η περιστροφή της μπάλας αυτής ανιχνεύεται από ποτενσιόμετρο ή από οπτικό αποκωδικοποιητή και χρησιμοποιείται για να προσδιορίζει πληροφορίες λειτουργίας.



Σχήμα 22 : Μηχανικό ποντίκι.

Τα μηχανικά ποντίκια μπορεί να παράγουν θόρυβο κατά την κίνηση και η σκόνη που βρίσκεται πάνω στο τραπέζι είναι δυνατόν να συγκεντρωθεί στο εσωτερικό του ποντικιού. Τα μηχανικά ποντίκια έχουν τη δυνατότητα να λειτουργούν σχεδόν σε κάθε επιφάνεια. Τα οπτικά ποντίκια χρησιμοποιούν οπτικούς ανιχνευτές στη θέση της μπάλας οι οποίοι συνοδεύονται από ένα πινακίδιο. Αυτά τα ποντίκια δεν έχουν κινητά μέρη και έτσι δεν προκαλούν θόρυβο και δεν συγκεντρώνουν σκόνη από την επιφάνεια εργασίας. Τα οπτικά ποντίκια είναι δυνατόν να παρουσιάσουν μικρότερη διακριτική ικανότητα από τα μηχανικά.

Χρήση των πλήκτρων του Ποντικιού

Οι Price και Cordova το 1983 [47] συνέκριναν δύο μεθόδους χρήσης πλήκτρων ποντικιού που υποδεικνύουν κατά πόσον τα μαθηματικά προβλήματα είναι σωστά ή λάθος. Σε όλες τις περιπτώσεις το "σωστό" υποδεικνύεται με απλό πάτημα του πλήκτρου. Στη μία περίπτωση το "λάθος" υποδεικνύεται με απλό πάτημα

διαφορετικού πλήκτρου του ποντικιού, ενώ στην άλλη περίπτωση το "λάθος" εισάγεται με δύο γρήγορα πατήματα του ίδιου πλήκτρου. Η απόδοση ήταν πάντα συνεπής σε όλες τις ανταποκρίσεις που απαιτούν απλά πατήματα πλήκτρων, αλλά δεν ήταν σωστή όταν απαιτούνταν διπλά πατήματα.

Οι Hill, Gunn, Martin και Schwartz το 1991 [25] μελέτησαν τα σχετικά προβλήματα που συναντούν οι δεξιόχειρες χρήστες, όταν εκτελούν διαφορετικές λειτουργίες με ποντίκια τριών πλήκτρων.

Η λειτουργία του απλού πατήματος εκτιμάται ότι είναι ευκολότερη από τη λειτουργία του διπλού πατήματος. Εκτιμήθηκε ότι και οι δύο αυτές λειτουργίες είναι ευκολότερες από τις άλλες λειτουργίες που μελετήθηκαν, όπως πίεση-κράτημα, πίεση-τράβηγμα, τριπλό πάτημα, πάτημα-πίεση-κράτημα και πάτημα-πίεση-τράβηγμα. Οι δύο τελευταίες διακρίνονται σαν οι πιο δύσκολες λειτουργίες. Το αριστερό πλήκτρο εκτιμάται πως είναι ευκολότερο στη χρήση από το κεντρικό πλήκτρο. Και τα δύο αυτά κουμπιά εκτιμώνται ότι είναι ευκολότερα στη χρήση από το δεξί κουμπί. Βέβαια, είναι αντιληπτό ότι οι πολύπλοκοι συνδυασμοί, που απαιτούν ταυτόχρονο πάτημα του αριστερού και του κεντρικού πλήκτρου, είναι ευκολότεροι από τη χρήση δεξιού πλήκτρου.

Πλεονεκτήματα και Μειονεκτήματα

Το ποντίκι έχει γίνει μια πολύ δημοφιλής περιφερειακή συσκευή του υπολογιστή. Το ποντίκι μπορεί να λειτουργήσει σε μικρές περιοχές επειδή είναι δυνατόν να ανασηκωθεί και να ξανατοποθετηθεί. Ο χρήστης μπορεί συνήθως να το χρησιμοποιεί, να εντοπίζει και να κουνά το ποντίκι, ενώ κοιτάζει την οθόνη. Υπάρχουν διαθέσιμα στην αγορά και ασύρματα ποντίκια που χρησιμοποιούν υπέρυθρες ακτίνες.

Επειδή το ποντίκι, εκτός από το πληκτρολόγιο, απαιτεί και για τον εαυτό του μια μικρή, αλλά επιπρόσθετη επιφάνεια, δεν είναι συμβατό με φορητούς υπολογιστές. Το ποντίκι λειτουργεί μόνο με σχετικό τρόπο και έτσι δεν είναι κατάλληλο στις εφαρμογές ζωγραφικής. Επιπλέον, η ζωγραφική με το ποντίκι δεν είναι τόσο φυσική όσο με το στυλό. Απαιτείται επίσης αρκετή πείρα για ουσιαστική χρήση του ποντικιού σε εφαρμογές ζωγραφικής.

Εφαρμογές

Το ποντίκι είναι πιο κατάλληλο σε εφαρμογές ένδειξης και επιλογής. Είναι ικανό να εκτελεί εφαρμογές σχεδίασης, αν και το πλακίδιο αφής παρουσιάζει καλλίτερη απόδοση σε τέτοιες εφαρμογές.

Ποντίκι χωρίς χέρια (No Hands Mouse)

Το ποντίκι χωρίς χέρια είναι μια συσκευή ένδειξης που λειτουργεί αποκλειστικά με τα πόδια. Η συσκευή αυτή δίνει στον χρήστη τη δυνατότητα πλήρους ελέγχου του κέρσορα χωρίς να χρειάζεται να παίρνει τα χέρια του από το πληκτρολόγιο ή το βλέμμα του από την οθόνη. Έτσι ελαχιστοποιείται "ο χρόνος ταξιδιού" που χρειάζεται ο χρήστης για να μετακινεί τα χέρια και το βλέμμα του από το πληκτρολόγιο και την οθόνη αντίστοιχα προς το ποντίκι και πίσω (Σχήμα 23).

Έτσι με τη χρήση αυτής της συσκευής, τα άτομα που έχουν προβλήματα δυσκινησίας στα χέρια μπορούν να έχουν πρόσβαση στον υπολογιστή με αρκετά εύκολο τρόπο.

Το ποντίκι χωρίς χέρια αποτελείται από δύο πετάλια τα οποία προσομοιώνουν τη λειτουργία του κανονικού ποντικιού. Το ένα πετάλιο χρησιμοποιείται για τη μετακίνηση του κέρσορα και το άλλο για την επιλογή (mouse click).



Σχήμα 23 : Ποντίκι χωρίς χέρια.

3.3.3. Ιχνόσφαιρες (Trackballs)

Τεχνολογίες

Η ιχνόσφαιρα αποτελείται από ένα σταθερό πλαίσιο, στο οποίο υπάρχει ενσωματωμένο ένα σφαιρίδιο, που έχει τη δυνατότητα να περιστρέφεται σε οποιαδήποτε κατεύθυνση με τα άκρα των δακτύλων (Σχήμα 24, 25). Η κίνηση της ιχνόσφαιρας ανιχνεύεται με οπτικούς αποκωδικοποιητές. Αυτή η περιστροφή παράγει και την πληροφορία εξόδου που καθορίζει την κίνηση του δρομέα πάνω στην οθόνη.

Η διάμετρος, η αδράνεια περιστροφής και η τριβή της περιστροφής της ιχνόσφαιρας μπορούν να ρυθμισθούν. Το κέρδος της κίνησης στην οθόνη προς την περιστροφή του σφαιριδίου πρέπει επίσης να ορισθεί. Επίσης, η ιχνόσφαιρα, όπως και το ποντίκι, λειτουργεί με σχετικό τρόπο. Απότομες και γρήγορες κινήσεις της μπάλας έχουν ως αποτέλεσμα μεγάλες κινήσεις του δρομέα πάνω στην οθόνη. Έτσι, με την κατάλληλη ρύθμιση είναι δυνατόν να καταλήξουμε σε μια ιδανική περίπτωση της ταχύτητας της κίνησης και της ακρίβειας της τοποθέτησης.

Πλεονεκτήματα και Μειονεκτήματα

Μεγάλες ιχνόσφαιρες σε μια κατάλληλη επιφάνεια εργασίας μπορεί να αποδεχτούν πολύ άνετες για παρατεταμένη χρήση, επειδή ο χρήστης έχει τη δυνατότητα να ξεκουράζει το αντιβράχιο, να κρατά τα χέρια σε μια θέση και να περιστρέφει και να σταματάει τη σφαίρα με τα δάκτυλα. Η ιχνόσφαιρα παρουσιάζει άμεση και αισθητή ανάδραση ως προς την ταχύτητα και την περιστροφή της μπάλας. Απαιτεί μικρό προκαθορισμένο χώρο και έχει τη δυνατότητα να ενωθεί με το πληκτρολόγιο, όπως επίσης μπορεί να λειτουργεί χωρίς να είναι ανάγκη να απομακρύνεται το βλέμμα του χρήστη από την οθόνη.



Σχήμα 24 : Ιχνόσφαιρα.



Σχήμα 25: Ιχνόσφαιρες

Είναι δυνατόν να επισημανθεί ως μειονέκτημα το γεγονός ότι η ιχνόσφαιρα είναι συσκευή περιστροφής και λειτουργεί με σχετικό τρόπο, πράγμα που περιορίζει τη χρησιμότητά της σε εργασίες σχεδίασης. Δεν μπορεί να χρησιμοποιείται στη σχεδίαση ή στην εισαγωγή κειμένου. Είναι επίσης δύσκολο να γίνει εκμετάλλευση πλήκτρων συσχετισμένων με την ιχνόσφαιρα χρησιμοποιώντας τον αντίχειρα στο βαθμό που τα δάκτυλα του ίδιου χεριού, που απαιτούνται για την περιστροφή της μπάλας ή για το σταθερό της κράτημα με τη βοήθεια της αλληλεπίδρασης των μυών και των άκρων, χρειάζονται για να επιτευχθεί ο έλεγχος και αυτών των ξεχωριστών ενεργειών (MacKenzie, Sellen και Buxton, 1991) [36].

Εφαρμογή

Η ιχνόσφαιρα είναι πιο κατάλληλη στις εφαρμογές ένδειξης και επιλογής. Δεδομένου ότι η ιχνόσφαιρα μπορεί κυρίως να χρησιμοποιηθεί για σχεδίαση αφηρημένων και απλών γραμμών και σκιαγραφημάτων, ο Parrish σε

μια εργασία του [44], προτείνει να γίνεται χρήση της ιχνόσφαιρας μόνο σε περίπτωση, που οι γραφικές εργασίες του σχεδίου δεν είναι πολύ αυστηρές.

3.3.4. Χειριστήρια Χειρός (Joysticks)

Τεχνολογίες

Το χειριστήριο αποτελείται από ένα μοχλό που στηρίζεται πάνω σε μια σταθερή βάση. Τα χειριστήρια μετακίνησης ανιχνεύουν την κίνηση του μοχλού με τη βοήθεια ενός ποτενσιόμετρου. Υπάρχει σχέση ανάμεσα στη μετακίνηση του μοχλού και του παραγόμενου σήματος εξόδου. Η μετακίνηση του μοχλού έχει τη δυνατότητα να γίνει προς τους δύο άξονες. Ο μοχλός μπορεί επίσης να έχει ενσωματωμένο ελατήριο για να επιστρέφει στη θέση του όταν αφήνεται ελεύθερος (Σχήμα 26).

Στα χειριστήρια που είναι βασισμένα σε ενεργοποιήσεις διακοπών ή στα δυαδικά χειριστήρια η κίνηση του μοχλού παράγει έξοδο, κλείνοντας έναν ή περισσότερους διακόπτες που είναι τοποθετημένοι ολόγυρα στη βάση. Το σήμα εξόδου το οποίο παράγεται από τη συσκευή κατά την κίνηση της στους δύο άξονες, δεν είναι ανάλογο με την ποσότητα μετακίνησης του μοχλού. Με την κίνηση του μοχλού προς μια κατεύθυνση ο δρομέας πάνω στην οθόνη μετακινείται με σταθερό ρυθμό. Όταν υπάρχει ενσωματωμένο και το ελατήριο, ο μοχλός γυρίζει στην αρχική του θέση (στο κέντρο) και ο δρομέας ακινητοποιείται.

Τα χειριστήρια πίεσης ή τα ισομετρικά χειριστήρια έχουν ένα αλύγιστο μοχλό που δεν κινείται σχεδόν καθόλου προς οποιαδήποτε κατεύθυνση. Για να ανιχνευθεί η κατεύθυνση της κίνησης και η ποσότητας της πίεσης που ασκήθηκε στον μοχλό χρησιμοποιείται ένα όργανο μέτρησης της έντασης. Ο δρομέας κινείται ανάλογα με την ποσότητα της πίεσης που ασκείται στο μοχλό. Όταν ο μοχλός απελευθερώνεται, η έξοδος γίνεται πάλι μηδέν.

Οι Foley και van Dam [16] το 1982 σε μια εργασία τους διατύπωσαν την άποψη ότι τα χειριστήρια είναι πιο σωστό να χρησιμοποιούνται σε εφαρμογές γραφικών σε κατάσταση λειτουργίας ελεγχόμενου ρυθμού, από ότι σε κατάσταση λειτουργίας ελεγχόμενης θέσης.



Σχήμα 26: Χειριστήριο Χειρός.

Σε κατάσταση λειτουργίας ελεγχόμενου ρυθμού η σταθερή μετακίνηση του χειριστηρίου εκτοπίσματος έχει σαν αποτέλεσμα σταθερό ρυθμό μετακίνησης του δρομέα. Αντίθετα σε κατάσταση λειτουργίας ελεγχόμενης θέσης η σταθερή μετακίνηση του χειριστηρίου εκτοπίσματος έχει σαν αποτέλεσμα συγκεκριμένη μετακίνηση του δρομέα.

Με τον ελεγχόμενο ρυθμό η μεγαλύτερη μετακίνηση του χειριστηρίου εκτοπίσματος έχει σαν αποτέλεσμα μεγαλύτερη ταχύτητα του δρομέα, ενώ μειώνοντας τη μετακίνηση (ή τη δύναμη) προς το μηδέν έχει σαν αποτέλεσμα την ακινητοποίηση του δρομέα. Οι Foley και van Dam [16] σημείωσαν την ανάγκη ενσωμάτωσης μιας μικρής αδρανούς ζώνης γύρω από την κεντρική θέση των χειριστηρίων ελεγχόμενου ρυθμού, όπου η ταχύτητα είναι μηδέν, για να μην υπάρξουν προβλήματα μηδενισμού της εισόδου του χειριστηρίου.

Πλεονεκτήματα και Μειονεκτήματα

Ένα προφανές πλεονέκτημα των χειριστηρίων είναι το γεγονός ότι απαιτούν ένα πολύ μικρό καθορισμένο χώρο για να λειτουργήσουν και είναι δυνατόν να κατασκευαστούν σε πολύ μικρό μέγεθος, ώστε να μπορούν να προσαρμοστούν στο πληκτρολόγιο. Επίσης, το χειριστήριο μπορεί να χρησιμοποιηθεί σε εκτεταμένο χρόνο χωρίς να προκαλεί μεγάλη κούραση στα χέρια και τα δάχτυλα. Ωστόσο, το μέτριο μέγεθος πολλών χειριστηρίων αφήνει το κέρδος του χειριστηρίου αρκετά ψηλό, για να επιτυγχάνονται ακριβείς τοποθετήσεις του δρομέα γρήγορα. Το χειριστήριο δεν μπορεί να χρησιμοποιηθεί για ψηφιοποίηση του σχεδίου.

Εφαρμογές

Τα χειριστήρια τείνουν να γίνουν οι πιο κατάλληλες συσκευές στις εφαρμογές συνεχούς κίνησης ή παρακολούθησης, όπου δεν απαιτείται πολύ μεγάλη ακρίβεια. Το μικρό τους μέγεθος και η ευκολία προσαρμογής τους με το πληκτρολόγιο τα έχουν αναγάγει σε δημοφιλείς επιλογές στις εφαρμογές φορητών υπολογιστών.

3.4. Συσκευές Ορίζοντα (Devices on the Horizon)

Όλες οι συσκευές εισόδου που αναλύθηκαν πιο πάνω ανήκουν στη κατηγορία των πολύ διαδομένων συσκευών. Είναι διαθέσιμες εμπορικά σε πλήθος υλοποιήσεων και η μέτρια τιμή τους δικαιολογεί τη μεγάλη χρήση τους σε μεγάλο αριθμό εφαρμογών. Υπάρχει και μια πληθώρα συσκευών που βρίσκονται σε στάδιο ανάπτυξης και οι οποίες, παρόλο που αυτή τη στιγμή χρησιμοποιούνται μόνο σε ειδικές εφαρμογές, είναι πολύ πιθανό να χρησιμοποιηθούν σε ευρύτερες εφαρμογές με την πρόοδο της τεχνολογίας και την ωρίμαση των τεχνικών αλληλεπίδρασης. Μια τέτοια συσκευή είναι ή συσκευή εισόδου ελεγχόμενη με τα μάτια.

3.4.1. Είσοδος ελεγχόμενη με τα μάτια (Eye-controlled Input)

Αρκετές ομάδες ερευνητών έχουν πραγματοποιήσει έρευνες σχετικά με τη χρησιμότητα των κινήσεων των ματιών ως μεθόδου εισόδου δεδομένων στον υπολογιστή. Για παράδειγμα, ο Jacob [27] το 1991 μελέτησε τη χρήση της τεχνολογίας παρακολούθησης ματιού και παρατήρησε την αντανάκλαση της υπέρυθρης ακτινοβολίας από τα μάτια (Σχήμα 27, 28).

Η εικονική γραμμή του βλέμματος υπολογίζεται από τη σχέση ανάμεσα στα δύο σημεία παρακολούθησης, το μάτι και τον ανιχνευτή φωτός. Αυτός ο ανιχνευτής του ματιού τοποθετείται μερικούς πόντους μακρύτερα από τον χρήστη και δεν τον ακουμπάει με κανένα τρόπο.



Σχήμα 27: Έλεγχος με τα Μάτια

Οι κανόνες υλοποίησης τέτοιων συσκευών εισόδου περιέχουν διατάξεις ανάδρασης και μεθόδους επιβεβαίωσης επιλογής. Η ανάδραση μπορεί να είναι συνεχής, όπου το υπολογισμένο σημείο της προσοχής εμφανίζεται συνεχώς πάνω στην οθόνη, ή μπορεί να είναι διακριτή, όπου ο στόχος εμφανίζεται μόνο όταν το βλέμμα είναι κοντά ή πάνω από το στόχο.

Ο Glenn σε μια έρευνά του [19] συνέκρινε τις δύο παραπάνω μεθόδους και κατέληξε στο συμπέρασμα πως δεν υπάρχουν σημαντικές διαφορές ανάμεσα τους, ούτε ως προς την ταχύτητα ούτε ως προς την ακρίβεια επιλογής στόχων. Οι Ware και Mikaelian σε μια εργασία τους [55] πρότειναν πως για την επιλογή αντικειμένων πάνω στην οθόνη είναι πιο δόκιμο να υπάρχει ένας ξεχωριστό διακόπτης, που να χρησιμοποιείται για την επιλογή του στόχου πάνω στο οποίο έχει προσηλωθεί το βλέμμα των ματιών του χρήστη.

Οι συσκευές εισόδου που ελέγχονται με τα μάτια είναι πρόσφορες στις εφαρμογές επιλογής και παρακολούθησης αντικειμένων, επειδή χρησιμοποιούν απλή κίνηση ματιών. Οι συσκευές αυτές μπορούν να ελαττώσουν το φόρτο εργασίας, ελευθερώνουν τα χέρια, τα οποία έτσι μπορούν εκτελέσουν άλλες εργασίες, ελαττώνεται ο χρόνος εκτέλεσης λόγω της μείωσης του χρόνου τοποθέτησης, κίνησης και ελέγχου που απαιτούνται για τη λειτουργία των χειροκίνητων συσκευών εισόδου. Επειδή κάποιος αριθμός ακούσιων κινήσεων των ματιών συνεχίζει να υπάρχει ακόμα και όταν τα μάτια προσπαθούν να σταθεροποιηθούν πάνω σε ένα σημείο, περιστασιακά η γραμμή των ματιών απλώς πλησιάζει κατά προσέγγιση στο σωστό σημείο οπτικής προσοχής.

Η τεχνολογική πολυπλοκότητα και το υψηλό κόστος των υπάρχοντων τεχνολογιών ελέγχου με τα μάτια, τις κάνουν κατάλληλες κυρίως για τις εφαρμογές, όπου η γρήγορη και άμεση επιλογή του στόχου είναι κυρίαρχης σημασίας.

