



**Πανεπιστήμιο Πελοποννήσου**  
**Σχολή Θετικών Επιστημών και Τεχνολογίας**  
**Τμήμα Επιστήμης και Τεχνολογίας Υπολογιστών**

## **Μεταπτυχιακή Διπλωματική Εργασία**

**Τεχνικές Αλληλεπίδρασης Ανθρώπου-Υπολογιστή:**

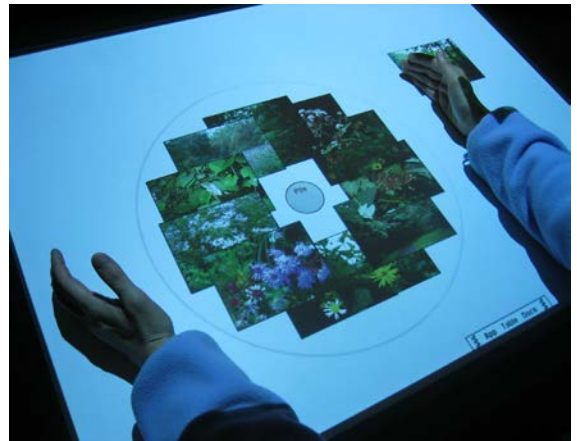
**Διάδραση με χειρονομίες**

**Σπανογιανόπουλος Σωτήριος**

**Επιβλέπων:**

**Λέπουρας Γεώργιος, Μαυρίδης Νικόλαος**

**Τρίπολη, Ιανουάριος 2013**



## ΠΕΡΙΕΧΟΜΕΝΑ

<b>A. Εισαγωγή.....</b>	<b>4</b>
A.1. Επικοινωνία με χειρονομίες.....	5
<i>Εξέλιξη της Ανθρώπινης Επικοινωνίας.....</i>	<i>5</i>
A.2. Διεπαφές με χειρονομίες .....	7
A.3. Αρχές σχεδίασης διεπαφών με χειρονομίες .....	8
- Το άγγιγμα του Μίδα.....	8
- Πολιτισμικά θέματα.....	8
- Σειρά εκτέλεσης επιμέρους ενεργειών.....	9
- Σχεδίαση με βάση το πλαίσιο χρήσης.....	9
- Ανάδραση.....	9
- Ολοκλήρωση και Αποσαφήνιση .....	9
<i>Σημαντικές εφαρμογές και παραδείγματα.....</i>	<i>9</i>
<i>Βασικά τεχνολογικά ζητήματα και περιορισμοί.....</i>	<i>10</i>
A.4. Τεχνικές αλληλεπίδρασης: δυνατότητες και περιορισμοί.....	11
<b>B. Συστήματα Διάδρασης με Χειρονομίες (εφαρμογές σε οικιακό περιβάλλον).....</b>	<b>12</b>
B.1. Εισαγωγή.....	12
B.2. Ιστορική Αναδρομή .....	13
B.3. Μέθοδοι σχεδίασης και αξιολόγησης.....	15
B.4. Συστήματα Αλληλεπίδρασης με Ανθρώπους και Διεπαφές που Βασίζονται στις Χειρονομίες .....	16
<b>Γ. Συγκριτική Αξιολόγηση Τεχνικών και Συμπεράσματα.....</b>	<b>43</b>
<b>Δ. Επίλογος και Μελλοντική Εργασία .....</b>	<b>46</b>
<b>Ε. Βιβλιογραφικές Πηγές.....</b>	<b>48</b>

## A. Εισαγωγή

Η συστηματική μελέτη της απόδοσης του ανθρώπου ξεκίνησε στις αρχές του εικοστού αιώνα, στα εργοστάσια, δίνοντας έμφαση σε χειρονακτικές εργασίες. Ο Β' Παγκόσμιος Πόλεμος αποτέλεσε το έναυσμα για την μελέτη της επικοινωνίας ανθρώπου-μηχανής, καθώς οι αντιμαχόμενες πλευρές προσπαθούσαν να παράγουν όσον το δυνατόν πιο αποτελεσματικά οπλικά συστήματα. Το 1949 ιδρύθηκε η Ερευνητική Κοινότητα για την Εργονομία. Ο όρος εργονομία αναφέρεται στην απόδοση του χρήστη στο περιβάλλον οποιουδήποτε συστήματος, υπολογιστικού, μηχανικού ή χειροκίνητου. Με την ευρεία διάδοση των υπολογιστών, ένας διαρκώς αυξανόμενος αριθμός ερευνητών επικεντρώθηκαν στην μελέτη της επικοινωνίας της διάδρασης (interaction) μεταξύ ανθρώπων και υπολογιστών, εξετάζοντας τις φυσικές, σωματικές και ψυχολογικές εκφάνσεις αυτής της διαδικασίας. Η έρευνα αυτή ξεκίνησε με τον όρο «επικοινωνία ανθρώπου - μηχανής» αλλά λόγω του ιδιαίτερου ενδιαφέροντος και της ανάπτυξης των υπολογιστών μετονομάστηκε σε «επικοινωνία ανθρώπου – υπολογιστή».