3.4.2. Έλεγχος Κεφαλής (Head Tracking)

Οι συσκευές ελέγχου με τη βοήθεια του κεφαλιού, είναι συστήματα ένδειξης, τα οποία παρέχουν πλήρη έλεγχο του ποντικιού των υπολογιστών στα άτομα που δεν μπορούν να χρησιμοποιήσουν τα χέρια τους, αλλά διαθέτουν πολύ καλό έλεγχο κίνησης κεφαλιού (Σχήμα 28).

Η κάθε κίνηση του κεφαλιού μεταφράζεται σε κίνηση του κέρσορα του ποντικιού πάνω στην οθόνη του υπολογιστή. Το κουτί ελέγχου που έχει ένα σχετικά μικρό μέγεθος, είναι δυνατόν να τοποθετηθεί παντού. Ο αισθητήρας φωτός μπορεί να προσαρμοστεί με πάρα πολλούς τρόπους, για παράδειγμα είναι δυνατόν να προσαρμοστεί στα ακουστικά κεφαλής και στα γυαλιά του χρήστη. Ο μικρός υπέρυθρος δέκτης προσαρμόζεται πάνω ή δίπλα στον υπολογιστή και συνδέεται με τη σειριακή (COM) πόρτα του υπολογιστή.

Με την κίνηση του κεφαλιού ο αισθητήρας του φωτός που είναι προσαρμοσμένος πάνω στο χρήστη, στέλνει ένα σήμα στον υπέρυθρο δέκτη, ο οποίος μεταφράζει την κίνηση αυτή σε κίνηση του ποντικιού.

Η νέα υπέρυθρη τεχνολογία παρέχει απαλή και ήρεμη κίνηση του κέρσορα και πολύ καλή ακρίβεια κίνησης. Αυτό το χαρακτηριστικό της συσκευής την κάνει ιδανική σχεδόν για όλες τις εφαρμογές.



Σχήμα 28 Έλεγχος Κεφαλής

Η προαιρετική χρήση ενός διακόπτη (switch), ο οποίος έχει τη δυνατότητα να ενεργοποιείται με απλό φύσημα ή ρούφηγμα, καθώς και κάποιου εικονικού πληκτρολογίου (virtual keyboard) επιτρέπει δακτυλογράφηση και άλλες λειτουργίες πληκτρολογίου.

3.4.3. Είσοδος με χειρονομίες (Gestural and Spatial Input)

Οι Zimmerman, Lanier, Blanchard, Bryson και Harvill το 1987 [59] κατασκεύασαν μια συσκευή εισόδου σαν γάντι, η οποία έχει τη δυνατότητα να εντοπίζει την θέση, τον προσανατολισμό και την κίνηση των χεριών

(Σχήμα 29). Οι ειδικοί αισθητήρες πάνω στα γάντια υπολογίζουν την απόκλιση των δαχτύλων. Η τοποθεσία και ο προσανατολισμός των χεριών μετριέται είτε με υπερηχητικά κύματα είτε με μαγνητικό αισθητήρα.

Τα γάντια επιτρέπουν στους χρήστες να αλληλεπιδρούν με τα αντικείμενα της οθόνης, όπως συμβαίνει στην πραγματικότητα. Ο χρήστης μπορεί να τα μαζέψει, να τα στρίψει, και να τα πετάξει.



Σχήμα 29: Είσοδος με χειρονομίες

Στην ανίχνευση της θέσης του χεριού, του προσανατολισμού και της κίνησής του, οι συσκευές εισόδου οι βασισμένες στις χειρονομίες προσφέρουν μια φυσική διαδικασία για είσοδο, προσδιορισμό και χειρισμό αντικειμένων, μια διαδικασία που ουσιαστικά εκτελείται κατευθείαν από το χέρι. Η δυνατότητα του χεριού για ταυτόχρονη κίνηση και περιστροφή μπορεί να χρησιμεύσει για ευέλικτη και αποτελεσματική εκτέλεση εφαρμογών που απαιτούν διαχείριση αντικειμένων πάνω στην οθόνη.

3.5. Συμπεράσματα

Η επιλογή της συσκευής εισόδου για μια συγκεκριμένη εφαρμογή περιλαμβάνει τις εξής θεωρήσεις. Πρώτον, πρέπει να καθοριστούν τα χαρακτηριστικά της εφαρμογής και τα χαρακτηριστικά των χρηστών, το περιβάλλον χρήσης και το υπάρχον υλικό. Επίσης, πρέπει να καθοριστούν οι αρχικές και οι μελλοντικές απαιτήσεις της εφαρμογής.

Στη συνέχεια, πρέπει να συγκρίνουμε τα χαρακτηριστικά των υποψήφιων συσκευών εισόδου με τις απαιτήσεις της εφαρμογής, για να μικρύνουμε τη λίστα των υποψήφιων συσκευών. Ο πίνακας 1 ανακεφαλαιώνει τα χαρακτηριστικά και τις σχετικές δυνατότητες των συσκευών εισόδου που παρουσιάστηκαν σε προηγούμενες παραγράφους.

Πίνακας 1. Πλεονεκτήματα και Μειονεκτήματα των κανονικών συσκευών εισόδου.

	Οθόνη Αφής	Στυλό Φωτός	Πλακίδια Γραφικών	Ποντίκια	Ιχνό σφαίρες	Χειριστήρια
Συντονισμός ματιών- χεριών	+	+	0	0	0	0
Απαλλαγή από προβλήματα παράλλαξης	-	-	+	+	+	+
Ελευθέρια στο χώρο εργασίας	-	-	0	0	+	+
Ελάχιστες απαιτήσεις χώρου	+	+	-	-	+	+
Ελάχιστες απαιτήσεις εκμάθησης	+	0	0	0	0	0
Άνεση στην παρατεταμένη χρήση	-	-	0	0	0	0
Δυνατότητα εξομοίωσης άλλων συσκευών	-	-	+	-	-	-
Κατάλληλο για:						
Γρήγορη ένδειξη	+	+	0	+	0	-
Ακριβή ένδειξη	-	-	+	+	+	0
Ζωγραφική	-	0	+	0	-	-
Ανίχνευση	-	-	+	-	-	-
Εισαγωγή αλφαριθμητικών δεδομένων	0	-	0	-	-	-

4. Διακόπτες και Τεχνικές Σάρωσης (Switches and Scanning Techniques)

4.1. Εισαγωγή

Τα άτομα που μπορούν να ελέγχουν με αξιοπιστία μια ή δύο μόνο φυσικές κινήσεις, δεν μπορούν να χρησιμοποιούν άλλες συσκευές εισόδου από τους διακόπτες (Σχήμα 30), μονούς, διπλούς, κλπ.



Σχήμα 30: Διακόπτης

Το μεγαλύτερο μέρος των διακοπών είναι σχεδιασμένο για τα άκρα και για το κεφάλι. Διακόπτες (που συνήθως καλούνται απλοί διακόπτες, επειδή μεταβιβάζουν μονά σήματα στον υπολογιστή) χρησιμοποιούνται για την επιλογή από διαφορετικές προκαθορισμένες επιλογές. Αυτές οι επιλογές παρουσιάζονται στο χρήστη με τη χρήση τεχνικών σάρωσης.

Για παράδειγμα, αν ο χρήστης διαθέτει τρεις επιλογές, η καθεμία από αυτές θα τονιστεί σε κάποιο συγκεκριμένο χρονικό διάστημα. Ο χρήστης επιλέγει πατώντας τον διακόπτη, όταν η επιλογή τονίζεται. Αυτή η μέθοδος μπορεί να χρησιμοποιηθεί για την επιλογή γραμμάτων, αριθμών ή εικόνων στην οθόνη του υπολογιστή. Η μέθοδος αυτή μπορεί να μην είναι και τόσο γρήγορος τρόπος επικοινωνίας αλλά ανοίγει τον έξω κόσμο για ανθρώπους, οι οποίοι δεν έχουν την κινητική ικανότητα να χειρίζονται πληκτρολόγια ή ποντίκια.

Στοιχεία των Διακοπών:

- Ποικιλία χρωμάτων
- Ποικιλία μεγεθών

Τρόποι ενεργοποίησης:

- Διαφορετικοί τρόποι πατήματος του διακόπτη
- Ταλάντωση του διακόπτη
- Πιέζοντας ή έλκοντας τον διακόπτη

- Ηχητικά ή με άνοιγμα και κλείσιμο των ματιών
- Με μικρή κίνηση των μυών
- Σχεδόν πάντα απαιτούν ειδική διεπαφή χρήσης (user interface)
- Μπορούν να χρησιμοποιηθούν για περιβαλλοντικό έλεγχο
- Μπορούν να προσαρμοστούν σε παιχνίδια, αναπηρικά καροτσάκια ή κρεβάτια

4.2. Σύνδεση Διακοπτών με τον Υπολογιστή

Οι διακόπτες συνδέονται με τον υπολογιστή ή με την ειδική συσκευή, ή δια μέσου κάποιας θύρας του υπολογιστή (σειριακή, παράλληλη κτλ) (Σχήμα 31).



Σχήμα 31: Συσκευή Σύνδεσης Διακοπτών

4.3. Παρουσίαση Αντικειμένων Επιλογής

Το ελάχιστο στοιχείο, το οποίο μπορεί να επιλεγεί, ονομάζεται τεμάχιο. Αυτό είναι δυνατόν να παρουσιαστεί με πολλούς και διάφορους τρόπους. Ας πάρουμε για παράδειγμα ένα απλό εικονικό αριθμητικό πληκτρολόγιο, όπου τα τεμάχια είναι απλώς οι αριθμοί από το ένα έως το εννέα, με τη μορφή πλήκτρων.

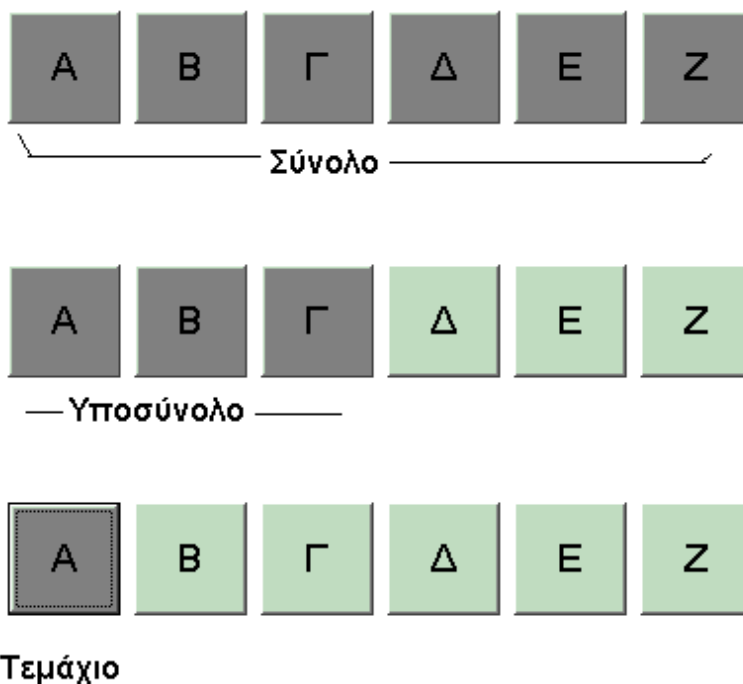
Αυτό μπορεί να παρουσιαστεί στην οθόνη σαν μια γραμμή από πλήκτρα και να χρησιμοποιηθεί μια μέθοδος σάρωσης, για να μπορεί ο χρήστης να επιλέγει από αυτά.

4.3.1. Σύνολο Επιλογών (Selection Set)

Το σύνολο όλων των τεμαχίων, από τα οποία υπάρχει δυνατότητα να γίνει η επιλογή, καλείται Σύνολο Επιλογών. Κάθε στιγμή μπορούμε να παρουσιάσουμε ένα υποσύνολο από το γενικό σύνολο επιλογών. Αφού η επιλογή είναι μια αργή διαδικασία είναι ανώφελο να παρουσιάσουμε περισσότερα τεμάχια από όσα θα χρησιμοποιηθούν (Σχήμα 32).

4.3.2. Τονισμός (Highlighting)

Για να μπορεί ο χρήστης πατώντας το διακόπτη να επιλέγει κάποιο τεμάχιο, πρέπει να υπάρχει στη διάθεσή του κάποια μέθοδος που να του υποδεικνύει ποιο τεμάχιο είναι ενεργοποιημένο τη δεδομένη στιγμή. Ο τονισμός είναι ο κύριος τρόπος, με τον οποίο γίνεται αυτό. Ο οπτικός τονισμός χρησιμοποιείται όταν το τεμάχιο (ή μια ομάδα τεμαχίων) ξεχωρίζει από τα άλλα. Η ενεργοποίηση του διακόπτη εκείνη τη στιγμή θα επιλέξει το τεμάχιο ή των ομάδα τεμαχίων που είναι τονισμένα.



Σχήμα 32: Σύνολο, Υποσύνολο και Τεμάχιο

4.3.3. Ανάδραση (Feedback)

Η ανάδραση είναι σημαντική γιατί υποδεικνύει πότε έχει εντοπιστεί η ενεργοποίηση του διακόπτη ή πότε εκτελείται η πράξη. Η οπτική ανάδραση είναι δυνατόν να αλλάζει κατά τον τονισμό. Είναι γενικά χρήσιμο να περιλαμβάνονται επιλογές για οπτική ή ακουστική ανάδραση.

4.4. Μέθοδοι Σάρωσης (Scanning Methods)

4.4.1. Μικρά σύνολα επιλογών (Small Selection Sets)

Στις περιπτώσεις που έχουμε μικρά σύνολα επιλογών, η μοναδική γραμμή τεμαχίων παρουσιάζεται συνήθως στο κάτω μέρος της οθόνης. Μετά από το πάτημα του διακόπτη αρχίζει η απλή σάρωση περνώντας από το ένα τεμάχιο στο άλλο κατά μήκος της γραμμής από αριστερά προς τα δεξιά.

Η σάρωση θα πρέπει να τονίζεται είτε αλλάζοντας το χρώμα είτε υπογραμμίζοντας το τεμάχιο. Μετά από ένα πέρασμα η σάρωση πρέπει να ξαναξεκινήσει από την αρχή.

4.4.2. Μεγάλα σύνολα επιλογών (Larger Selection Sets)

Τα ευθύγραμμα πλέγματα τεμαχίων σαρώνονται με έναν από τους εξής δύο τρόπους:

1. Η σάρωση γίνεται με παρόμοιο τρόπο, όπως για μικρά σύνολα επιλογών, αλλά με τη σάρωση να περνάει από τη μία γραμμή στην άλλη, από πάνω αριστερά προς κάτω δεξιά.
2. Η σάρωση παίρνει τη μορφή της σάρωσης γραμμής-στήλης με το πρώτο πέρασμα της να γίνεται για όλη τη στήλη (ή τη γραμμή). Μετά από την επιλογή που γίνεται με την ενεργοποίηση του διακόπτη, τα τεμάχια της επιλεγμένης γραμμής σαρώνονται, όπως τα μικρά σύνολα επιλογών. Επίσης, όλη η διαδικασία σάρωσης πρέπει να ξαναρχίσει από την αρχή, όταν γίνεται επιλογή κάποιου τεμαχίου στο τέλος της δεύτερης φάσης.

4.4.3. Πτυσσόμενα μενού

Στις περιπτώσεις που έχουμε πτυσσόμενα μενού προτείνεται η σάρωση να γίνεται πάνω στα τεμάχια των μενού με τονισμό. Όταν επιλεγεί κάποιο μενού, ανοίγει το αντίστοιχο υπό-μενού. Το υπό-μενού τότε σαρώνεται με τον ίδιο τρόπο που έχει γίνει και στα πιο πάνω μενού. Αυτό επαναλαμβάνεται μέχρις ότου επιλεγεί το τελικό τεμάχιο. Είναι σημαντικό να επισημάνουμε ότι ο χρήστης πρέπει να έχει τη δυνατότητα να επιστρέφει ιεραρχικά πίσω στο μενού από οποιαδήποτε σημείο.

4.5. Τεχνικές Σάρωσης (Scanning Technique)

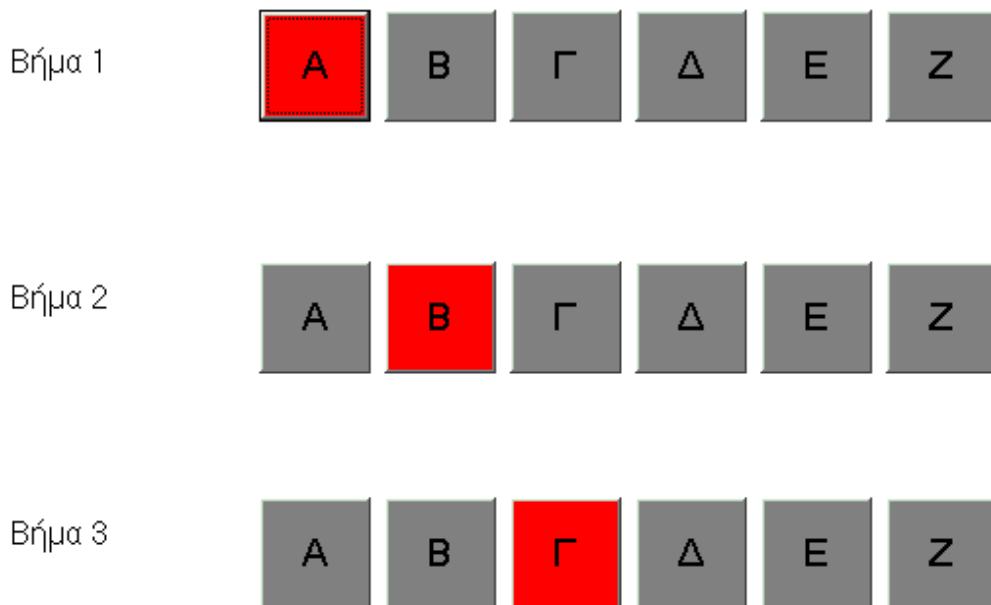
Η σάρωση είναι η διαδικασία κατά την οποία ο τονισμός των τεμαχίων κινείται από το ένα τεμάχιο στο άλλο με μια συγκεκριμένη σειρά. Η σάρωση συνήθως αρχίζει από τα αριστερά και κινείται προς τα δεξιά. Στα πλέγματα αρχίζει από πάνω αριστερά και μετακινείται προς κάτω δεξιά, είτε κατά γραμμή είτε κατά στήλη.

4.5.1. Απλή Σάρωση (Simple Scan)

Σε αυτή την περίπτωση τα τεμάχια τονίζονται ένα-ένα με τη σειρά. Η επιλογή λαμβάνει χώρα όταν το ζητούμενο τεμάχιο τονιστεί. Υπάρχει μια πληθώρα τρόπων μετακίνησης του τονισμού από το ένα τεμάχιο στο άλλο (Σχήμα 33).

4.5.2. Ομαδική Σάρωση (Group Scan)

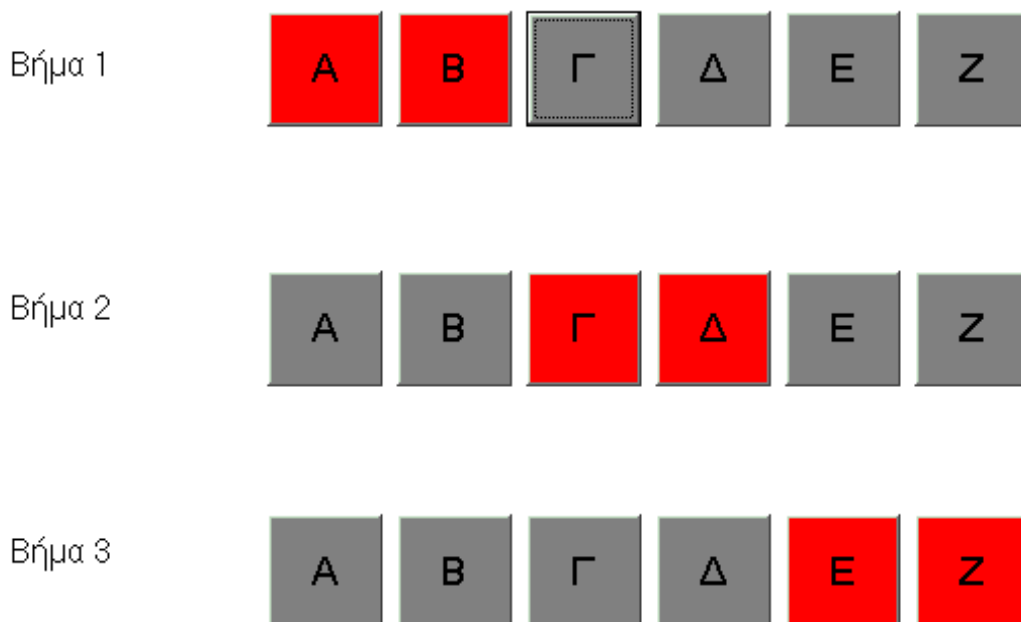
Στην περίπτωση της ομαδικής σάρωσης όλη η επιφάνεια σάρωσης χωρίζεται σε υπό-ομάδες. Κάθε στιγμή τονίζεται μια ολόκληρη υπό-ομάδα. Οι χρήστες στην αρχή επιλέγουν την υπό-ομάδα και μετά σαρώνονται όλα τα τεμάχια της υπό-ομάδας ξεχωριστά. Θεωρητικά είναι δυνατό να έχουμε ιεραρχίες από υπό-ομάδες. Τότε πρώτα επιλέγουμε την υπό-ομάδα και μετά την υπό-ομάδα της υπό-ομάδας και ούτω καθ' εξής, μέχρις ότου φτάσουμε σε επίπεδο με μοναδικό τεμάχιο (Σχήμα 34).