Η επικοινωνία ανθρώπου-υπολογιστή άπτεται πολλών πεδίων αλλά ειδικά στην σχεδίαση υπολογιστικών συστημάτων αποτελεί ζήτημα κεφαλαιώδους σημασίας. Για όλα τα άλλα πεδία (όπως για παράδειγμα στη Ψυχολογία ή την Ιατρική) μπορεί να θεωρηθεί ως ένας εξειδικευμένος κλάδος της διαδικασίας σχεδίασης παρά τις πληροφορίες ζωτικής σημασίας που μπορεί να παρέχει. Στη σχεδίαση όμως υπολογιστικών συστημάτων πρέπει να αποτελεί αναπόσπαστο κομμάτι της διαδικασίας αυτής. Από αυτήν την άποψη λοιπόν, η επικοινωνία ανθρώπου-υπολογιστή περιλαμβάνει τη σχεδίαση, την υλοποίηση και την αξιολόγηση διαδραστικών συστημάτων σε σχέση με τον χρήστη.

Όταν μιλάμε για «επικοινωνία ανθρώπου-υπολογιστή», δεν αναφερόμαστε απαραίτητα σε ένα μεμονωμένο χρήστη με ένα προσωπικό υπολογιστή. Με τον όρο χρήστη (user) μπορεί να εννοούμε ένα μεμονωμένο χρήστη, μια ομάδα χρηστών που δουλεύουν μαζί ή μια αλληλουχία χρηστών, καθένας εκ των οποίων ασχολείται με κάποιο επιμέρους θέμα μιας εργασίας. Με τον όρο υπολογιστής (computer) αναφερόμαστε σε οποιαδήποτε τεχνολογία υπολογιστών – από τους προσωπικούς υπολογιστές γενικής χρήσης, μέχρι τα μεγάλης κλίμακας υπολογιστικά συστήματα, τα συστήματα ελέγχου διεργασιών (process control systems), ή τα ενσωματωμένα συστήματα (embedded systems). Το σύστημα μπορεί να περιλαμβάνει και άλλα μέρη εκτός από τους υπολογιστές. Τέλος με τον όρο διάδραση (interaction) εννοούμε οποιαδήποτε επικοινωνία μεταξύ ενός χρήστη και ενός υπολογιστή ή μιας μηχανής, άμεση ή έμμεση. Η άμεση διάδραση συνεπάγεται ένα διάλογο με ανάδραση και έλεγχο καθ' όλη τη διάρκεια της εκτέλεσης της εργασίας. Η έμμεση διάδραση μπορεί να περιλαμβάνει μαζική επεξεργασία (πχ. με τη διαμεσολάβηση συγκεκριμένων προϊόντων που προσδιορίζουν με την αποδοχή ή την απόρριψή τους τη σχέση χρήστη και στόχου) ή επεξεργασία στο παρασκήνιο. Το σημαντικό γεγονός όμως της όλης διαδικασίας είναι ότι ο χρήστης αλληλεπιδρά/επικοινωνεί με τον υπολογιστή ή τη μηχανή προκειμένου να πετύχει κάποιο στόχο.

Επιγραμματικά, το σημαντικότερο ζήτημα, όσον αφορά την σχεδίαση ενός συστήματος επικοινωνίας ανθρώπου-μηχανής είναι η τοποθέτηση του χρήστη στη βάση της σχεδίασης του συστήματος. Οι ανάγκες, προτιμήσεις και ικανότητες των χρηστών κατά τη λειτουργία ενός συστήματος, πρέπει να είναι οι παράγοντες που θα κατευθύνουν τη σχεδίαση των συστημάτων αυτών. Δεν θα πρέπει οι χρήστες να αλλάζουν τον τρόπο με τον οποίο χρησιμοποιούν ένα σύστημα, προκειμένου να μπορέσουν να χρησιμοποιήσουν ένα τέτοιο σύστημα. Αντίθετα, ιδανικά το σύστημα θα πρέπει να είναι σχεδιασμένο με τέτοιο τρόπο, ώστε να ικανοποιεί τις απαιτήσεις τους (Αβούρης, 2000).