Σχήμα 33: Απλή Σάρωση

4.5.3. Κατευθυνόμενη Σάρωση (Directed Scan)

Η κατευθυνόμενη σάρωση χρησιμοποιείται για να εξομοιώσει την άμεση επιλογή. Η ακολουθία των τονισμών ελέγχεται από το χρήστη με τη λειτουργία του διακόπτη. Η λειτουργίες των απλών ή των διπλών διακοπών είναι δυνατών να περιλαμβάνουν υποδείξεις κατεύθυνσης της σάρωσης και διαδικασίες επιλογής που υλοποιούν τη σάρωση.



Σχήμα 34: Ομαδική Σάρωση

4.5.4. Αυτόματη Σάρωση (Autoscan)

Στην περίπτωση της αυτόματης σάρωσης ο τονισμός μετακινείται αυτόματα από το ένα τεμάχιο ή την υπό-ομάδα στο επόμενο με προκαθορισμένο ρυθμό. Αυτό συνεχίζεται ώσπου ο χρήστης να ενεργοποιήσει το διακόπτη. Τότε επιλέγεται η τρέχουσα υπό-ομάδα ή το τεμάχιο. Ο χρόνος μεταξύ των διαδοχικών μετακινήσεων του τονισμού καλείται Χρόνος Σάρωσης. Η συνήθης κλίμακα χρόνου σάρωσης είναι μεταξύ 0.1 και 10 δευτερόλεπτα με βήμα 0.1 δευτερόλεπτο.

4.5.5. Σάρωση Μετάπτωσης (Toggle Scan)

Σε αυτή την περίπτωση η σάρωση αρχίζει με την πρώτη ενεργοποίηση του διακόπτη και συνεχίζεται σαν αυτόματη σάρωση μέχρι την επόμενη ενεργοποίηση του διακόπτη. Αυτός ο τρόπος σάρωσης είναι απαραίτητος στις περιπτώσεις που οι διακόπτες παρέχουν μόνο στιγμιαία επαφή.

4.5.6. Σάρωση Μοναδικού Βήματος (Single Step Scan)

Στη σάρωση μοναδικού βήματος ο τονισμός κινείται σε κάθε χρονική στιγμή από το ένα τεμάχιο στο άλλο με κάθε ενεργοποίηση του διακόπτη και η επιλογή γίνεται με το δεύτερο διακόπτη. Αυτός ο τύπος σάρωσης δεν είναι κατάλληλος για συστήματα με ένα μόνο διακόπτη.

4.5.7. Επανεκκίνηση Σάρωσης (Restarting Scan)

Η σάρωση μπορεί να ξαναρχίσει, είτε συνεχίζοντας από τη θέση της τελευταίας επιλογής, είτε ξεκινώντας από το αρχικό τεμάχιο. Μπορεί να υπάρχουν ειδικοί λόγοι για την επιλογή του ενός ή του άλλου τρόπου, αλλά ο χρήστης πρέπει να μπορεί να επιλέγει μόνος του τον τρόπο σάρωσης.

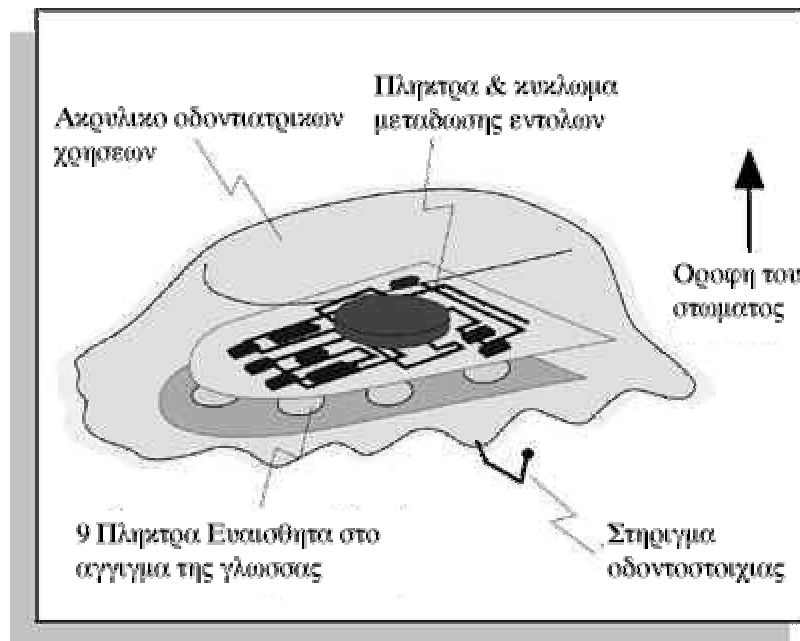
4.6. Διακόπτης με άγγιγμα της γλώσσας (Tongue Touch Keypad)

Οι διακόπτες με άγγιγμα της γλώσσας είναι προσαρμόσιμες συσκευές ελέγχου για άτομα με τετραπληγία, και αποτελούνται από δύο κύρια υποσυστήματα:

1. Μια μοναδική διεπαφή (διακόπτης), ο οποίος δεν έχει κανένα εξωτερικό καλώδιο, τοποθετείται μέσα στο στόμα και δίνει στο χρήστη τη δυνατότητα ελέγχου με τη χρήση της γλώσσας και
2. Ένα σύστημα ενσωμάτωσης ελέγχου (Continuous Integration Control) που μεταφράζει τα σήματα που προέρχονται από το ασύρματο διακόπτη σε εντολές. Με τα σήματα αυτά ο χρήστης μπορεί να ελέγχει μια σειρά από διαφορετικές συσκευές

Ο διακόπτης με άγγιγμα της γλώσσας είναι μια πρωτότυπη συσκευή που εκμεταλλεύεται την ταχύτητα και την ακρίβεια της γλώσσας. Είναι μια συσκευή με εννέα πλήκτρα που συνδέονται με ηλεκτρονικά κυκλώματα τα οποία μεταφράζουν το κάθε άγγιγμα της γλώσσας σε μια εντολή για το σύστημα ελέγχου.

Οι εντολές του διακόπτη μεταφέρονται ασύρματα στον αποδέκτη, ο οποίος τοποθετείται απαρατήρητα πίσω από το κεφάλι του χρήστη (Σχήμα 35)



Σχήμα 35: Διακόπτης με άγγιγμα της γλώσσας

5. Τεχνολογίες Ομιλίας

5.1. Εισαγωγή

Στο κεφάλαιο αυτό επιχειρείται μία σύντομη παρουσίαση των τεχνολογιών ομιλίας όπως αυτές υφίστανται σήμερα. Σκοπός δεν είναι μια εκτεταμένη παρουσίαση αλλά η κάλυψη των διαφορετικών τεχνολογιών και δυνατοτήτων τους προκειμένου να εξετασθούν οι εφαρμογές των τεχνολογιών αυτών σαν τρόπος πρόσβασης στα υπολογιστικά συστήματα. Κατ' αρχάς θα ήταν σκόπιμο να αποσαφηνισθεί ο όρος “τεχνολογίες ομιλίας”, όπως αυτός χρησιμοποιείται στην παρούσα έκθεση. Ο όρος αυτός θα μπορούσε να θεωρηθεί ότι προέρχεται είτε από τον όρο “voice technology” είτε από τον όρο “speech technology”. Πάντως, οι δύο αυτοί όροι έχουν διαφορετική σημασία. Με τον όρο “speech technology” εννοούμε εκείνη την τεχνολογία, η οποία αφορά στην αναγνώριση και στην παραγωγή ομιλίας σε ένα σύστημα. Ο όρος “voice technology” (τεχνολογία φωνής) αναφέρεται στο πώς η τεχνολογία ομιλίας μπορεί να ολοκληρωθεί, για παράδειγμα με ένα περιβάλλον βάσεων δεδομένων. Οι τεχνολογίες ομιλίας μπορούν να ταξινομηθούν σε δύο βασικούς κλάδους, την αναγνώριση και την παραγωγή ομιλίας.

5.2. Αναγνώριση ομιλίας

Γενικά, η ποιότητα ενός συστήματος αναγνώρισης ομιλίας χαρακτηρίζεται από μία σειρά παραμέτρων οι οποίες επηρεάζουν την ακρίβεια αναγνώρισης λέξεων / φράσεων και την ταχύτητα απόκρισης του συστήματος. Η πρώτη παράμετρος είναι η πιθανή αναγνώριση από το σύστημα λέξεων ανεξάρτητα από τον ομιλητή ή σε διαφορετική περίπτωση αν θα πρέπει το σύστημα να εκπαιδευθεί ξεχωριστά για να αναγνωρίζει κάθε έναν από τους ομιλητές (speaker independent / speaker dependent). Είναι φανερό ότι για εξειδικευμένες εφαρμογές, το σύστημα μπορεί να εκπαιδευτεί ξεχωριστά για κάθε ομιλητή. Για παράδειγμα, ο χειρισμός μίας βιντεοκάμερας από έναν αστροναύτη της NASA θα μπορούσε να ανήκει σε αυτή την κατηγορία. Από την άλλη μεριά, σε εφαρμογές κοινής πρόσβασης (public access) είναι επιθυμητό η αναγνώριση να γίνεται ανεξαρτήτως του ομιλητή, ώστε να μην απαιτείται εκπαίδευση. Το πλεονέκτημα των πρώτων συστημάτων όμως είναι ότι προσφέρουν περισσότερη ακρίβεια στην αναγνώριση και ιδιαίτερα σε περιπτώσεις μεγάλων λεξιλογίων. Επίσης, είναι πολύ αποτελεσματικά στο να αντιμετωπίζουν τις διαφορετικές διαλέκτους μιας γλώσσας, καθώς και ομιλητές, που δεν χρησιμοποιούν τη μητρική τους γλώσσα.

Μία δεύτερη σημαντική παράμετρος είναι το μέγεθος του λεξιλογίου που υποστηρίζεται. Στην αναγνώριση ομιλίας το απαιτούμενο λεξιλόγιο μπορεί να αποτελείται από μερικές δεκάδες λέξεις έως και εκατοντάδες χιλιάδες λέξεις. Η τελευταία περίπτωση είναι επιθυμητή σε εφαρμογές υπαγόρευσης (τα σημερινά συστήματα προσφέρουν λεξιλόγια των 250.000-300.000 λέξεων με ενεργό τμήμα τους 50.000 λέξεις). Αν και το ευρύ λεξιλόγιο είναι επιθυμητό σε πολλές εφαρμογές, εν τούτοις συνεπάγεται και ορισμένα μειονεκτήματα. Έτσι, όσο αυξάνεται το μέγεθος του λεξιλογίου, τόσο μειώνεται η ακρίβεια αναγνώρισης των λέξεων αφού αυξάνεται η πιθανότητα ύπαρξης λέξεων που είναι ηχητικά παραπλήσιες. Επίσης, μειώνεται και η ταχύτητα απόκρισης του συστήματος λόγω της εκτενέστερης αναζήτησης που πρέπει να πραγματοποιηθεί μεταξύ των εναλλακτικών εκδοχών-λέξεων που αντιστοιχούν στο σήμα εισόδου προκειμένου να καταλήξει το σύστημα στην σωστή εκδοχή.

Πολύ σημαντικό είναι και το αν η ομιλία είναι διακριτή (discrete) ή συνεχής (continuous), δηλαδή εάν οι λέξεις προφέρονται μεμονωμένα, ή με πιο φυσικό τρόπο ενωμένες η μία με την άλλη όπως στην καθομιλουμένη. Η αναγνώριση συνεχούς ομιλίας μέχρι πρότινος θεωρείτο δυσκολότερη από την διακριτή ομιλία. Στη συνεχή ομιλία οι λέξεις όχι μόνο δεν χωρίζονται από παύσεις αλλά είναι δυνατόν παύσεις να υπάρχουν και μέσα σε μία λέξη. Έχοντας πληροφορία για τα όρια της κάθε λέξης, η αναζήτηση της σωστής εκδοχής απλοποιείται κατά πολύ. Στην συνεχή ομιλία όμως αυτή η πληροφορία δεν είναι διαθέσιμη, και έτσι η εύρεση της σωστής εκδοχής εκτείνεται σε εύρος αρκετών λέξεων, κάνοντας την αναζήτηση αρκετά πολύπλοκη. Πρόσφατες

σχετικά εξελίξεις, όπως ο εντοπισμός λέξεων (word spotting) και το subword modelling [Markowitz 1996] έχουν προσδώσει ιδιαίτερη ευελιξία στην αναγνώριση συνεχούς ομιλίας. Έτσι σιγά-σιγά ο διαχωρισμός αυτός αρχίζει και γίνεται λιγότερο σημαντικός.

Πολύ σημαντικό είναι επίσης το αν η επικοινωνία γίνεται μέσω τηλεφώνου. Στην περίπτωση αυτή, η χαμηλή ποιότητα του σήματος λόγω του περιορισμένου εύρους συχνοτήτων και του περιβάλλοντα θορύβου δυσχεραίνει την αναγνώριση. Πάντως, πρόσφατες εξελίξεις έχουν επιτρέψει την τηλεφωνική αναγνώριση ομιλίας με αρκετή ευρωστία ακόμα και σε θορυβώδη περιβάλλοντα.

Η τεχνολογία αναγνώρισης ομιλίας μπορεί να έχει διαφορετικούς στόχους και επομένως διαφορετικές απαιτήσεις και εφαρμογές. Διακρίνονται τέσσερις βασικοί στόχοι αναγνώρισης:

- υπαγόρευση,
- κατανόηση ομιλίας,
- αναγνώριση ομιλητή, και
- αναγνώριση γλώσσας.

5.2.1. Υπαγόρευση

Ένας λόγος για την χρήση της τεχνολογίας αυτής είναι η υπαγόρευση κειμένων (dictation). Σκοπός της λειτουργίας αυτής είναι ο χρήστης του συστήματος να μπορεί να δημιουργήσει κάποιο κείμενο (όπως ένα ηλεκτρονικό γράμμα ή το ιατρικό αρχείο ενός ασθενή) χωρίς να χρειάζεται να το πληκτρολογήσει, είτε γιατί η πληκτρολόγηση θα προσέφερε μικρότερη ταχύτητα, είτε γιατί η υπαγόρευση θα ήταν εργονομικά σωστότερη για κάποια συγκεκριμένη εφαρμογή. Το βασικό χαρακτηριστικό της υπαγόρευσης είναι ότι το μήνυμα του ομιλητή θα πρέπει να αναγνωρισθεί με ακρίβεια, λέξη προς λέξη.

Στην υπαγόρευση είναι φανερό ότι το λεξιλόγιο θα πρέπει εν γένει να είναι εκτενές ώστε να καλύπτει όλες τις λέξεις που μπορεί να χρησιμοποιήσει ο ομιλητής. Επίσης είναι πολύ σημαντικό να υποστηρίζεται αναγνώριση συνεχούς ομιλίας, αφού η διακριτή κάνει την υπαγόρευση πολύ αφύσικη. Αν και είναι επιθυμητή η αναγνώριση ομιλίας ανεξάρτητα του ομιλητή, αυτή δεν έχει ακόμα επιτευχθεί με ικανοποιητικά μικρό σφάλμα αναγνώρισης. Πάντως, σε αρκετά από τα συστήματα υπαγόρευσης η αναγνώριση ανεξάρτητα του ομιλητή δεν είναι πανάκεια και μεταφράζεται ως μικρός χρόνος εκπαίδευσης του συστήματος (μερικά λεπτά). Σε συστήματα κοινής πρόσβασης, η αναγνώριση ανεξάρτητα του ομιλητή είναι βασική προϋπόθεση.

Οι σημερινές δυνατότητες υπαγόρευσης σε προσωπικούς υπολογιστές είναι αρκετά σημαντικές. Από τα διάφορα εμπορικά πακέτα που κυκλοφορούν, εδώ θα αναφερθούν τέσσερα πολύ γνωστά: το Dragon NaturallySpeaking της Dragon Systems, το IBM ViaVoice Pro Millenium Edition, το Voice Xpress Plus της Lernout & Hauspie, και το FreeSpeech2000 της Philips. Όλα υποστηρίζουν συνεχή ομιλία, η αρχική τους εκπαίδευση είναι πολύ σύντομη και το λεξιλόγιο ευρύ. Τα ποσοστά αναγνώρισης είναι παραπλήσια όπως φαίνεται στον Πίνακα 5.1. Οι όποιες διαφορές δεν είναι πραγματικές γιατί τα δεδομένα έχουν ληφθεί από διαφορετικές πηγές ([H22], [H23], [H24], [H25]).

Συνοψίζοντας, όσον αφορά στις εμπορικές εφαρμογές, η τρέχουσα κατάσταση μας παρέχει δυνατότητες ευρύτατου λεξιλογίου και συνεχούς ομιλίας. Ο μόνος περιορισμός είναι ότι η αναγνώριση δεν πραγματοποιείται ανεξάρτητα του ομιλητή ακόμα. Στον τομέα των πολύ μεγάλων λεξιλογίων και της αναγνώρισης συνεχούς ομιλίας ανεξάρτητα του ομιλητή αναμένεται η ανάπτυξη νέων τεχνικών με χρήση μεθοδολογιών δυναμικού προγραμματισμού (dynamic programming) [Deshmukh 1999], οι οποίες βρίσκονται σε ερευνητικό επίπεδο προς το παρόν. Αυτές οι τεχνικές πολύ πιθανόν σύντομα να υιοθετηθούν και σε εμπορικό επίπεδο.

Σύστημα	Naturally Speaking	ViaVoice	voiceXpress	FreeSpeech 2000
Λεξιλόγιο (αριθμός λέξεων)	250.000	260.000	220.000	300.000
Ποσοστό αναγνώρισης	98%	94%	98%	υψηλό
Χρόνος αρχικής εκπαίδευσης (λεπτά)	5	Σύντομη	12	15
Γλώσσες που υποστηρίζονται	Αγγλικά Γερμανικά Γαλλικά Ισπανικά Ιταλικά Ολλανδικά Αυστραλέζικα Γιαπωνέζικα Σουηδικά	Αγγλικά Γερμανικά Γαλλικά Ισπανικά Ιταλικά Κινέζικα Γιαπωνέζικα	Αγγλικά Γερμανικά Ολλανδικά Γαλλικά Κινέζικα Ισπανικά	Αγγλικά Γερμανικά Αυστριακά Ολλανδικά Γαλλικά Βελγικά Ισπανικά Ιταλικά Πορτογαλλικά Βραζιλιάνικα Σουηδικά

Πίνακας 5.1: Σύγκριση συστημάτων υπαγόρευσης

5.2.2. Κατανόηση ομιλίας

Σε αντίθεση με την υπαγόρευση, η κατανόηση ομιλίας αποσκοπεί στο να εξάγει το νόημα της φράσης του ομιλητή, παρά να αναγνωρίσει κάθε μία από τις λέξεις που την αποτελούν. Αυτό προσφέρει ιδιαίτερα πλεονεκτήματα. Για παράδειγμα, η τεχνική “word spotting” ανιχνεύει συγκεκριμένες λέξεις στην φράση του ομιλητή. Αυτό σημαίνει ότι ο ομιλητής μπορεί να μιλάει με πιο φυσικό τρόπο στο σύστημα, χωρίς να πρέπει να συντάξει τη φράση του με ένα συγκεκριμένο τρόπο, μια που υπάρχουν λέξεις-κλειδιά μέσα στην φράση που πρέπει να αναγνωρισθούν και όχι κάθε μία από τις λέξεις που αποτελούν την φράση. Επιπλέον, το λεξιλόγιο μπορεί να είναι σαφώς μικρότερο αφού αφορά μόνο συγκεκριμένες λέξεις. Επίσης η πιθανότητα να αναγνωρισθεί μία λέξη σωστά είναι μεγαλύτερη από το να αναγνωρισθεί μια ολόκληρη πρόταση σωστά. Σαν αποτέλεσμα, η αναγνώριση ομιλίας μπορεί να γίνει και ανεξάρτητα του ομιλητή, σε αντίθεση με τα συστήματα υπαγόρευσης.

Σε αυτή την κατηγορία ανήκει η αναγνώριση διαλεκτικής ομιλίας. Η ειδοποιός διαφορά με τα συστήματα υπαγόρευσης είναι ότι στην κατανόηση ομιλίας είναι επιθυμητή η κατανόηση του *περιεχομένου* μιας φράσης, δηλαδή του νοήματός της. Για το λόγο αυτό, η αναγνώριση ομιλίας θα πρέπει να ολοκληρωθεί με την επεξεργασία φυσικής ομιλίας (natural language processing), προκειμένου οι φράσεις του ομιλητή να μπορούν να αναγνωρισθούν από τις λέξεις κλειδιά που περιέχουν, *αλλά και* βάσει των συμφραζομένων που έχουν προηγηθεί στον διάλογο. Έτσι, στην περίπτωση της κατανόησης ομιλίας, το σύστημα έχει πρόσβαση σε λεξικό, σύνταξη, σημασιολογία, πραγματολογία, γλωσσολογικό περιβάλλον και προσωδία [Κουρουπέτρογλου 1998].

Όσον αφορά την τρέχουσα κατάσταση σε επίπεδο ακαδημαϊκής έρευνας είναι ενδιαφέρον το παράδειγμα του συστήματος JUPITER [Zue 2000], το οποίο παρέχει πληροφορίες για τον καιρό. Μάλιστα είναι ανεξάρτητο ομιλητή και η ποιότητα της ομιλίας είναι τηλεφωνική. Ακολουθεί ένα απόσπασμα διαλόγου μεταξύ του JUPITER και ενός χρήστη του συστήματος (Πίνακας 5.2) στο οποίο φαίνεται ότι το σύστημα κατανοεί τον διάλογο καθώς επίσης και θυμάται το τι έχει ήδη ειπωθεί προκειμένου να ερμηνεύσει την τρέχουσα εντολή του χρήστη.

Ομιλητής	Yes, I would like the weather forecast for London England, please
JUPITER	In London, in England, Wednesday. Partly cloudy skies with periods of sunshine. High 82 and low 63. Is there something else?
Ομιλητής	What is that in degrees Celsius, please?
JUPITER	In London, in England, Wednesday. High 28 Celsius and low 17 Celsius. What else?

Πίνακας 5.2: Κατανόηση ομιλίας στο σύστημα JUPITER

Το σύστημα JUPITER έχει λεξιλόγιο περίπου 2000 λέξεις. Από τον Μάιο του 1997 έχει δεχθεί περισσότερα από 100.000 τηλεφωνικές κλήσεις, πετυχαίνοντας ακρίβεια 89% για αναγνώριση λέξεων (σε έμπειρους χρήστες περισσότερο από 98%) και σωστή κατανόηση του περιεχομένου των φράσεων 80% για τους αρχάριους χρήστες και πάνω από 95% για τους πεπειραμένους. Είναι ενδεικτικό ότι παρά το ευρύ λεξιλόγιο, την τηλεφωνική ποιότητα ομιλίας, την αναγνώριση ανεξάρτητα του ομιλητή και τη χρήση συνεχούς ομιλίας, το σύστημα πετυχαίνει ποσοστά αναγνώρισης 95% και 98% για περιεχόμενο και λέξεις αντίστοιχα. Βέβαια θα πρέπει να σημειωθεί ότι για non-native ομιλητές τα αποτελέσματα είναι αρκετά χαμηλότερα.

Για την εγγραφή και παροχή πληροφοριών που αφορούν στο συνέδριο της αυτόματης αναγνώρισης και κατανόησης ομιλίας (Workshop on Automatic Speech Recognition and Understanding – ASRU’99 [H26]) υλοποιήθηκε ένα διαλογικό σύστημα που επιτρέπει πρόσβαση σε πληροφορίες για διαμονή, μεταφορά, κοινωνικές δραστηριότητες που αφορούν το συνέδριο. Επίσης, οι συγγραφείς μπορούν να πληροφορηθούν για την κατάσταση των άρθρων που έχουν παραδώσει. Τα μέλη της IEEE μπορούν να εγγραφούν στο συνέδριο λέγοντας απλά τον αριθμό τους στην IEEE.

5.2.3. Αναγνώριση ομιλητή

Η προσωπική αναγνώριση είναι η διαδικασία συσχέτισης ενός συγκεκριμένου ατόμου με κάποια ταυτότητα. Μπορεί να έχει την μορφή εξακρίβωσης (αγγλικοί όροι verification ή authentication) δηλαδή να απαντάει στην ερώτηση “είμαι όντως αυτός που υποστηρίζω ότι είμαι;”, όπως επίσης μπορεί να έχει και την μορφή ταυτοποίησης (recognition ή identification) όπου η ερώτηση που τίθεται στο σύστημα είναι “ποιος είμαι;” [Jain 2000].

Εδώ μας ενδιαφέρει περισσότερο η εξακρίβωση (authentication) ενός ατόμου. Παραδοσιακά αυτή γίνεται με δύο τρόπους, είτε με κάποιο αντικείμενο όπως ταυτότητα ή πιστωτική κάρτα, είτε στηρίζεται σε κάποια γνώση που πρέπει να έχει το άτομο αυτό, όπως ένα σύνθημα ή ο κωδικός (PIN) μιας κάρτας. Όμως, τελευταία έχουν εξελιχθεί διάφορες τεχνικές βιομετρικής αναγνώρισης (biometric identification), οι οποίες στηρίζονται στα φυσικά χαρακτηριστικά και την συμπεριφορά του ατόμου προς αναγνώριση. Επειδή αρκετά από αυτά τα χαρακτηριστικά είναι ιδιαίτερα στον καθένα μας, η βιομετρική αναγνώριση είναι από τη φύση της πιο αξιόπιστη από τις συμβατικές μεθόδους αναγνώρισης. Μία σημαντική βιομετρική τεχνική είναι η αναγνώριση μέσω αποτυπώματος φωνής (voice print). Στο υποκεφάλαιο αυτό, με τον όρο αναγνώριση ομιλητή αναφερόμαστε στην τεχνολογία αυτόματης αναγνώρισης ομιλητή, η οποία κάνει χρήση της φωνής ως εργαλείο-κλειδί για την εξακρίβωση της ταυτότητας ενός ομιλητή. Ο σκοπός δεν είναι η αναγνώριση των λεγομένων του ομιλητή αλλά η εξακρίβωση της ταυτότητάς του.

Τα κύρια μειονεκτήματα της χρησιμοποίησης της φωνής είναι ότι τα φυσικά της χαρακτηριστικά είναι μεταβλητά και τροποποιούνται εύκολα από τα χαρακτηριστικά μετάδοσης και μικροφώνου καθώς και από τον περιβάλλοντα θόρυβο. Εάν όμως ένα σύστημα αποδεχθεί μια μεγάλη διαφοροποίηση στη φωνή του ομιλητή, ίσως αποδεχθεί με τον ίδιο τρόπο τη φωνή ενός διαφορετικού ομιλητή εάν είναι αρκετά όμοια.

Ως αποτέλεσμα των περιορισμών αυτών, τα σημερινά συστήματα συνδυάζουν πολλές φορές την βιομετρική με την παραδοσιακή αναγνώριση. Ένα χαρακτηριστικό παράδειγμα είναι το Sprint Voice FONCARD [H27] που

στηρίζεται σε τεχνολογία της Texas Instruments [H28]) όπου ο ομιλητής προφέρει ένα συνθηματικό που αποτελείται από οικείες λέξεις ή ψηφία. Το περιεχόμενο του συνθηματικού αναγνωρίζεται από ένα speaker independent σύστημα, και παράλληλα ένα δεύτερο σύστημα συγκρίνει το φωνητικό αποτύπωμα του ομιλητή με κάποια προ-αποθηκευμένα.

Ένα σενάριο αναγνώρισης ομιλητή δίνει η Philips με το λογισμικό της SpeechWave [H29]. Την πρώτη φορά που ο ομιλητής καλεί το σύστημα αναγνώρισης, πρέπει να εγγραφεί με αυτό. Θα πρέπει να επαναλάβει τον κωδικό του αρκετές φορές για να μπορέσει να ληφθεί ένα αξιόπιστο αποτύπωμα της φωνής του. Ο παρακάτω διάλογος παρατίθεται ως παράδειγμα από την Philips (Πίνακας 5.3).

Σύστημα	To enrol to the voice banking system, we first have to take your voice print. Please say your six-digit voice banking ID number now.
Ομιλητής	One two three four five six.
Σύστημα	Please, repeat your voice banking ID number now.
Ομιλητής	One two three four five six. (φυσικά η φωνή του ομιλητή διαφέρει αυτή την φορά)
Σύστημα	Your voice print has been taken. Please say your voice banking ID number now.
Ομιλητής	One two three four five six. (φυσικά η φωνή του ομιλητή διαφέρει και αυτή την φορά)
Σύστημα	You are admitted to the system. How may I help you today?

Πίνακας 5.3: Εξακρίβωση χρήστη στο σύστημα SpeechWave

Η εταιρεία Nuance [H30] στο προϊόν της “Nuance Verifier” [H31] προσφέρει τις εξής επιλογές διεπαφής (Πίνακας 5.4):

Τύπος διεπαφής	Προτεινόμενος τρόπος εγγραφής	Προτεινόμενος τρόπος εξακρίβωσης
User defined password	Speak password 2-3 times	Speak password 1-2 times
User name	Speak your name 2-3 times	Speak your name 1-2 times
Account number + random digits	Speak account number + 3-5 sets of random digits	Speak account number plus 1-2 sets of random digits
Random phrases	Speak several sentences of text	Speak 2-3 sets of phrases

Πίνακας 5.4: Τρόποι εγγραφής και εξακρίβωσης του προϊόντος Nuance Verifier

Γενικά οι τιμές των βιομετρικών συστημάτων μπορεί να κυμαίνονται, όπως επίσης και οι δυνατότητές τους. Σύμφωνα με το συγκριτικό τεστ του περιοδικού PC Magazine [H32] υπάρχουν απλά και φθηνά συστήματα όπως το voicecrypt 2.01 της εταιρείας Veritel [H33], το οποίο είναι χρήσιμο για να προσφέρει κάποια επιπλέον προστασία σε αρχεία, αλλά δεν είναι απαραβίαστο. Στο σύστημα αυτό ο χρήστης προφέρει το όνομά του και μετά απαντάει σε μία από τις ερωτήσεις που έχει προκαθορίσει. Πάντως με τις προδιαγραφές εργοστασίου είναι δυνατόν να ξεγελαστεί το σύστημα, ενώ άμα αυξηθεί η ευαισθησία του, τότε δεν μπορεί να ξεγελαστεί αλλά γίνονται ενοχλητικές οι συχνές απορρίψεις του γνήσιου χρήστη. Η τιμή του voicecrypt είναι περίπου \$50 για κάθε θέση αναγνώρισης. Από την άλλη μεριά το προϊόν Citadel GateKeeper της εταιρείας IntelTrak Technologies Inc. [H34] είναι ένα από τα πιο ασφαλή βιομετρικά συστήματα που δοκίμασε το εν λόγω

περιοδικό, και μπορεί να καλύψει μέχρι και 5000 χρήστες. Βέβαια η τιμή του κυμαίνεται μεταξύ \$50000 και \$150000.

5.2.4. Αναγνώριση γλώσσας

Ο σκοπός της αναγνώρισης γλώσσας είναι το σύστημα να συμπεράνει τη γλώσσα του ομιλητή από ένα απόσπασμα ομιλίας. Στον τομέα αυτό γίνονται τελευταία αρκετές πρόοδοι και ιδιαίτερα σε ερευνητικό επίπεδο. Για παράδειγμα, το σύστημα PRLM-P [Zissman 1997] για 12 γλώσσες αναγνώρισε τη γλώσσα σωστά στο 74,3% των περιπτώσεων για αποσπάσματα των 30 δευτερολέπτων, και 53,4% για αποσπάσματα των 10 δευτερολέπτων, ενώ ήταν αρκετά ακριβές (πάνω από 90%) για συγκρίσεις ανά δύο γλώσσες. Επίσης, για 4 γλώσσες (βρετανικά αγγλικά, ισπανικά, γαλλικά και γερμανικά) το ποσοστό λαθών ήταν γύρω στο 10% για 10 δευτερόλεπτα ομιλίας [Corredor-Ardoy 1997]. Όσον αφορά τα τρέχοντα εμπορικά συστήματα, το σύστημα αναγνώρισης γλώσσας της Sanders [H35] έχει τεχνικά χαρακτηριστικά διάκρισης 11 γλωσσών (85.4% στα 10 δευτερόλεπτα και 93.3% στα 50 δευτερόλεπτα).

5.2.5. Ανθρώπινοι παράγοντες στην αναγνώριση ομιλίας

Η αναγνώριση ομιλίας μπορεί να κάνει μια εφαρμογή πιο εύχρηστη, χρήσιμη και φυσική. Παρ' όλα αυτά η τεχνολογία αυτή μόλις τα τελευταία χρόνια έχει αρχίσει να εφαρμόζεται. Τούτο συμβαίνει διότι για μια επιτυχημένη εφαρμογή δεν αρκεί να υπάρχει μόνο η κατάλληλη τεχνολογία που θα την υποστηρίξει, αλλά επιπλέον πρέπει να ληφθούν υπόψη μια σειρά από ανθρώπινους παράγοντες (human factors) που σχετίζονται με αυτήν. Σε πολλές καινοτομικές τεχνολογίες είναι μεγάλος ο πειρασμός να επικεντρωθεί ο σχεδιασμός σε θέματα υλικού και τεχνολογίας λογισμικού και το ζήτημα των ανθρωπίνων παραγόντων να θεωρηθεί δευτερεύον ή εντελώς άσχετο. Κι όμως, ο σωστός σχεδιασμός βασισμένος σε ανθρώπινους παράγοντες, αυξάνει την πιθανότητα επιτυχίας μιας εφαρμογής και μειώνει το κόστος των μετέπειτα τροποποιήσεών της. Οι ανθρώπινοι παράγοντες που αφορούν την αναγνώριση ομιλίας μπορούν να κατηγοριοποιηθούν σχετικά με:

- τις δυνατότητες και τους περιορισμούς της εκάστοτε τεχνολογίας,
- τους στόχους και την δομή του καθήκοντος (task) που πρέπει να υποστηριχθεί, και
- τα χαρακτηριστικά των χρηστών [Markowitz 1996].

Θέματα τεχνολογίας αναγνώρισης ομιλίας

Τα κριτήρια, που σχετίζονται με την τεχνολογία της αναγνώρισης ομιλίας σχετίζονται με τη σωστή αναγνώριση λέξεων (όπως στην περίπτωση των ομοήχων), δηλαδή της πληροφορίας που δεν παρέχεται καθαυτή από το ακουστικό σήμα, με την ταχύτητα απόκρισης του συστήματος (επιθυμητή είναι η αναγνώριση με όσο το δυνατόν μικρότερη χρονική υστέρηση), με την ανάγκη για διακριτή ή συνεχή ομιλία, με τις διακυμάνσεις στην φωνή του κάθε ομιλητή (intra-speaker variability) και μεταξύ των ομιλητών (speaker dependent/independent recognition), με τον περιβάλλοντα θόρυβο και τέλος με την ποιότητα της ομιλίας λόγω της συσκευής (για παράδειγμα, τηλεφωνική ομιλία).

Έχουν συντελεστεί σημαντικές πρόοδοι στα θέματα αυτά, με αποτέλεσμα να διατίθενται σήμερα μια σειρά εμπορικών πακέτων για πλήθος εφαρμογών, καθώς και να αρχίζει να υιοθετείται η τεχνολογία αυτή σε διάφορες υπηρεσίες. Πάντως, θα πρέπει μία εφαρμογή να εξετάζεται και από την πλευρά του αν η υφιστάμενη τεχνολογία μπορεί να την υποστηρίξει ικανοποιητικά. Εάν ισχύουν οι προϋποθέσεις, τότε θα πρέπει να εξετασθεί το κατά πόσο ένα καθήκον θα βελτιωθεί από την εφαρμογή της τεχνολογίας αυτής.

Στόχοι και δομή καθηκόντων

Όσον αφορά τους στόχους ενός καθήκοντος, θα πρέπει να εξετασθεί ένας αριθμός παραμέτρων. Είναι πολύ σημαντικός ο τύπος της πληροφορίας που ο χρήστης προσπαθεί να μεταδώσει στο σύστημα. Συνήθως αυτός πρόκειται για λέξεις ή φράσεις που περιγράφουν εντολές ή τιμές ή κάποιο κείμενο. Η ομιλία μπορεί να

χρησιμοποιηθεί για να περιγράψει γραφική πληροφορία (graphics) μέχρι κάποιο βαθμό, όμως όχι για να μεταδώσει την πληροφορία αυτή καθ' αυτή. Πάντως, τις περισσότερες φορές, ο χρήστης καλείται να δώσει εντολές ή τιμές ή να υπαγορεύσει κάποιο κείμενο, οπότε η ομιλία ενδείκνυται .

Συνυφασμένη με το πρόβλημα της πληροφορίας από το χρήστη στο σύστημα είναι η πληροφορία που πρέπει να δοθεί στο χρήστη από το σύστημα, για τουλάχιστον δύο λόγους. Αφενός, ο χρήστης εμπλέκεται σε έναν διάλογο με το σύστημα, το οποίο σημαίνει ότι οι απαντήσεις του και η συμπεριφορά του εξαρτώνται από τις πληροφορίες που λαμβάνει από το σύστημα. Αφετέρου, πολλές φορές οι χρήστες συμπεριφέρονται σαν να έχουν κάποιον ανθρώπινο συνομιλητή απέναντί τους, κάτι το οποίο επηρεάζει τον τρόπο με τον οποίο απευθύνονται προς το σύστημα. Για παράδειγμα, στα AT&T Bell Laboratories, κατ' αρχάς δοκιμάστηκε μία λαμπερή φωνή για τις προτροπές του συστήματος, βάσει του σκεπτικού ότι είναι πιο κοντά προς την ανθρώπινη επικοινωνία και εν γένει πιο καλαίσθητη. Ο τόνος της φωνής όμως περισσότερο προέτρεπε τους χρήστες να συζητήσουν με το σύστημα παρά να λένε τις λέξεις κλειδιά. Γρήγορα αυτή αντικαταστάθηκε από μία λιγότερο φιλική φωνή [Markowitz 1996].

Οι επιλογές που έχει ο χρήστης σε κάποιο σημείο της πορείας της αλληλεπίδρασής του με το σύστημα είναι ένας ακόμα σημαντικός παράγοντας. Πώς μπορεί να προσπελάσει ο χρήστης αυτές τις επιλογές; Πώς μπορεί για παράδειγμα να διαφοροποιήσει τα δεδομένα-τιμές που έχει να δώσει από τις εντολές μετακίνησης ανάμεσα στα πεδία μιας ηλεκτρονικής φόρμας;

Παρά τα υψηλά ποσοστά επιτυχούς αναγνώρισης, λάθη μπορούν να συμβούν σχετικά συχνά. Δύο είναι τα σχετικά ερωτήματα. Κατ' αρχάς, πώς και πότε ο χρήστης καταλαβαίνει ότι έχει συμβεί ένα λάθος στην αναγνώριση. Δεύτερον, όταν το λάθος γίνει αντιληπτό, πώς μπορεί να διορθωθεί; Αυτό εξαρτάται από τον τύπο του λάθους (παράγραφος 5.2.7).

Επίσης, εργονομικοί παράγοντες που αφορούν τη σωστή τοποθέτηση της συσκευής καταγραφής ομιλίας ως προς τον χρήστη. Για παράδειγμα δεν μπορούμε να ζητήσουμε από ένα χρήστη να κρατάει προσεκτικά ένα μικρόφωνο την ώρα που οδηγεί. Επίσης, σε ένα εργοστάσιο το προστατευτικό κράνος μπορεί να εμποδίζει τη σωστή τοποθέτηση του μικροφώνου. Σε ένα κιόσκι πληροφοριών ο χρήστης μπορεί διαρκώς να αλλάζει τη θέση του και την απόσταση του από το μικρόφωνο.

Όλοι αυτοί οι παράγοντες θα πρέπει να ληφθούν υπόψη, προκειμένου να αποφασιστεί η καταλληλότητα της αναγνώρισης ομιλίας για ένα συγκεκριμένο καθήκον ή η χρήση κάποιου διαφορετικού τρόπου εισαγωγής (input modality) το οποίο πιθανώς να παρέχει το σύστημα, όπως η πληκτρολόγηση. Δηλαδή η χρήση της τεχνολογίας αναγνώρισης ομιλίας δεν είναι πανάκεια για κάποιο καθήκον.