Για την σχεδίαση ενός διαδραστικού συστήματος απαιτούνται γνώσεις από πολλά επιστημονικά πεδία: την ψυχολογία και τη γνωστική επιστήμη (γνώση για την αντίληψη του χρήστη, την γνωστική του ικανότητα και την ικανότητα επίλυσης προβλημάτων), την εργονομία (γνώση για τις φυσικές ικανότητες του χρήστη), την κοινωνιολογία (κατανόηση του ευρύτερου πλαισίου μέσα στο οποίο λαμβάνει χώρα η διάδραση), την επιστήμη των υπολογιστών (ανάπτυξη της απαιτούμενης τεχνολογίας), τη σχεδίαση γραφικών (για την δημιουργία μιας αποτελεσματικής παρουσίασης της διεπιφάνειας), την συγγραφή τεχνικών κειμένων (για την σύνταξη των συνοδευτικών εγχειριδίων του συστήματος) κ.α. Για το λόγο αυτό, παρόλο που ο όρος επικοινωνία ανθρώπου-μηχανής περιλαμβάνει όλα τα ανωτέρω μέρη, στην πράξη υπάρχουν εξειδικευμένοι επιστήμονες που επικεντρώνουν το ενδιαφέρον τους σε καθένα από τους παραπάνω τομείς. Έτσι, το τελικό

σύστημα, αποτελεί το προϊόν της συνεργασίας εξειδικευμένων επιστημόνων όλων των παραπάνω κλάδων (Βολουδάκης και Μαρμακόπος, 2005).

## **A.1. Επικοινωνία με χειρονομίες**

### *Εξέλιξη της Ανθρώπινης Επικοινωνίας*

Οι χειρονομίες είναι προϊόν της αλληλεπίδρασης μεταξύ των ανθρώπων και αποτελούνται από κινήσεις του σώματος και του προσώπου, οι οποίες αντικαθιστούν την προφορική επικοινωνία. Οι χειρονομίες είναι απαραίτητες για τη φυσική και παραστατική ανθρώπινη επικοινωνία και προσθέτουν έμφαση στην ομιλία τόσο οπτικά όσο και συναισθηματικά. Μπορεί να διεγείρουν τα συναισθήματά, και ως εκ τούτου να δώσουν ζωντάνια στη φωνή.

Σε μερικούς πολιτισμούς οι άνθρωποι χειρονομούν πιο ελεύθερα από άλλους που έχουν διαφορετική προέλευση. Εντούτοις, όταν μιλούν, ουσιαστικά όλοι οι άνθρωποι αλλάζουν την έκφραση του προσώπου τους και κάνουν κάποιου είδους χειρονομίες. Επίσης, οι ιδέες και τα συναισθήματα δεν μεταδίδονται μόνο με τη φωνή, αλλά και με χειρονομίες και εκφράσεις του προσώπου. Γενικά, οι χειρονομίες υπάγονται σε δύο γενικές κατηγορίες: στις περιγραφικές και τις εμφατικές. Οι περιγραφικές χειρονομίες εκφράζουν ενέργεια ή δείχνουν διαστάσεις και τόπο. Οι εμφατικές χειρονομίες εκφράζουν συναισθήματα και πεποίθηση. Τονίζουν τις ιδέες, τις ζωντανεύουν και τους δίνουν επιπρόσθετη ισχύ. Οι εμφατικές χειρονομίες είναι ουσιαστικές, αλλά μπορούν εύκολα να μετατραπούν σε ιδιόρρυθμο χαρακτηριστικό της ομιλίας.

Η κύρια ταξινόμηση των χειρονομιών γίνεται με βάση τη σημασία, τη λειτουργία και την περιγραφή τους. Η ταξινόμιά τους βέβαια προχωρά σε μεγάλο βάθος διακρίνοντας πολλές κατηγορίες χειρονομιών που μπορούν να χρησιμοποιηθούν σε αντίστοιχες διεπαφές και επαφές.