Οι χρήστες του συστήματος

Μία σημαντική διαφοροποίηση είναι μεταξύ άπειρων και έμπειρων χρηστών, οι οποίοι θα πρέπει να αντιμετωπιστούν διαφορετικά. Η εμπειρία ενός χρήστη στην εκτέλεση ενός καθήκοντος σημαίνει ότι το λεξιλόγιο που θα χρησιμοποιήσει μπορεί να είναι εξειδικευμένο. Εάν όμως οι χρήστες δεν επιθυμούν να εξειδικευθούν στο συγκεκριμένο καθήκον, τότε δεν θα πρέπει να χρησιμοποιείται η γλώσσα των ειδικών. Πρέπει επίσης να ληφθούν υπόψη οι δεξιότητες (και τα τυχόν προβλήματα στην όραση ή την κίνηση) των ομάδων χρηστών στους οποίους απευθύνεται το σύστημα. Επιπλέον, θα πρέπει να γίνει διαχωρισμός μεταξύ των τακτικών και των περιστασιακών χρηστών. Στην πρώτη περίπτωση θα πρέπει να παρέχονται τεχνικές “power user” όπως η δυνατότητα να ανταποκριθεί ο χρήστης στο σύστημα διακόπτοντάς το (barge in). Για τους περιστασιακούς χρήστες, είναι καλό το σύστημα να τους δίνει πιο εμπειρισταωμένες προτροπές για να τους βοηθάει στις απαντήσεις τους.

Σημασία έχει επίσης η πλατφόρμα που θα χρησιμοποιηθεί. Όπως αναφέρεται στο επόμενο κεφάλαιο, για τους σκοπούς αυτής της έκθεσης, οι πλατφόρμες που ενδιαφέρουν είναι το τηλέφωνο (σταθερό/κινητό) κατά κύριο λόγο, το τυπικό τερματικό, το κιόσκι πληροφοριών, και διάφορα εναλλακτικά interfaces για extra-ordinary χρήστες.

Είναι σημαντικό ότι ένα σύστημα με αναγνώριση ομιλίας θα χρησιμοποιηθεί για την εκπλήρωση ενός ή περισσοτέρων καθηκόντων. Όσον αφορά το χρήστη, σημασία έχουν τα καθήκοντα καθαυτά και όχι το πώς εισάγονται οι εντολές ή δεδομένα στο σύστημα. Δηλαδή, σε ένα επιτυχημένο σύστημα αναγνώρισης ομιλίας, ο χρήστης δεν θα χρειάζεται να στρέψει την προσοχή του στην ομιλία αλλά στο καθήκον που πρέπει να επιτελεστεί.

5.2.6. Τύποι εφαρμογών της αναγνώρισης ομιλίας

Διακρίνουμε τους εξής διαφορετικούς τύπους εφαρμογών:

- εντολών-και-ελέγχου (command-and-control),
- εισαγωγής δεδομένων (data entry),
- πρόσβασης δεδομένων (data access), και
- υπαγόρευσης (dictation).

Στο [Markowitz 1996] διακρίνεται και η ομιλία από το τηλέφωνο, αλλά όσον αφορά τη συγκεκριμένη έκθεση οποιαδήποτε από τις τέσσερις προαναφερθείσες κατηγορίες μπορεί να γίνεται είτε τηλεφωνικά είτε από κάποιο κανονικό τερματικό ή, ίσως, από κάποιο κίосκι πληροφοριών.

Οι εφαρμογές εντολών-και-ελέγχου σχετίζονται εν γένει με την κατανόηση ομιλίας. Το λεξιλόγιο μπορεί να είναι περιορισμένο ή εκτενές, ανάλογα την εφαρμογή. Ειδικά όσον αφορά στις τηλεφωνικές εφαρμογές (και όχι αυτές που απευθύνονται σε λίγους εξειδικευμένους χρήστες) χρησιμοποιούνται τεχνικές ανίχνευσης λέξεων-κλειδιών (keyword spotting) αντί για γραμματικές πεπερασμένων αυτόματων (finite-state grammars), διότι αυτό επιτρέπει περισσότερη ελευθερία και αμεσότητα στον ομιλητή. Επιπλέον, οι εφαρμογές αυτές χρησιμοποιούν speaker-independent τεχνολογία για να περιορίσουν ή να εξαλείψουν την ανάγκη εκπαίδευσης σε κάθε χρήστη. Γενικά για τους χρήστες αυτούς προτιμάται αναγνώριση συνεχούς ομιλίας (αλλά βέβαια αναγνώριση διακριτών λέξεων-κλειδιών).

Η αναγνώριση ομιλίας είναι ίσως ο μόνος τρόπος εισόδου που επιτρέπει στους χρήστες να κρατήσουν το βλέμμα τους, τα χέρια τους και, εν γένει, την σκέψη τους σε αυτό που κάνουν ενώ παράλληλα εισάγουν δεδομένα. Εδώ όμως θα πρέπει να τονίσουμε, ότι το σύστημα δεν απασχολεί την σκέψη του ομιλητή μόνο όταν ο τελευταίος μπορεί να απευθύνεται σε αυτό με φυσιολογική φωνή και υπάρχει μικρό ποσοστό λαθών (για ένα παράδειγμα με το MATHS workstation και ανάλυση βάσει ψυχολογικών μοντέλων αναφερθείτε στο [Edwards 1999]). Στην εισαγωγή δεδομένων το λεξιλόγιο είναι περιορισμένο τις περισσότερες φορές, οπότε είναι αρκετά πιθανόν ο παραπάνω στόχος να επιτευχθεί.

Η πρόσβαση δεδομένων με φωνή παρέχει τη δυνατότητα στους χρήστες να ανακτήσουν επιθυμητές πληροφορίες οποιαδήποτε στιγμή. Οι οδηγίες του χρήστη αναγνωρίζονται και κατόπιν μεταφράζονται στη γλώσσα της βάσεως δεδομένων. Εάν το σύστημα είναι περιορισμένης χρήσης τότε μπορούν να χρησιμοποιηθούν γραμματικές πεπερασμένων αυτομάτων. Εν γένει όμως, αυτά τα συστήματα προορίζονται για κοινή χρήση από το τηλέφωνο οπότε και υιοθετούν τεχνικές word spotting. Αυτά τα συστήματα πρέπει επίσης να είναι speaker independent και ανθεκτικά στον περιβάλλοντα θόρυβο. Παραδείγματα αποτελούν η πρόσβαση ενός τραπεζικού λογαριασμού τηλεφωνικά, οι τιμές των μετοχών στο χρηματιστήριο, κλπ.

Στην υπαγόρευση, η ομιλία του χρήστη μεταφράζεται σε κείμενο. Θεωρητικά, το λεξιλόγιο πρέπει να είναι εκτενές. Πιο σημαντικό όμως είναι το λεξιλόγιο να περιέχει τις λέξεις που χρησιμοποιούνται στο κείμενο, έτσι ώστε αυτές να μην αναγνωρίζονται λανθασμένα. Οπωσδήποτε θα πρέπει να χρησιμοποιείται συνεχής ομιλία. Θα πρέπει επίσης να παρέχονται εργαλεία διόρθωσης, προσθήκης και πλοήγησης στο κείμενο.

5.2.7. Τύποι λαθών αναγνώρισης

Οι τύποι των λαθών που μπορούν να συμβούν κατά την αναγνώριση ομιλίας είναι πολύ σημαντικοί διότι σε συνδυασμό με το προς εκτέλεση καθήκον επηρεάζουν το πότε ο χρήστης αντιλαμβάνεται κάποιο λάθος αναγνώρισης και το ποιες ενέργειες πρέπει να αναλάβει για να διορθώσει το λάθος αυτό. Τρεις βασικές κατηγορίες λαθών είναι η διαγραφή (deletion), η υποκατάσταση (substitution) και η εισαγωγή (insertion). Για παράδειγμα, εάν ο χρήστης πει “1234” το σύστημα θα καταλάβει “124”, “1235”, και “12384” αντίστοιχα για κάθε τύπο λάθους.

Επίσης, σημαντικά λάθη είναι ο διαχωρισμός (split) και η ένωση (merge) λέξεων, και η αποδοχή μιας λέξης που δεν ανήκει στο λεξιλόγιο (false acceptance). Στις δύο πρώτες περιπτώσεις παράδειγμα είναι η κατανόηση της αγγλικής λέξης “randemonium” ως “point and mention” και της φράσης “lamp and” ως “lamproon”. Σε συστήματα που στηρίζονται σε word spotting, υπάρχουν δύο ειδικότερες περιπτώσεις όσον αφορά την αποδοχή μιας λέξης εκτός λεξιλογίου. Το σύστημα μπορεί να ανιχνεύσει σε ένα τμήμα εισόδου μια λέξη η οποία δεν υπάρχει στην πραγματικότητα, όπως επίσης, μπορεί να αποτύχει να ανιχνεύσει μια λέξη η οποία έχει πραγματικά ειπωθεί.

5.2.8. Συστήματα αναγνώρισης ελληνικής ομιλίας

Το Ινστιτούτο Επεξεργασίας Λόγου (ΙΕΛ) εκπόνησε το 1997 ερευνητικό έργο με τίτλο “τηλεφωνική πρόσβαση βάσεων πληροφοριών με αναγνώριση φωνής” [H36]. Μακροπρόθεσμα θα αναπτυχθεί ένα πιλοτικό σύστημα μιας βάσης πληροφοριών (π.χ. διευθύνσεις και τηλέφωνα μουσείων, οργανισμών κλπ.), όπου θα έχει πρόσβαση ο χρήστης μέσω ενός περιορισμένου αριθμού λέξεων. Το σύστημα μπορεί να επεκταθεί σε πραγματικά συστήματα βάσεων δεδομένων με μεγάλο όγκο πληροφοριών, όπου για την προσπέλαση θα απαιτείται ένα ευρύτερο λεξιλόγιο. Επίσης, στα πλαίσια του έργου “ΕΥΛΟΓΟΝ” αναπτύχθηκε σύστημα το οποίο δίνει πληροφορίες για γεωγραφικά δεδομένα της Ελλάδος με φωνή και οπτική απεικόνιση. Η πρόσβαση στο σύστημα από το χρήστη γίνεται μόνο με φωνή χωρίς χρήση πληκτρολογίου.

Το Πολυτεχνείο Κρήτης είναι ανάδοχος του έργου “Λογοτυπογραφία – Δημιουργία Ελληνικού Συστήματος Υπαγόρευσης και εφαρμογή του στη Βελτίωση Ροής Εργασίας στο Δημοσιογραφικό Χώρο” εγκεκριμένο από τη ΓΓΕΤ (ΕΠΕΤ II, 1999-) και με συνεργαζόμενους φορείς το ΙΕΛ και την εταιρεία Χ.Κ. Τεγόπουλος Εκδόσεις ΑΕ. Το έργο αποσκοπεί στη δημιουργία υποδομής για την αναγνώριση συνεχούς λόγου στην Ελληνική γλώσσα, την προσαρμογή ενός επιτυχημένου συστήματος αναγνώρισης συνεχούς ομιλίας μεγάλου λεξιλογίου στην ελληνική και την εγκατάστασή του σε εκδοτικό περιβάλλον, την έρευνα και ανάπτυξη συστημάτων μεγάλης αναγνωρισιμότητας και την προετοιμασία του εδάφους για τη δημιουργία ενός 100% ελληνικού συστήματος αναγνώρισης φωνής. Περαιτέρω δραστηριότητες της Ομάδας Αναγνώρισης Φωνής στον Τομέα Τηλεπικοινωνιών του Πολυτεχνείου Κρήτης περιλαμβάνουν τη δημιουργία τηλεφωνητή ρομπότ όπου ο ομιλητής αναφέρει με φυσική ομιλία το όνομα της υπηρεσίας ή του ατόμου με το οποίο θέλει να συνδεθεί τηλεφωνικά (δημοσίευμα εφημερίδας Ελευθεροτυπία, 22/4/99, σελ. 60) όπως επίσης και το χειρισμό ενός ιατρικού ρομπότ με ομιλία (σελ. 49), στο οποίο αναπτύχθηκε εφαρμογή ανοικτής ακρόασης καθοδήγησης λαπαροσκοπικού ρομποτικού βραχίονα με φυσική ομιλία στην ελληνική γλώσσα ανεξάρτητου ομιλητή και ταυτόχρονοι μηχανισμοί απόρριψης ανεπιθύμητων ήχων.

Στο Τμήμα Πληροφορικής του Πανεπιστημίου Αθηνών ολοκληρώθηκε ερευνητικό έργο που είχε στόχο την δημιουργία ενός συστήματος πρόσβασης σε βάσεις δεδομένων με αναγνώριση ομιλίας (speaker independent, διακριτή ομιλία). Τα δεδομένα που ανακτώνται είναι τα τηλέφωνα των μελών του Τμήματος Πληροφορικής του Πανεπιστημίου Αθηνών. Το περιβάλλον που χρησιμοποιείται για την αναγνώριση ομιλίας είναι το HTK της εταιρείας Entropic. Η πρόσβαση γίνεται απομακρυσμένα μέσω τηλεφώνου. Στο πλαίσιο ενός νέου ερευνητικού έργου του ΕΠΕΤ II (2000-) με τίτλο ΥΓΙΟΡΟΜΠΟΤ στο ίδιο Τμήμα Πληροφορικής θα μελετηθεί η χρήση του περιβάλλοντος HTK ή του περιβάλλοντος CSLU [H37] με τελικό στόχο την αναγνώριση συνεχούς ομιλίας στην επικοινωνία ανθρώπου με ένα ρομπότ.

Στο Ε.Μ.Π. λαμβάνει χώρα το πρόγραμμα “ΑΚΟΔΙΦΩΝ” (ΓΓΕΤ, ΕΠΕΤ ΙΙ, 2000-) που αφορά την βελτίωση της φωνιμικής αναγνωρισιμότητας με ανάπτυξη και χρησιμοποίηση ακουστικών παραμέτρων διαμόρφωσης και φράκταλς).

Η Ομάδα Γλωσσικής Τεχνολογίας του Πανεπιστημίου Πατρών [H38] έχει μακρά παράδοση στις ερευνητικές περιοχές της αναγνώρισης ομιλίας, αναγνώρισης και επιβεβαίωσης ομιλητή καθώς και στη μοντελοποίηση της γλώσσας ως μέσο βελτίωσης της απόδοσης της αναγνώρισης ομιλίας. Έχει αναπτύξει τις εξής εφαρμογές:

- Ένα σύστημα αναγνώρισης ομιλητή ανεξάρτητο κειμένου και γλώσσας με βάση νευρωνικά δίκτυα. Ένα σύνολο εργαλείων κατάλληλων για σχολιασμό ομιλίας.
- Ένα διαλογικό σύστημα για την αυτοματοποίηση των υπηρεσιών τηλεφωνικού κέντρου, που αυτοματοποιεί την συλλογή δεδομένων για εταιρείες ασφάλισης αυτοκινήτων.
- Ένα Διαλογικό σύστημα για την αυτοματοποίηση των τηλεφωνικών υπηρεσιών καταλόγου.

Παρατηρούμε λοιπόν ότι σε ερευνητικό επίπεδο υπάρχει έντονη προσπάθεια σε όλους τους τύπους εφαρμογών αναγνώρισης ομιλίας (υπαγόρευσης, πρόσβασης δεδομένων, εντολών-και-ελέγχου). Όμως υπάρχουν λύσεις και σε εμπορικό επίπεδο. Για παράδειγμα, η εταιρεία Inter Engineering Computer Security [H39] προωθεί το σύστημα BioID της εταιρείας Dialog Communication Systems [H40], που αναγνωρίζει αυτόματα ένα άτομο βασιζόμενο σε μοναδικά χαρακτηριστικά του προσώπου, της κίνησης των χειλιών και τη φωνή. Επίσης, ο Τομέας Συστημάτων Γλωσσικής Τεχνολογίας της εταιρείας Knowledge A.E. [H41] έχει να επιδείξει μια σειρά προϊόντων όπως το SESAMI (εξακρίβωση ομιλητή ανεξάρτητη κειμένου μέσω τηλεφώνου) και το ORPHEUS (διακριτής ομιλίας, μεσαίου λεξιλογίου, ανεξάρτητου ομιλητή, υψηλής ακρίβειας, μέσω τηλεφώνου). Επίσης, θα πρέπει να τονισθεί ότι αρκετά από τα ερευνητικά έργα που υποστηρίζονται από τη ΓΓΕΤ αποσκοπούν τελικά στην ανάπτυξη προϊόντων ή τεχνογνωσίας απαραίτητης για αυτά. Οπότε, σύντομα αναμένεται να υπάρξουν ακόμη περισσότερα εμπορικά προϊόντα που θα εστιάζονται στην αναγνώριση ελληνικής γλώσσας.

5.3. Παραγωγή ομιλίας

Η παραγωγή ομιλίας μπορεί να γίνει με τρεις διαφορετικές μεθόδους [Κουρουπέτρογλου 1998]:

- κωδικοποίηση κυματομορφής (waveform coding),
- ανάλυση-σύνθεση (analysis-synthesis), και
- σύνθεση με κανόνες (synthesis by rules).

Ο Πίνακας 5.5 [Κουρουπέτρογλου 1998] προσφέρει μια σύγκριση αυτών των μεθόδων.

Γνωρίσματα	Κωδικοποίηση κυματομορφής	Ανάλυση-σύνθεση	Σύνθεση με κανόνα
Ποιότητα			
{κατανοητότητα	Υψηλή	Υψηλή	Υψηλή
{φυσικότητα	Υψηλή	Μέτρια	Μέτρια
Μέγεθος λεξιλογίου	Μικρό(<500)	Μεγάλο(1000ς)	Απεριόριστο
Ρυθμός bit	24-64 kbps	2,4-9,6 kbps	50-75 bps
Μήκος ομιλίας που αποθηκεύεται σε 1 Mbit μνήμης	15-40s	10s-7min	Απεριόριστο
Αποθηκευόμενες μονάδες	Λέξεις, Συλλαβές, Προτάσεις	Λέξεις, Συλλαβές, Προτάσεις	Φωνήματα, Συλλαβές, κλπ.

Πολυπλοκότητα	Χαμηλή	Μέτρια	Υψηλή
Κυρίως υλικό	Μνήμη	Μνήμη και Επεξεργαστής	Επεξεργαστής

Πίνακας 5.5: Μερικά χαρακτηριστικά των διαφορετικών μεθόδων σύνθεσης ομιλίας

5.3.1. Κωδικοποίηση κυματομορφής

Με την μέθοδο αυτή, ενότητες όπως λέξεις ή φράσεις της ανθρώπινης φωνής κατ' αρχάς ηχογραφούνται και αποθηκεύονται ψηφιακά. Κατόπιν, οι επιθυμητές προτάσεις ομιλίας συντίθενται εντοπίζοντας και συρράφοντας τις κατάλληλες ενότητες. Αν και η μέθοδος αυτή είναι χαμηλής πολυπλοκότητας, εν τούτοις η παραγωγή ομιλίας με φυσικότητα απαιτεί ιδιαίτερη προσοχή στις ηχογραφήσεις για τους λόγους που αναφέρονται στις επόμενες παραγράφους.

Μία βασική συνιστώσα είναι το μέγεθος των ενοτήτων που χρησιμοποιούνται. Θεωρητικά μπορούν να κυμαίνονται από φωνήματα έως ολόκληρες φράσεις. Το μέγεθος των ενοτήτων επηρεάζει άμεσα την ποιότητα της παραγόμενης ομιλίας και την ευελιξία του συστήματος παραγωγής. Εάν οι ενότητες είναι μεγάλες (προτάσεις ή φράσεις) τότε η ποιότητα ομιλίας είναι καλή, όμως ο αριθμός των διαφορετικών φράσεων που μπορούν να παραχθούν είναι πολύ περιορισμένος. Στην περίπτωση χρήσης πολύ μικρών ενοτήτων, όπως συλλαβές ή φωνήματα, μπορεί να συντεθεί μια μεγάλη κλίμακα λέξεων και φράσεων, αλλά η ποιότητα ομιλίας εκφυλίζεται κατά πολύ.

Στην ενδιάμεση περίπτωση χρήσης λέξεων ως ενότητες, υπάρχουν διάφορα προβλήματα [Κουρουπέτρογλου 1998]. Πρώτον, η φράση που εκφωνείται από κάποιον ομιλητή είναι πολύ διαφορετική από την εκφώνηση των ανεξάρτητων λέξεων που την αποτελούν. Σε μια πρόταση οι λέξεις έχουν την μισή διάρκεια από ότι όταν εκφωνούνται μεμονωμένα, γεγονός που κάνει την ομιλία που βασίζεται σε ενότητες λέξεων να μοιάζει επώδυνα αργή. Δεύτερον, ο επιτονισμός (intonation) μιας πρότασης και η προσωδία (prosody) , που εξαρτώνται από συντακτικούς, σημασιολογικούς και πραγματολογικούς παράγοντες, δείχνουν ιδιαίτερα αφύσικοι στην απλή αλληλουχία λέξεων ακόμα και αν είναι αποθηκευμένες αρκετές διαφοροποιήσεις της ίδιας λέξης.

Για τους λόγους αυτούς, σε πρακτικά συστήματα χρησιμοποιείται ένας συνδυασμός φράσεων καθώς και μεμονωμένων λέξεων και φράσεων για τις οποίες αποθηκεύονται πολλές διαφοροποιήσεις. Μία σειρά επιδείξεων τέτοιων συστημάτων μπορούν να βρεθούν στο [H42]. Ο λόγος για την επιλογή της σύνθεσης με κυματομορφή από τα συστήματα αυτά είναι η υψηλή ποιότητα ομιλίας που τελικά προσφέρουν συγκριτικά με τις άλλες μεθόδους. Από την άλλη μεριά όμως έχουν και ένα πολύ σημαντικό μειονέκτημα, ότι δηλαδή δεν μπορούν να μετατρέψουν οποιοδήποτε κείμενο σε ομιλία, μια που καλύπτουν μόνο συγκεκριμένες λέξεις ή φράσεις. Τούτο σημαίνει ότι, αν (ως τυπικό παράδειγμα) σε ένα σύστημα χρησιμοποιούνται συγκεκριμένες προτροπές (prompts) και ίσως και κάποια ψηφία, τότε η κωδικοποίηση κυματομορφής προσφέρεται. Αν όμως το σύστημα χρησιμοποιείται για να διαβάσει τα εισερχόμενα μηνύματα σε ηλεκτρονικό ταχυδρομείο, τότε η μέθοδος αυτή δεν είναι κατάλληλη, αφού το κείμενο μπορεί να είναι οτιδήποτε. Πάντως, ακόμα και το πλεονέκτημα της αυξημένης φυσικότητας πρόσφατα έχει αρχίσει να ατονεί λόγω των εξελίξεων στις τεχνολογίες σύνθεσης με κανόνες.

5.3.2. Ανάλυση-σύνθεση

Με τη μέθοδο αυτή, τα κύματα ομιλίας της ηχογραφημένης ανθρώπινης φωνής μετατρέπονται σε ακολουθίες παραμέτρων και αποθηκεύονται, ενώ ο συνθέτης ομιλίας οδηγείται από τέτοιες παραμέτρους τοποθετημένες σε αλληλουχία προκειμένου να παράγουν το επιθυμητό μήνυμα. Αφού οι ενότητες αποθηκεύονται μέσω παραμέτρων, η ποσότητα της πληροφορίας είναι πολύ μικρότερη από ότι στη μέθοδο κωδικοποίησης

κυματομορφής (όπως φαίνεται και από τη διαφορά στο γνώρισμα μήκους ομιλίας στον Πίνακα 5.5 του υποκεφαλαίου 5.3). Όμως η φυσικότητα της ομιλίας εκφυλίζεται.

5.3.3. Σύνθεση με κανόνες

Στη σύνθεση με κανόνες οι κυριότεροι παράμετροι για τις θεμελιώδεις μικρές ενότητες ομιλίας, όπως οι συλλαβές ή τα φωνήματα, αποθηκεύονται και κατά την παραγωγή της ομιλίας συνδέονται μεταξύ τους με τη χρήση κανόνων. Περαιτέρω κανόνες που προκύπτουν από τον επιτονισμό και την προσωδία εφαρμόζονται κατά τη συρραφή των ενοτήτων (concatenation) και επηρεάζονται και από τις γειτονικές ενότητες.

Θεωρητικά, η πιο ευέλικτη σύνθεση με κανόνες θα μπορούσε να στηρίζεται σε φωνήματα μόνο τα οποία είναι περιορισμένα σε κάθε γλώσσα (συνήθως λιγότερα από 50). Δυστυχώς όμως οι κανόνες σύνθεσης με φωνήματα ως ενότητες είναι τόσο περίπλοκοι ώστε είναι δύσκολο να ληφθεί υψηλής ποιότητας ομιλία. Για τον λόγο αυτό χρησιμοποιούνται μεγαλύτερες ενότητες για συρραφή, όπως οι δίφθογγοι. Η σύνθεση βάσει διφθόγγων παράγει καταληπτή ομιλία, η οποία όμως δεν είναι εντελώς φυσική όσον αφορά στην ποιότητά της (βλ. Πίνακα 5.5).

Μία πρόσφατη εξέλιξη, η επιλογή ενότητας (unit selection) η οποία μπορεί να είναι σε μέγεθος από μία φράση έως και μικρότερη από ένα δίφθογγο δίνει τη δυνατότητα υψηλής ποιότητας, σε αντίθεση με τη ρομποτική που προσφέρει η κλασική σύνθεση διφθόγγων. Τεχνικά, η σύνθεση με διφθόγγους είναι μια υποπερίπτωση της σύνθεσης με επιλογή ενότητας. Οι δύο όμως αυτές μέθοδοι διαφέρουν σε ένα βασικό χαρακτηριστικό. Στη σύνθεση με επιλογή ενότητας, για κάποια συγκεκριμένη ενότητα υπάρχουν περισσότερα του ενός παραδείγματα (ενώ στη σύνθεση με δίφθογγους υπάρχει μόνο ένα παράδειγμα για κάθε δίφθογγο). Ένας κατάλληλος μηχανισμός χρησιμοποιείται για να επιλέξει την ακουστικά καταλληλότερη ενότητα με τα σωστά προσωδιακά χαρακτηριστικά, ώστε να μην απαιτούνται επιπλέον τροποποιήσεις [Black 1997]. Βέβαια αυξάνονται οι απαιτήσεις γνώσεως σε βάθος γλωσσολογικών και προσωδιακών μοντέλων. Ένα βασικό μειονέκτημα (προς το παρόν) είναι ότι η δημιουργία μιας φωνής είναι ακριβή και χρονοβόρα. Για παράδειγμα στον συνθέτη RealSpeak (παράγραφος 5.3.4), μία φωνή παίρνει περίπου δύο μήνες ή και περισσότερο για να γραφτεί σε επανειλημμένες ηχογραφήσεις σε στούντιο καταλήγοντας σε 2Mb από ήχους-τμήματα από ανθρώπινη ομιλία που θα χρησιμοποιηθούν από τον αλγόριθμο. Επίσης, με τη μέθοδο αυτή, κάθε φωνή θα πρέπει να κατασκευαστεί από την αρχή. Δεν μπορεί δηλαδή να τροποποιηθεί η χροιά μιας υφιστάμενης φωνής για να δώσει μια καινούργια.

Πάντως, όσον αφορά στην χρήση της τεχνολογίας παραγωγής ομιλίας, εφόσον υπάρχουν συστήματα που προσφέρουν έτοιμες φωνές, η χρήση της σύνθεσης με κανόνες, όχι μόνο είναι κατά πολύ πιο ευέλικτη από την κωδικοποίηση κυματομορφής, αλλά σύντομα θα παράγει ποιότητα εφάμιλλη ή και καλύτερη της μεθόδου αυτής.

5.3.4. Εμπορικές εφαρμογές παραγωγής ομιλίας

Όσον αφορά στην παραγωγή ομιλίας, κυκλοφορούν διάφοροι συνθέτες σε υλικό και λογισμικό. Ενδεικτικά αναφέρονται οι συνθέτες των εταιριών AT&T [H43], Bell Laboratories [H44], [H45], Eloquence [H46], Apple [H47], [H48], [H49], [H50]. Εδώ θα παρουσιάσουμε συνοπτικά το συνθέτη λογισμικού L&H RealSpeak [H51] της εταιρείας Lernout & Hauspie [H52], που χρησιμοποιεί την τεχνολογία σύνθεσης με κανόνες και επιλογή ενοτήτων για να προσφέρει μία σειρά από γλώσσες – προς το παρόν διατίθεται σε αγγλικά, γερμανικά και γαλλικά. Οι βασικές ενότητες στο συνθέτη αυτόν δεν περιορίζονται σε διφθόγγους, αλλά περιλαμβάνουν και συλλαβές, καθώς και ακόμα μεγαλύτερες ακολουθίες φωνημάτων. Ο συνδυασμός των κατάλληλων ενοτήτων και γλωσσολογικών στοιχείων δίνει πολύ καλής ποιότητας ομιλία. Επίδειξη για τα αμερικάνικα αγγλικά μπορεί να βρεθεί στο [H53].

Φυσικά, η σύνθεση ομιλίας έχει διαφορετικά χαρακτηριστικά ανά γλώσσα. Έτσι, ένας αγγλικός συνθέτης ομιλίας δεν μπορεί να μιλήσει ελληνικά με ελληνική προφορά. Για τον λόγο αυτό, είναι σημαντικό να

παρουσιασθούν οι ερευνητικές προσπάθειες στον ελληνικό χώρο αλλά και οι εμπορικές εφαρμογές σύνθεσης που διατίθενται αυτή την στιγμή.

Ελληνικοί συνθέτες ομιλίας

Στον ελλαδικό χώρο έχουν αναπτυχθεί μια σειρά συνθετών:

Ο συνθέτης ομιλίας UoATalk του Πανεπιστημίου Αθηνών είναι ένα σύστημα μετατροπής ηλεκτρονικού κειμένου σε ομιλία που αναπτύχθηκε από την Ομάδα Επικοινωνίας με Ομιλία του Τμήματος Πληροφορικής. Το κείμενο μπορεί να προέρχεται από οποιαδήποτε εφαρμογή λογισμικού ενώ το παραγόμενο σήμα προσομοιάζει τη φυσική ομιλία του ανθρώπου, αποδίδοντάς της ρυθμιζόμενα χαρακτηριστικά και χροιά. Το σύστημα UoATalk είναι ανοικτό με αρθρωτή αρχιτεκτονική που του επιτρέπει την προσθήκη ή τη μεταβολή αρθρωμάτων. Κάθε ένα άρθρωμα υλοποιεί διακριτές λειτουργίες της ομιλίας ή της ανάλυσης του δοθέντος ηλεκτρονικού κειμένου. Ο συνθέτης χωρίζεται σε δύο τμήματα, κάθε ένα από τα οποία αποτελείται από επιμέρους αρθρώματα. Το τμήμα Επεξεργασία Φυσικής Γλώσσας περιλαμβάνει αρθρώματα όπως (α) ο Προ-Επεξεργαστής Κειμένου ο οποίος μετατρέπει σε λέξεις τα στοιχεία του κειμένου που δεν αντιστοιχούν άμεσα σε αντίστοιχους φθόγγους, (β) ο Μετατροπέας Χαρακτήρων σε φθόγγους που προσδίδει σε κάθε ένα γράμμα ένα μοναδικό ακουστικό χαρακτηριστικό, (γ) ο Γραμματικός Αναλυτής ο οποίος αναγνωρίζει γραμματικά κάθε λέξη, (δ) ο Συντακτικός Αναλυτής ο οποίος αναλύει την σύνταξη, (ε) η Συνάρθρωση η οποία φροντίζει να επικολλήσει δύο γειτονικές μονάδες (ακολουθίες) φωνητικών φθόγγων, (στ) ο Επιτονισμός ο οποίος τονίζει κάθε μία λέξη, (ζ) η Προσωδία η οποία προσδίδει προσωδιακά χαρακτηριστικά σε κάθε φράση και (η) τα Χαρακτηριστικά Ομιλητή που προσδίδουν προσωπικά χαρακτηριστικά στην παραγόμενη ομιλία. Το τμήμα Σύνθεσης Κυματομορφής αποτελείται από το άρθρωμα Επεξεργασίας Σήματος το οποίο παράγει την τελική κυματομορφή. Η τεχνολογίες που χρησιμοποιούνται και μελετώνται είναι κυρίως δύο: η πρώτη χρησιμοποιεί έναν συνθέτη formant (τύπου Klatt) και προσομοιάζει το φασματογράφημα της φυσικής ομιλίας του ανθρώπου μέσω μιας διάταξης χρονικά μεταβαλλόμενων ψηφιακών φίλτρων. Κάθε μονάδα φωνητικών φθόγγων είναι ένα σύνολο παραμέτρων που οδηγούν αυτά τα φίλτρα. Η δεύτερη χρησιμοποιεί τη μέθοδο συρραφής κυματομορφών (τύπου PSOLA) η οποία συνενώνει ηχογραφημένα τμήματα φυσικής ομιλίας προκειμένου να παράγει το τελικό σήμα ομιλίας. Στην περίπτωση αυτή κάθε μονάδα φωνητικών φθόγγων είναι μία ξεχωριστή κυματομορφή. Το τμήμα Επεξεργασίας Σήματος αποτελείται και από επιμέρους αρθρώματα επεξεργασίας της παραγόμενης κυματομορφής που μπορούν να ρυθμίζουν τα χαρακτηριστικά και τη χροιά.

Στα πλαίσια των ερευνητικών έργων ΑΙΝΕΙΑΣ (ΕΠΕΤ II, 1999-) και M-PIRO (IST, 2000-) στο Τμήμα Πληροφορικής του Πανεπιστημίου Αθηνών μελετάται και βελτιώνεται η συμπεριφορά των επιμέρους αρθρωμάτων, ενώ ερευνάται και η ανάπτυξη νέων. Μέρος των μελετών αποτελεί και η δυνατότητα αυτόματης ρύθμισης με βάση τα χαρακτηριστικά κάποιου ομιλητή. Για αυτό τον σκοπό το σύστημα σύνθεσης έχει την δυνατότητα να συνεργάζεται με προγράμματα ανάλυσης ομιλίας και κειμένου, από τα οποία εξάγει πληροφορίες για κάθε άρθρωμα.

Η Ομάδα Γλωσσικής Τεχνολογίας του Πανεπιστημίου Πατρών [H38] έχει μακρά παράδοση στις ερευνητικές περιοχές της σύνθεσης ομιλίας από κείμενο. Έχει αναπτύξει τα ακόλουθα προγράμματα και εφαρμογές:

- Ένα σύστημα σύνθεσης ομιλίας από Νεοελληνικό κείμενο με βάση τον αλγόριθμο σύνθεσης MBROLA
- Ένα σύστημα σύνθεσης ομιλίας από Νεοελληνικό κείμενο με βάση το συνθέτη formant Klutt.

Το Ινστιτούτο Επεξεργασίας του Λόγου (ΙΕΛ) [H54] προσφέρει (σε εμπορικό επίπεδο) τον “Εκφωνητή” [H55] ο οποίος στηρίζεται σε συρραφή διφθόγγων. Το ΙΕΛ επίσης έχει εκκινήσει από το 1999 μία σειρά από θεμελιώδη ερευνητικά έργα που αφορούν στην σύνθεση ομιλίας στην ελληνική γλώσσα. Ένα σημαντικό έργο είναι το “ΕΥΦΟΝΟΝ – Δημιουργία Υψηλής Ποιότητας Συνθετικής Φωνής και Χρησιμοποίησή της στην

Κινητή Τηλεφωνία” με συνεργαζόμενο φορέα την Panafon – Ελληνική Εταιρεία Τηλεπικοινωνιών Α.Ε. Το έργο αυτό αποσκοπεί

- Στη βελτίωση και τον εμπλουτισμό του συστήματος σύνθεσης φωνής από κείμενο για τα ελληνικά βασισμένο στην τεχνική formants, η οποία έχει αναπτυχθεί στο ΙΕΛ.
- Τη δημιουργία ενός συστήματος σύνθεσης φωνής από κείμενο για την ελληνική βασισμένο σε τεχνικές στο πεδίο του χρόνου.
- Τη δημιουργία συστήματος ανάγνωσης Ηλεκτρονικού Ταχυδρομείου, με χρήση του καλύτερου από τους δύο παραπάνω συνθέτες φωνής.

Αν και αυτή τη στιγμή δεν υπάρχει συνθέτης ελληνικής ομιλίας που να χρησιμοποιεί συρραφή με επιλογή ενοτήτων, κάτι τέτοιο αναμένεται στο άμεσο μέλλον, μέσα από το ευρωπαϊκό πρόγραμμα M-PIRO [H56].

Συμπερασματικά, ήδη υπάρχει δυνατότητα σύνθεσης της ελληνικής ομιλίας και σε εμπορική μορφή, και στο άμεσο μέλλον αναμένεται υψηλής ποιότητας συνθετική ομιλία στα ελληνικά.

5.3.5. Ανθρώπινοι παράγοντες

Όπως και στην αναγνώριση ομιλίας, έτσι και στη σύνθεση, η χρήση ομιλίας στη διεπαφή δεν είναι πανάκεια. Κατ’ αρχάς σημαντικός είναι ο τύπος της πληροφορίας που πρέπει να παρουσιαστεί. Μερικοί τύποι πληροφορίας, όπως η φωτογραφία ενός προϊόντος, δεν μπορούν να μεταδοθούν ηχητικά παρά μόνον οπτικά. Φυσικά είναι δυνατόν να περιγραφούν μέχρι ενός σημείου. Το ίδιο ισχύει και για διαγράμματα, χάρτες, σχήματα. Διαφορετική αντίληψη αποκτά ο χρήστης που βλέπει μπροστά του ολόκληρο το διάγραμμα και διαφορετική αυτός που προσπαθεί να το φανταστεί βάσει μιας περιγραφής με ομιλία. Σε αυτές τις περιπτώσεις, εάν υπάρχει η δυνατότητα χρήσης κάποιας διαφορετικής μεθόδου παρουσίασης, τότε ίσως αυτή θα πρέπει να προτιμηθεί.

Στην περίπτωση που η επικοινωνία πρέπει να γίνει οπωσδήποτε ακουστικώς, τότε πάλι θα πρέπει να ληφθούν διάφοροι παράγοντες υπόψη. Ένα σημαντικό στοιχείο είναι ότι η ομιλία μερικές φορές είναι χρονοβόρα και προτιμάται η χρήση σύντομων ήχων [Mitsopoulos 1999] ή συνδυασμός ομιλίας και ήχων [Stevens 1996]. Για παράδειγμα, είναι δυνατή η ηχητική περιγραφή ενός γραφήματος ενώ η λεκτική θα ήταν και χρονοβόρα και δύσληπτη [H57].

Θα πρέπει να παρατηρηθεί μια διαφορά μεταξύ των γραφικών διεπαφών (GUIs) και των διεπαφών που χρησιμοποιούν ομιλία. Στην πρώτη περίπτωση η πληροφορία παρουσιάζεται στατικά και είναι διαθέσιμη ανά πάσα στιγμή στον χρήστη. Έτσι μπορεί να παρουσιαστεί για παράδειγμα μια ολόκληρη σελίδα κειμένου με γραφικά και ο χρήστης να την εξερευνήσει κατά βούληση. Δεν ισχύει όμως το ίδιο για την ομιλία και τον ήχο, για τα οποία η παρουσίαση είναι σειριακή και η πληροφορία πρέπει να συγκρατείται στη μνήμη του χρήστη. Οδηγίες για τη φωνητική παρουσίαση μπορούν να βρεθούν στο [Pitt 1996].

5.4. Ολοκληρωμένα διαλογικά συστήματα με ομιλία

Το ζητούμενο εδώ είναι το πώς ολοκληρώνεται η τεχνολογία σύνθεσης και παραγωγής σε μία διεπαφή που χρησιμοποιεί ομιλία. Παράδειγμα αποτελεί η Philips και τα διαλογικά συστήματά της (π.χ. σε κινητά ή συσκευές νοικοκυριού). Το πιο σημαντικό χαρακτηριστικό αυτών των συσκευών είναι η ευρωστία, δηλ. μπορούν να δουλεύουν κάτω από όλες τις πιθανές συνθήκες.

Πώς μετράμε την επιτυχία ενός διαλογικού συστήματος; Στην Philips χρησιμοποιούν τον όρο “workability”. Αυτός διαφέρει από το ποσοστό λαθών που πολλές φορές επισημαίνεται σε διάφορα περιοδικά για υπολογιστές ή τεχνικά χαρακτηριστικά. Για διαλογικά συστήματα, το πραγματικό θέμα δεν είναι το ποσοστό λαθών στην αναγνώριση. Εκείνο που πρέπει να μετρηθεί είναι το ποσοστό επιτυχίας του διαλόγου, δηλαδή το ποσοστό των χρηστών που αποκτούν την πληροφορία που απαιτούν. Αυτό το ποσοστό εξαρτάται όχι μόνο από την ποιότητα

του συστήματος αναγνώρισης ομιλίας, αλλά και από την ύπαρξη ενός συστήματος κατανόησης ομιλίας και ενός έξυπνα σχεδιασμένου διαλόγου [Joscelyne 1999]. Σε άλλες περιπτώσεις, όπως σε συσκευές που θα χρησιμοποιηθούν μέσα σε κάποιο όχημα, ένας από τους πιο σημαντικούς παράγοντες είναι η ευρωστία στον περιβάλλοντα θόρυβο. Εξίσου σημαντική είναι η έξυπνη σχεδίαση της διεπαφής που διευκολύνει την αλληλεπίδραση. Το ποσοστό λαθών δεν είναι καθαυτό σημαντικό. Ζωτικής σημασίας είναι να αποκτήσει ο χρήστης την εντύπωση ότι το προϊόν δουλεύει αποτελεσματικά.

5.4.1. Πολυτροπία

Τα συστήματα που χρησιμοποιούν ομιλία μπορεί να είναι πολυτροπικά. Η πολυτροπία (multimodality) μπορεί να υπάρχει τόσο όσον αφορά στη χρήση διαφορετικών αισθητήριων καναλιών (sensory channels), αποκλειστικά δηλαδή την είσοδο ή την έξοδο του συστήματος, όσο και στον συνδυασμό διαφορετικών αισθήσεων για την είσοδο / έξοδο του συστήματος.

Για την έξοδο του συστήματος μπορεί να γίνει συνδυασμός εικόνας / ήχου. Χαρακτηριστικά παράδειγμα είναι τα Verbots [H58] και το “visual text-to-speech” της AT&T [H59]. Πάντως χρειάζεται προσοχή στο συνδυασμό και το χρονοισμό. Για παράδειγμα, εάν η γραφική απεικόνιση του προσώπου είναι υψηλής ευκρίνειας τότε ακόμα και μικροί αποσυγχρονισμοί μεταξύ της ομιλίας και της κίνησης των χειλιών εντοπίζονται από το χρήστη, κάτι το οποίο δεν συμβαίνει για παράδειγμα με μικρών διαστάσεων πρόσωπα. Για περισσότερες πληροφορίες απευθυνθείτε στο [Massaro 1998].

Το πιο πρόσφατο αξιόλογο παράδειγμα στο χώρο αποτελεί η εικονική παρουσιάστρια ειδήσεων στο διαδίκτυο Ananova [H60]. Στην ουσία πρόκειται για ένα πολύπλοκο σύστημα με τη δική του μηχανή αναζήτησης ειδήσεων στο διαδίκτυο. Κάθε είδηση τυγχάνει επεξεργασίας από το σύστημα και μαρκάρεται κατάλληλα έτσι ώστε οι εκφράσεις της παρουσιάστριας να αντικατοπτρίζουν το νόημα του εκφωνούμενου κειμένου (για παράδειγμα να μην είναι χαρούμενη σε μία λυπηρή είδηση). Η μαρκαρισμένη είδηση τροφοδοτείται στο σύστημα παραγωγής ομιλίας από κείμενο αφενός και στο σύστημα προσομοίωσης προσώπου αφετέρου που παράγει τις εκφράσεις του προσώπου της Ananova και συγχρονίζεται η ομιλία με την εικόνα.

Όσον αφορά στον συνδυασμό διαφορετικών τρόπων εισόδου (input modalities) είναι πιθανό για παράδειγμα να χρησιμοποιείται η ομιλία μαζί με κάποιες χειρονομίες (π.χ. deixis), όπως πολλές φορές συμβαίνει και στην καθημερινή ομιλία. Πολύ πρακτικά, σε μία διεπαφή χρησιμοποιούμε και φωνή και το mouse ταυτόχρονα: “δείξε μου τις πτήσεις προς αυτή την πόλη” ταυτόχρονα κάνοντας κλικ στο όνομα της πόλης. Πολλές φορές όμως τρόποι αλληλεπίδρασης που εμφανίζονται φυσιολογικοί εγκυμονούν κινδύνους. Παράδειγμα αποτελεί το σύστημα MATIS που χρησιμοποιεί τον προαναφερθέντα τρόπο δήξεως όπως αναλύεται με το ψυχολογικό μοντέλο Interactive Cognitive Subsystems (ICS) [Barnard 1994].

Πολυτροπία μπορούμε να έχουμε και με την έννοια ότι ολοκληρώνεται η ομιλία σαν είσοδος με την οπτική παρουσίαση προκειμένου να προσφέρει μία πιο έξυπνη και αποτελεσματική διεπαφή πρόσβασης διατηρώντας τα πιθανά πλεονεκτήματα της οπτικής παρουσίας. Βέβαια η χρήση τηλεφώνου υπονοεί εν γένει ότι τα κανάλια επικοινωνίας που μπορούν να χρησιμοποιηθούν είναι ο ήχος και η ομιλία.

5.4.2. Οι προθέσεις των εταιρειών

Η Philips στα πλαίσια της στρατηγικής της να παραμείνει ένας βασικός παράγων στην αγορά των καταναλωτικών ηλεκτρονικών αγαθών θεωρεί την τεχνολογία ομιλίας ως τεχνολογία-κλειδί, μια που είναι βασικό συστατικό των εύχρηστων διεπαφών. Μακροπρόθεσμα, όλα τα συστήματα της Philips που επιδέχονται μία φωνητική διεπαφή θα είναι εξοπλισμένα με αυτή [Joscelyne 1999]. Επίσης η Intel (όπως και η Microsoft) έχουν συμμαχήσει με τη βελγική Lernout & Hauspie [H61] που ειδικεύεται στην τεχνολογία ομιλίας. Η τελευταία εταιρία μάλιστα πρόσφατα (28/3/2000) αγόρασε την αντίστοιχη της αμερικάνικη Dragon Systems Inc [H62]. Για το ίδιο λόγο συμμαχούν η IBM και η NOKIA.

Η εταιρεία παραγωγής εξοπλισμού για αυτοκίνητα Delphi Automotive Systems πρόσφατα ανακοίνωσε τη συνεργασία της με την εταιρεία Palm. Θέμα της συνεργασίας είναι η προσθήκη τεχνολογιών ομιλίας στους υπολογιστές χειρός PalmTops προκειμένου οι οδηγοί να μπορούν να τους χρησιμοποιούν διεξάγοντας με φωνητικές εντολές τις εργασίες τους κατά τη διάρκεια της οδήγησης. Μελλοντικά οι οδηγοί, όχι μόνο θα έχουν πρόσβαση στα δεδομένα του PalmTop τους με φωνητικό τρόπο, αλλά επιπλέον και πρόσβαση στο διαδίκτυο.

Παρατηρείται λοιπόν μία έντονη κινητικότητα στο χώρο από τις μεγαλύτερες μάλιστα εταιρείες, με συνασπισμούς και εξαγορές. Τούτο οφείλεται στο γεγονός ότι οι υφιστάμενες τεχνολογίες έχουν ωριμάσει κατά πολύ και οι εταιρείες διαβλέπουν μεγάλα κέρδη από την εκμετάλλευσή τους.

5.5. Χρήστες με ειδικές ανάγκες

Οι τεχνολογίες ομιλίας έχουν βρει εφαρμογή στον τομέα της προσθετικής επικοινωνίας (augmentative communication) που περιλαμβάνει ανάπηρους και ηλικιωμένους χρήστες με προβλήματα στην όραση ή κινητικά προβλήματα. Οι βασικές εφαρμογές μπορούν να ταξινομηθούν σε πέντε κατηγορίες [Kouroupetroglou et al 1995]:

- Βοηθήματα επικοινωνίας για άτομα που δε μιλάνε, έχουν προβλήματα ομιλίας, ή είναι κωφά (voice output communication aids – VOCAs) τα οποία υποκαθιστούν την ομιλία. Σε αυτή την περίπτωση ενδιαφέρει (όσον αφορά την τεχνολογία ομιλίας) η ποιότητα της σύνθεσης. Στον ελληνικό χώρο ισχύουν τα όσα προαναφέρθηκαν στην παράγραφο 5.3.4.
- Αναγνώριση ομιλίας για γραφή, προγραμματισμό, έλεγχο περιβάλλοντος και συστήματα σχεδιασμού με τη βοήθεια υπολογιστή (Computer Aided Design – CAD) για άτομα με κινητικά προβλήματα. Η τεχνολογία αναγνώρισης φωνής μπορεί να αυξήσει το ρυθμό γραπτής επικοινωνίας σε σύγκριση με αυτόν που μπορεί να επιτευχθεί είτε με κανονικό πληκτρολόγιο είτε με πληκτρολόγιο με πρόβλεψη λέξεων. Επίσης μπορεί να αντικαταστήσει το ποντίκι σε γραφικά περιβάλλοντα.
- Σύνθεση ομιλίας ως βοήθημα για ανάγνωση, γραφή και προγραμματισμό για χρήστες τυφλούς ή με προβλήματα όρασης.
- Μέθοδοι και συστήματα εκπαίδευσης ομιλίας για κωφούς.
- Επεξεργασία ομιλίας σε βοηθήματα ακοής και αφής και εμφυτεύσεις στον κοχλία του αυτιού που μετατρέπουν την ομιλία σε απτικά ή οπτικά σήματα ή δημιουργούν νέα ηχητικά σήματα τα οποία διοχετεύονται απευθείας στον κοχλία.

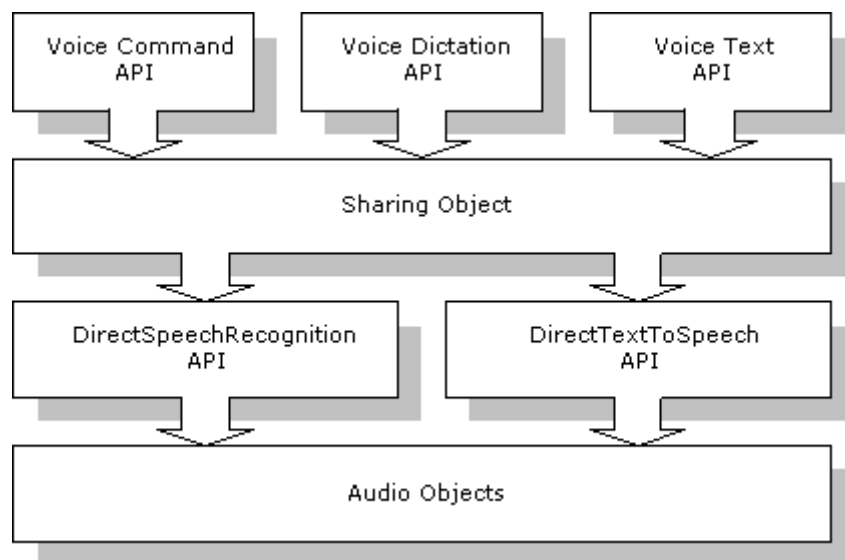
5.6. Πρωτόκολλα και πρότυπα επικοινωνίας

Ένα σημαντικό σημείο είναι το πώς το εκάστοτε σύστημα αναγνώρισης ή σύνθεσης ομιλίας μπορεί να ολοκληρωθεί με την υπάρχουσα διεπαφή και να επικοινωνήσει με τα υπόλοιπα κομμάτια της εφαρμογής. Αυτή την στιγμή υπάρχει μία πληθώρα από πρότυπα, τα περισσότερα από τα οποία στηρίζονται στην γλώσσα XML.

Κατ' αρχάς θα αναφέρουμε δύο πρότυπα που χρησιμοποιούνται για την επικοινωνία εφαρμογών που απαιτούν παραγωγή ομιλίας και συσκευών παραγωγής ομιλίας. Εδώ και μερικά χρόνια η εταιρεία Arkenstone Inc. [H63] διέγινε το εξής πρόβλημα: στην αγορά υπήρχαν πολλοί συνθέτες ομιλίας (και οθόνες Braille), ο καθένας όμως με τη δική του διεπαφή. Η εταιρεία αυτή δημιούργησε τις προδιαγραφές SSIL (Speech Synthesizer Interface Library) οι οποίες υιοθετήθηκαν από τους περισσότερους κατασκευαστές. Έτσι για να χρησιμοποιήσει μια εφαρμογή έναν οποιονδήποτε συνθέτη ομιλίας αρκεί να χρησιμοποιήσει τις προδιαγραφές SSIL. Ένα βασικό πρόβλημα με το SSIL είναι ότι είχε γραφεί για 16bit εφαρμογές, οι οποίες δεν μπορούσαν να είναι ταυτόχρονα ενεργές. Μία εξέλιξη αυτών των προδιαγραφών είναι οι προδιαγραφές SAM (Synthesizer Access Manager) οι οποίες επιτρέπουν το έξυπνο μοίρασμα του συνθέτη ομιλίας (και τυχόν οθονών Braille) μεταξύ διαφορετικών εφαρμογών οι οποίες προσπαθούν να τα προσπελάσουν ταυτόχρονα [H64]. Έτσι το SAM έχει αρχιτεκτονική 32bit που του επιτρέπει να χρησιμοποιηθεί με τα Windows9x και Windows NT

προσφέροντας στην ουσία μια πλατφόρμα ανεξάρτητη του υλικού και επίσης μία διεπαφή SAM-to-SSIL για να υποστηρίξει συσκευές οι οποίες δεν υποστηρίζουν ακόμα απευθείας SAM.

Όσον αφορά στα περιβάλλοντα Windows της Microsoft υπάρχει ήδη ένα πρωτόκολλο επικοινωνίας για τις εφαρμογές που απαιτούν χρήση ομιλίας, το Microsoft Speech API [H65]. Αυτό αποτελείται από 8 βασικά τμήματα τα οποία ονομάζονται ως εξής: Voice Command, Voice Dictation, Voice Text, Voice Telephony, DirectSpeechRecognition, DirectTextToSpeech και Audio Objects. Το τμήμα Voice Telephony παρέχει σύνθεση και αναγνώριση ομιλίας, σύνθεση κυματομορφών τύπου wave (της Microsoft) και DTMF σε απλής ή πολλαπλών γραμμών τηλεφωνικές συσκευές ομιλίας. Η αρχιτεκτονική των τμημάτων φαίνεται στο Σχήμα 36.



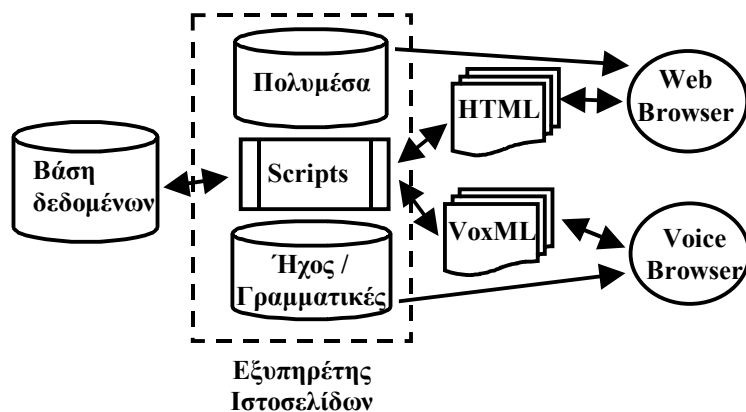
Σχήμα 36: Η αρχιτεκτονική του Microsoft Speech API

Γενικά, κεντρικό ρόλο στα πρωτοκόλλα που αφορούν τη χρήση ομιλίας έχει η γλώσσα XML (eXtensible Markup Language [H66]), η οποία είναι μία απλή διάλεκτος της SGML και διάδοχος της HTML. Ειδικά στο ηλεκτρονικό εμπόριο, αλλά και γενικότερα στο διαδίκτυο, υπάρχει μία σαφέστατη τάση υιοθέτησης της XML. Σε αυτό το στάδιο, ήδη υπάρχουν εξειδικευμένες επεκτάσεις της XML που αφορούν την ομιλία.

Μία σημαντική προσπάθεια έγινε από την Sun Microsystems το 1997 η οποία ανέπτυξε το πρωτόκολλο Java Speech Markup Language, ή JSML. Η JSML περιλαμβάνει στοιχεία για το πώς πρέπει να ειπωθεί κάποιο κείμενο, την προφορά λέξεων, προσωδία, κλπ [H67]. Η JSML είναι ένα υποσύνολο της XML. Επειδή είναι υποσύνολο της XML, η JSML αποκτά αυτόματα ένα στάνταρτ και επεκτάσιμο συντακτικό το οποίο δεν εξαρτάται από το Java Speech API [H68].

Η εταιρεία Motorola αναπτύσσει την VoxML Voice Markup Language η οποία υποστηρίζει αναγνώριση ομιλίας για είσοδο και σύνθεση ομιλίας για παρουσίαση πληροφοριών [H69]. Η γλώσσα αυτή όπως και η JSML βασίζεται στο στάνταρτ XML. Σύμφωνα με τη σχηματική διάταξη που προτείνει η voxML (στο “VoxML White Paper” [H70]), ο web server αποστέλλει στο χρήστη μία HTML σελίδα, εάν η πρόσβαση γίνεται από κανονικό browser ή μία σελίδα voxML, εάν η πρόσβαση γίνεται μέσω τηλεφώνου ή άλλης ηχητικής διεπαφής, όπως φαίνεται στο Σχήμα 37.

Το Μάρτιο του 2000, το φόρουμ της voiceXML [H71] ανακοίνωσε την έκδοση 1.0 των προδιαγραφών της. Η voiceXML είναι αποτέλεσμα κοινής προσπάθειας 79 υποστηρικτών της, μεταξύ των οποίων οι AT&T, Motorola, IBM, Lucent Technologies και πολλά άλλα σημαντικά ονόματα στο χώρο [H72]. Προς το παρόν η έκδοση 1.0 εξετάζεται από τα μέλη του φόρουμ, αλλά είναι πιθανό να αποτελέσει το στάνταρτ κομμάτι της XML όσον αφορά στην ομιλία.



Σχήμα 37: Αρχιτεκτονική προτεινόμενη από την Motorola

Ένα πολύ σημαντικό πρωτόκολλο είναι το Wireless Application Protocol (WAP), ένα προϊόν συνασπισμού των γιγάντων στις κινητές τηλεπικοινωνίες Ericsson, Motorola και Nokia. Το πρωτόκολλο αυτό αντιμετωπίζει μια σειρά προβλημάτων των συσκευών της κινητής τηλεφωνίας, όπως η ασφάλεια στην επικοινωνία, το ασταθές δίκτυο, αλλά και εργονομικούς παράγοντες, όπως το μικρό μέγεθος οθόνης και η χρήση τεχνολογίας ομιλίας για καλύτερες δυνατότητες εισόδου και εξόδου πληροφορίας στη διεπαφή. Αναμένεται να διαδραματίσει κεντρικό ρόλο στο ηλεκτρονικό εμπόριο με συσκευές κινητής τηλεφωνίας.

6. Τεχνολογίες Ελέγχου με Εγκεφαλικά Κύματα (Brain Control Technologies)

Τα συστήματα ελέγχου με εγκεφαλικά κύματα αντιπροσωπεύουν το επόμενο βήμα στην εξέλιξη της διεπαφής εισόδου μεταξύ ανθρώπου και μηχανής. Ήδη έχουν εμφανιστεί στο εμπόριο τα πρώτα από τα συστήματα της κατηγορίας αυτής. Τα συστήματα αυτά είναι τεχνολογίες ελέγχου, που λειτουργούν με τον εγκέφαλο και ενσωματώνουν την κίνηση ματιών, την ενεργοποίηση μυών και την ανίχνευση βίο-δυναμικών κυμάτων του εγκεφάλου στο μέτωπο των χρηστών, ενώ είναι σε θέση να παράγουν σήματα εισόδου στον υπολογιστή που μπορούν να χρησιμοποιηθούν σε μια πληθώρα εφαρμογών.

Η διεπαφή των συστημάτων ελέγχου με εγκεφαλικά κύματα είναι ένας σύνδεσμος επικοινωνίας μεταξύ ανθρώπου και μηχανής. Με τη χρήση αυτής της διεπαφής οποιοσδήποτε που επιθυμεί να εξερευνήσει τον κόσμο ελέγχου μέσω εγκεφάλου, έχει τη δυνατότητα να χειριστεί πλήθος υπολογιστικών προγραμμάτων και οποιαδήποτε ηλεκτρική συσκευή κατευθείαν από το κέντρο ελέγχου του, τον εγκέφαλο.

Οι διεπαφές αυτές δίνουν τη δυνατότητα ελέγχου του υπολογιστή και οποιασδήποτε ηλεκτρικής συσκευής χωρίς χέρια. Τα συστήματα ελέγχου με εγκεφαλικά κύματα είναι εύκολα και ευχάριστα στη χρήση και καθόλου δύσκολα στην εκμάθηση. Ο χρήστης φορά την κορδέλα (headband) στο κεφάλι, συνδέει το άλλο άκρο στον υπολογιστή και αφήνει το μυαλό να κάνει τα υπόλοιπα (Σχήμα 38)



Σχήμα 38: Σύστημα ελέγχου με εγκεφαλικά κύματα

Τα σήματα του εγκεφάλου και του σώματος ανιχνεύονται από τους αισθητήρες που βρίσκονται πάνω στην κορδέλα. Τα σήματα αυτά ενισχύονται, ψηφιοποιούνται και εκπέμπονται στον υπολογιστή σαν πολλαπλά Εγκεφαλοδάκτυλα, (Brainfingers), για να επηρεάσουν την ανάδραση της οθόνης, τον έλεγχο του ποντικιού, να ελέγξουν σχεδόν οποιαδήποτε εφαρμογή, να παίξουν μουσικά όργανα, να ενεργοποιήσουν περιφερειακές συσκευές και πολλά άλλα.

Η διεπαφή του συστήματος ελέγχου με εγκεφαλικά κύματα είναι η βάση για την τεχνολογία επιθετικής προσπέλασης (breakthrough) που ανιχνεύει και ανταποκρίνεται σε στιγμιαία ηλεκτρικά σήματα, τα οποία είναι επακόλουθο της λειτουργίας του εγκεφάλου και των ευαίσθητων μυών.

Τρεις αισθητήρες που βρίσκονται στην κορδέλα ανιχνεύουν τη δραστηριότητα του εγκεφάλου. Οι μονάδες διεπαφών των συστημάτων ελέγχου εγκεφάλου ενισχύουν και μεταφράζουν τα δεδομένα σε ξεχωριστές

συχνότητες και τα μεταδίδουν στη σειριακή θύρα του υπολογιστή. Στη συνέχεια το λογισμικό του συστήματος τα επεξεργάζεται και εμφανίζει τις συχνότητες αυτές σαν 10 συνεχή σήματα που ονομάζονται Εγκεφαλοδάκτυλα (Σχήμα 39). Επιπλέον, παράγονται και 4 ξεχωριστά σήματα από κάποιες συγκεκριμένες κινήσεις των ματιών και του προσώπου.



Σχήμα 39: Εγκεφαλοδάκτυλα

Τα Εγκεφαλοδάκτυλα έχουν τη δυνατότητα να ελέγχουν το δρομέα του ποντικιού πάνω στην οθόνη και μπορούν επίσης να προγραμματιστούν από το χρήστη να στέλνουν εντολές που έχουν σχέση με τα πλήκτρα του ποντικιού και του πληκτρολογίου. Οι κινήσεις των ματιών και οι χειρονομίες του προσώπου είναι ξεχωριστά σήματα ελέγχου, τα οποία χρησιμοποιούνται από το χρήστη για να προγραμματιστούν οι λειτουργίες του ποντικιού και του πληκτρολογίου.

Οποιαδήποτε εφαρμογή ή παιχνίδι είναι δυνατόν να ελεγχθεί χωρίς χέρια (hands-free), με τη χρήση συστημάτων με εγκεφαλικά κύματα.

Αναφορές

1. Avons S. E., Beveridge M. C., Hickman A. T. and Hitch G. J.: Considerations on using a lighthpen-interactive system with young children, Behavior Research Methods & Instrumentation, 15, pp. 75-78, 1983.
2. Bayes D. L. and Thompson R. A.: An experimental teleterminal - the software strategy. Bell System Technical Journal, pp. 121-144, January, 1983.
3. Beddoes M. P. and Hu Z.: A chord stenograph keyboard: A possible solution to the learning problem in stenography. IEEE Transactions on Systems, Man and cybernetics, 24, pp. 953-960, 1994.
4. Bolt R. A.: The human interface: Where people and computers meet. Belmont, CA: Lifetime Learning Publications, 1984.
5. Brain Actuated Technologies Inc. <http://www.brainfingers.com/index.html>
6. Brigham F. R. and Clark N.: Comparison of initial learning and acceptance: STR ergonomic keyboard vs. standard keyboard. (653-ITT-00894), Essex, England, ITT Europe, 1986.
7. Clare C. R. "Human factors: A most important ingredient in keyboard designs". EDN Magazine (Electrical Design News), 21(8), 99-102, 1976.
8. Brown E., Buxton W. and Murtagh K.: Windows on tablets as a means of achieving virtual input devices, Computer graphics, 19, pp. 225-230, 1985.
9. Beddoes M. P. and Hu Z. A chord stenograph keyboard: A possible solution to the problem in stenography. IEEE Transactions and Systems, Man, and Cybernetics, 24, 953-960, 1994.
10. Cakir A., Hart D. J. and Stewart T. F. M.: Visual Display Terminals. New York, NY, John Wiley, 1980.
11. Colven D. and Detheridge T.: A Common Terminology for Switch Controlled Software, A guide for software developers, ACE Center, NCET, 1990.
12. Crick Software <http://www.cricksoft.com/switches.htm>
13. Deininger R. L.: Human factors engineering studies of the design and use of pushbutton telephone sets. Bell Systems Technical Journal, 12(3), pp. 67-79, 1960.
14. Dvorak A.: There is a better typewriter keyboard. National Business Education Quarterly, XII-2, pp. 51-58 and 66, December 1943.
15. Emmons W. H.: A comparison of cursor-key arrangements (box versus cross) for VDUs. In E. Grandjean (Ed.), Ergonomics and Health in Modern Offices, pp. 214-219, London, UK, Taylor and Frances, 1984.
16. Foley J. D. and van Dam A.: Fundamentals of interactive computer graphics. Reading MA, Addison-Wesley, 1982.
17. Galitz W. O.: CRT keyboard human factors evaluation: Study II. Univac, System Application Engineering, Roseville DOD, February 1966.
18. Getschow C. O., Rosen M. J. and Goodenough-Trepagnier C.: A systematic approach to design of a minimum distance alphabetical keyboard. In RESNA 9th Annual Conference, pp. 396-398, Minneapolis, MN, RESNA, 1986.
19. Glen F. A., Iavecchia H. P., Ross L. V., Stokes J. M., Weiland W. J., Weisis D. and Zakland A. L. "Eye-voice-controlled interface". In Proceedings of the Human Factors Society 30th Annual Meeting (pp. 322-326). Santa Monica, CA: Human Factors and Ergonomics Society, 1986).

20. Goodman D., Dickinson J. and Francas M.: Human factors in keypad design. In Proceedings of the Human Factor Society 27th Annual Meeting, pp. 191-195, Santa Monica, CA, Human Factors Society, 1983.
21. Gopher D. and Raij D.: Typing with a two-hand chord keyboard: Will the QWERTY become obsolete? IEEE Transactions on Systems, Man and Cybernetics, 18, pp. 601-609, 1988.
22. Hall A. D., Cunningham J. B., Roache R. P. and Cox J. W.: Factors affecting performance using touch-entry systems, Tactual recognition fields and system accuracy, Journal of applied psychology, 73, 711-720, 1988.
23. Hatamian M. and Brown E. F.: A new light pen with subpixel accuracy, AT&T Technical Journal, 64, pp. 1065-1075, 1985.
24. Helander M. G., Landauer T. K. and Prabhu P. V.: Handbook of Human-Computer Interaction, Second Completely Revised Edition, Published by Elsevier Science B. V., 1997.
25. Hill G. W., Gunn W. A., Martin S. L. and Schwartz D. R.: Perceived difficulty and user control in mouse usage. In Proceedings of the Human Factors Society 35th Annual Meeting, pp. 295-299, Santa Monica, CA: Human Factors and Ergonomics Society, 1991.
26. Hirsch R. S.: Effects of standard versus alphabetical keyboard formats on typing performance. Journal of Applied Psychology, 54, pp. 484-490, 1970.
27. Jacob R. J. K.: The use of eye movements in human-computer interaction techniques: What you look at is what you get. ACM Transactions of Information Systems, 9, pp. 152-169, 1991.
28. Jahns D. W., Litewka J, Farrand W. P. and Hargreaves W. R.: Learning curve and performance analysis for the Kinesis TM ergonomic keyboard - a pilot studies. Presented as a poster at the HFS 35th Annual Meeting, San Francisco, Ca, Copies available from Kinesis, September 2-6, 1991.
29. Kinkeard R.: Typing speed, keying rates and optimal keyboard layouts. In Proceedings of the Human Factors Society 19th Annual Meeting, pp. 159-161, Santa Monica, CA, 1975.
30. Kroemer K. H. E.: Human engineering the keyboard. Human Factors, 14, pp. 51-63, 1972.
31. Lewis J. R.: Meta-analysis of preference for split versus standard keyboards: Findings from 1972 to 1993 (Tech. Report 54.899). Boca Raton, FL, International Business Machines Corp., 1995.
32. Lewis J. R.: Typing-key layouts for single-finger or stylus input: Initial user preference and performance. Tech. Report 54.729, Boca Ration, FL, International Business Machines Corp., 1994.
33. Loricchio D. F. and Kennedy P. J.: Keyspace and user productivity. In Abridged Proceedings of Poster Sessions of the third International Conference on Human-Computer Interaction, p. 48, New York, NY, Elsevier, 1987.
34. Loricchio D. F. and Lewis J. R User assessment of standard and reduced-size numeric keypads. In Proceedings of the Human Factors Society 35th Annual Meeting (pp. 251-252), Santa Monica, CA: Human Factors Society, 1991.
35. Lutz M. C. and Chapanis A.: Expected locations of digits and letters on ten-button keysets. Journal of Applied Psychology, 39, pp. 314-317, 1955.
36. MacKenzie I. S., Sellen A. and Buxton W.: A comparison of input devices in elemental pointing and dragging tasks. In Proceedings of the CHI '91 Conference on Human Factors in Computing Systems, pp. 161-166, New York, ACM, 1991.
37. Magyar R. L.: Comparison of user performance and preference for cursor keypad configuration and numeric keypad position (Tech. Memorandum 08.162). Lexington, KY, International Business Machines Corp., 1986.

38. Magyar R. L.: Effect of curved, flat and stepped keybutton configurations on keyboard preference and throughput performance. Internal IBM report, 1985.
39. Matias E., MacKenzie I. S. and Buxton W.: Half-QWERTY: A one-handed keyboard facilitating skill transfer from QWERTY. In Conf. Proc. on Human Factors in Computing Systems - CHI '93, pp.88-94, New York, NY, Association for Computing Machinery. 1993.
40. Murray J. T., Van Praag J. and Gilfoil D.: Voice versus keyboard control of cursor motion, Proc. Human Factors Society-Twenty seventh Annual Meeting, 103, 1983.
41. Najjar L. J., Stanton B. C. and Bowen C. D.: Keyboard heights and slopes for standing typists. Tech. Report 85-0081, Rockville, MD, International Business Machines Corp., 1988.
42. Nakaseko M, Grandjean E, Hunting W and Gierer R.: Studies on ergonomically designed alphanumeric keyboards. Human Factors, 27, pp. 175-187, 1985.
43. Norman D. A. and Fisher D.: Why alphabetic keyboards are not easy to use: Keyboards layout doesn't much matter. Human Factors, 24, pp. 509-519, 1982.
44. Parrish R. N., Gates J. L., Munger S. J., Grimma P. R. and Smith L. T. Development of design guidelines and criteria for user/operator transactions with battlefield automated system, Phase II Final Report: Volume II, Prototype handbook for combat and materiel developers (Tech. Report). Alexandria, VA, U.S. Army Research Institute for the behavioral and Social Sciences, 1982.
45. Pausch R. and Leatherby J. H.: An empirical study: Adding voice input to a graphical editor, Journal of the American Voice Input-Output Society 9, pp. 55-66, 1991.
46. Paci A. M. and Gabrielli L.: Some experiences in the field of design of VDU work station. In E. Grandjean (Ed.), Ergonomics and Health in Modern Offices, pp. 391-399, Philadelphia, PA, Taylor & Francis, 1984.
47. Price L. A. and Cordova C. A.: Use of mouse buttons. In Proceedings of the CHI '83 Conference on Human, Factors in Computing Systems, pp. 262-266, New York, ACM, 1983.
48. Rosch W. L. Keyboard ergonomics for IBMs. PC Magazine, 3(19), 110-112, 1984.
49. Schulze L. J. H. and Snyder H. L.: A comparative evaluation of five touch entry devices (Tech. Report HFL-83-6). Blacksburg: Virginia Polytechnic institute and State University, Department of Industrial Engineering and Operations Research, October 1983.
50. Shneiderman B.: Designing the User Interface, Strategies for Effective Human-Computer Interaction, Second Edition, Addison-Wesley Publishing Company, 1992.
51. Smith W. J. and Cronin D. T.: Ergonomic test of the Kinesis keyboard. In Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, pp. 318-322, Santa Monica, CA, Human Factors and Ergonomics Society, 1993.
52. Strong E. P. A comparative experiment in simplified keyboard retraining and standard keyboard supplementary training. Washington, DC, General Services Administration, 1956.
53. Suther T. W. and McTyre J. H.: Effect on operator performance at thin profile keyboard slopes of 5, 10, 15 and 25 degrees. In Proceedings of the Human Factors Society 26th Annual Meeting, pp. 430-434, Santa Monica, CA, Human Factors Society, 1982.
54. Valk M. A.: An experiment to study touch-screen "button" design. In Proceeding of the Human Factors Society 29th Annual Meeting, pp. 127-131, Santa Monica, CA, Human Factors and Ergonomic Society, 1985.
55. Ware C. and Mikaelian H. H. An evaluation of an eye tracker as a device for computer input. In Proceedings of the CHI+Gi 1987 Conference on Human Factory in Computing Systems and Graphics Interface, pp. 183-188, New York, ACM, 1987.

56. Whitfield D., Ball R. G. and Bird J. M.: Some comparisons of on-display and off-display touch input devices for interaction with computer generated displays, *Ergonomics*, 26, pp. 1033-1053, 1983.
57. Wiklund M. E., Dumas J. S. and Hoffman L. R.: Optimizing a portable terminal keyboard for combined one-handed and two-handed use. In *Proceedings of the Human Factors Society 31st Annual Meeting*, pp. 585-589, Santa Monica, CA, Human Factors Society, 1987.
58. Yamada H. A historical study of typewriters and typing methods: From the Position of planning Japanese parallels (Tech. Report 80-05). Department of Information Science, Faculty of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo Tokyo, 113 Japan, 1980.
59. Zimmerman T. J., Lanier J., Blanchard C., Bryson S. and Harvill Y.: A hand gesture interface device. In *Proceedings of the CHI+GI 1987 Conference on Human Factors in Computing Systems and Graphics Interface*, pp. 189-192, New York: ACM, 1987.
60. Buessen J., Product development of an ergonomic keyboard. *Behaviour and Information Technology*, 3, pp. 387-390, 1984.
61. Barnard, P. and J. May (1994). Interactions with Advanced Graphical Interfaces and the Deployment of Latent Human Knowledge. *Interactive Systems: Design, Specification, and Verification*. F. Paterno, Springer: 15-49.
62. Black A. and Taylor P. (1997). Automatically clustering similar units for unit selection in speech synthesis. In *Eurospeech97*, Rhodes, Greece, Vol. 2, pp. 601-604
63. Corredor-Ardoy C., Gauvain J.L., Adda-Decker M. Lamel L. (1997). Language identification with language-independent acoustic models. In *Proceedings of Eurospeech97*, Rhodes, Greece. Vol. 1, pp. 55-58
64. Deshmukh N., Ganapathiraju A., Picone J. (1999). Hierarchical Search for Large-Vocabulary Conversational Speech Recognition. *IEEE Signal Processing Magazine*, September 1999, pp. 84-107
65. Edwards, A. D. N., E. Mitsopoulos, et al. (1999). Discussing mathematics with my computer. *Proceedings of Talking to Computers II*, Sheffield, U.K.
66. Jain A., Hong L., Pankanti S. (2000) Biometric Identification. *Communications of the ACM*, February 2000/Vol 43, No 2, pp. 91-98
67. Joscelyne, A. (1999). The power of speech. *Le Journal*, the journal of record for human language technology. Link [H108].
68. Kalakota R. and Robinson M. (1999). *E-Business: Roadmap for Success*. Addison-Wesley, 1999.
69. Kouroupetroglou G. and Nemeth G. (1995). Speech Technology for Disabled and Elderly People. Chapter in the book "Telecommunications for All", Ed. Patrick Roe, Published by the European Commission - Directorate General XIII, Catalogue number: CD-90-95-712-EN-C, 1995.
70. Κουρουπέτρογλου Γ. (1998). *Μαθήματα Επεξεργασίας Ομιλίας*, Τμήμα Πληροφορικής, Πανεπιστήμιο Αθηνών, 1998.
71. Manaris B. & Harkreader A. (1998): SUITEKeys: A Speech Understanding Interface for the Motor-Control Challenged. In *Proceedings of the Third International ACM SIGCAPH Conference on Assistive Technologies*, April 15-17, 1998, Marina del Rey, CA. USA, pp. 108-115.
72. Markowitz J.A. (1996). *Using Speech Recognition*. Prentice Hall PTR, New Jersey, 1996.
73. Massaro, D.W. (1998) *Perceiving talking faces: from speech perception to a behavioral principle*, A Bradford Book, The MIT Press, Cambridge, Massachusetts, London, England
74. Pitt, I. (1996). *The Principled Design of Speech-Based Interfaces*. Doctoral Thesis. Department of Computer Science. York, University of York.

75. Stevens, R. D. (1996). Principles for the Design of Auditory Interfaces to Present Complex Information to Blind People. Department of Computer Science. York, University of York.
76. Zissman M.A. (1997). Predicting, diagnosing and improving automatic language identification performance. In Proceedings of Eurospeech97, Rhodes, Greece. Vol. 1, pp. 51-54
77. Zue V., Seneff S., Glass J.R., Polifroni J., Pao C., Hazen T.J., Hetherington L. (2000) Jupiter: A Telephone-Based Conversational Interface for Weather Information. IEEE Transactions on Speech and Audio Processing, Vol 8., No 1, January 2000, pp. 85-96

Αναφορες σε Ιστοσελιδες

- H22 συνθέτης NaturallySpeaking
<http://www.dragonsys.com/products/naturallyspeaking/standard/index.html>
- H23 άρθρο <http://www.zdent.com/pcweek/reviews/0504/04lern.html>
- H24 Συνθέτης ομιλίας ViaVoice
<http://www-4.ibm.com/software/speech/millennium/standard/index.html>
- H25 Συνθέτης ομιλίας FreeSpeech 2000
http://www.speech.philips.com/ud/get/Pages/h072lus_heat.htm
- H26 Συνέδριο ASRU'99 <http://asru99.research.att.com>
- H27 Texas Instruments Voice Authorization
<http://www.ti.com/corp/docs/press/company/1995/95301.shtml>
- H28 Texas Instruments <http://www.ti.com>
- H29 Speech at Philips <http://www.speech.philips.com>
- H30 Nuance <http://www.nuance.com>
- H31 Nuance Verifier <http://www.nuance.com/index.htm?SCREEN=verifier>
- H32 Άρθρο για βιομετρικά συστήματα
<http://www.zdnet.com/products/stories/reviews/0,4161,386987,00.html>
- H33 Veritel Voice Verification <http://www.veritelcorp.com>
- H34 Intelitrak Biometric Solutions <http://www.intelitrak.com>
- H35 Language Identification by Sanders http://www.sanders.com/spard/voice/v_li.html
- H36 Έργο ΙΕΛ φωνητικής πρόσβασης σε βάσεις δεδομένων <http://www.ilsp.gr/database.html>
- H37 Centre for Spoken Language Understanding <http://cslu.cse.ogi.edu/>
- H38 Ομάδα Γλωσσικής Τεχνολογίας, Πανεπιστημίου Πατρών
<http://slt.wcl.ee.upatras.gr/moderngreek/index.html>
- H39 Βιομετρική τεχνολογία στην Ελλάδα <http://www.inter.gr/gr/products/DCS/DCS.htm>
- H40 Dialog Communication Systems (βιομετρικές τεχνολογίες) <http://www.bioid.com>
- H41 Συστήματα Γλωσσικής Τεχνολογίας από την Knowledge S.A.
<http://www.knowledge.gr>
- H42 SpeechWorks demos <http://www.speechworks.com/demos/index.cfm>
- H43 AT&T Advanced Speech Products Group <http://www.att.com/aspg>
- H44 Bell Labs text-to-speech <http://www.bell-labs.com/project/tts/voices.html>
- H45 Sensimetrics <http://www.sens.com>
- H46 ETI-Eloquence text-to-speech <http://www.eloq.com>
- H47 Apple text-to-speech <http://www.apple.com/macOS/speech>
- H48 Fonix speech technologies <http://www.acuvoice.com>

- H49 ORATOR speech synthesizer <http://www.bellcore.com/ORATOR>
- H50 Telia text-to-speech <http://www.promotor.telia.se/infovox/listen.htm>
- H51 Άρθρο <http://www.zdnet.co.uk/pcmag/trends/1999/06/realspeak.html>
- H52 Lernout & Hauspie speech technologies <http://www.lhs.com>
- H53 RealSpeak <http://www.lhs.com/realspeak/demo.cfm>
- H54 Ινστιτούτο Επεξεργασίας Λόγου <http://www.ilsp.gr>
- H55 Εκφωνητής <http://www.ilsp.gr/ekfonitis.html>
- H56 Έργο M-PIRO <http://www.ltg.ed.ac.uk/mpiro/>
- H57 sound graphs <http://simsrv.cs.uni-magdeburg.de/~pitt/soundgraph.html>
- H58 verbots <http://vperson.com/>
- H59 AT&T Animated Head <http://www.research.att.com/~osterman/AnimatedHead/>
- H60 ANANOVA <http://www.ananova.com>
- H61 άρθρο <http://www.techweb.com/wire/story/reuters/REU19990506S0003>
- H62 άρθρο http://www.lhs.com/news/releases/20000328_dragon.asp
- H63 Arkenstone SSIL protocol <http://www.arkenstone.org>
- H64 SAM protocol <http://www.vi-access.com/sam.htm>
- H65 Microsoft Speech API <http://www.microsoft.com/IIT/OnlineDocs/intro2SAPI.html>
- H66 XML <http://www.xml.org>
- H67 JSML <http://java.sun.com/products/java-media/speech/forDevelopers/JSML/>
- H68 JSML <http://java.sun.com/products/java-media/speech/forDevelopers/JSML/JSML.html>
- H69 voxML <http://www.voxml.com>
- H70 voxML <http://www.voxml.com/downloads/VoxMLwp.pdf>
- H71 voiceXML <http://www.voicexml.org>
- H72 voiceXML <http://www.voicexml.org/pr20000307-1.html>