Οι περισσότερες χειρονομίες δεν είναι ούτε φυσικές ούτε είναι εύκολο να τις μάθει κανείς και να τις θυμάται. Ακόμα και το απλό κούνημα του κεφαλιού είναι πολύπλοκο όταν περιπλέκονται πολλοί πολιτισμοί. Για παράδειγμα, οι Δυτικοί που ταξιδεύουν στην Ινδία δυσκολεύονται να κατανοήσουν τον ινδικό τρόπο κουνήματος του κεφαλιού ο οποίος είναι μια διαγώνια μείξη του δυτικού κάθετου κουνήματος που σημαίνει «ναι» και του οριζόντιου κουνήματος που σημαίνει «όχι». Παρομοίως, οι χειρονομίες κουνήματος του χεριού για να πούμε «γεια», «αντίο», «έλα εδώ» κ.α. πραγματοποιούνται διαφορετικά σε διαφορετικές κουλτούρες (Αναφορές).

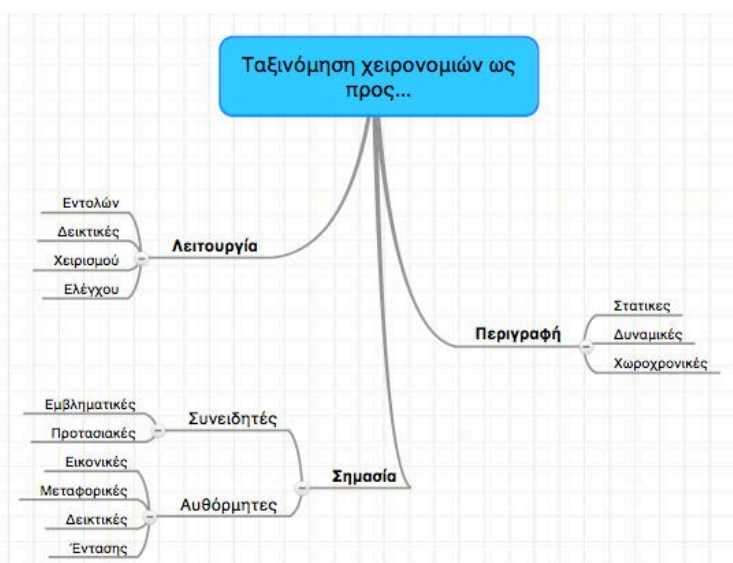
Ο Kimura (1993) περιγράφει ότι υπάρχουν ενδείξεις ότι οι κινήσεις (και κυρίως των χεριών) βοήθησαν στην εξέλιξη της ομιλίας και του συστήματος επικοινωνίας μεταξύ των ανθρώπων. Η έρευνα του δείχνει ότι η μοντελοποίηση των χειρονομιών και του λόγου δημιουργήθηκαν σαν φόρμες έκφρασης μέσω ενός συστήματος όπου η προφορική επικοινωνία δεν αποτελεί τον κύριο λόγο από τον οποίο δημιουργήθηκαν οι χειρονομίες.

Γι' αυτό το λόγο πολλοί έχουν μελετήσει τη σχέση επικοινωνίας λόγου και χειρονομιών, όπως ο Kendon (1988), ο οποίος ισχυρίζεται ότι οι χειρονομίες εμπεριέχουν μέσα τους μια «προφορικότητα». Σημειώνετε επίσης ότι τα προφορικά στοιχεία της επικοινωνίας που υπάρχουν σε έναν λόγο αντικαθίστανται πολλές φορές από νοήματα ή χειρονομίες, κάτι που ενισχύει και την προηγούμενη άποψη ότι οι χειρονομίες και η ομιλία δημιουργήθηκαν από ένα ολοκληρωμένο ανθρώπινο σύστημα επικοινωνίας.

Επίσης, ο McNeill (1992) συμπέρανε ότι δεν υπάρχει “γλώσσα” του σώματος, αλλά χειρονομίες που συμπληρώνουν τον προφορικό λόγο. Τέτοιες χειρονομίες εμπεριέχουν επιπλέον πληροφορίες, πέρα από το λόγο, σχετικά με τις εσωτερικές προθέσεις του ομιλητή, αλλά δυστυχώς δεν μπόρεσε να τις τυποποιήσει και να τις αναλύσει, ώστε να εξάγει κάποια συμπεράσματα για τη σχέση και τη σύνδεση τους με τον προφορικό λόγο. Πιο συγκεκριμένα αναφέρει ότι οι χειρονομίες είναι ένα κομμάτι ενσωματωμένο με τη γλώσσα όπως οι λέξεις, οι φράσεις και οι προτάσεις. Βασικά ισχυρίστηκε ότι οι χειρονομίες και η γλώσσα είναι ένα σύστημα. Για να το αποδείξει έκανε ένα πείραμα στο οποίο ένας ομιλητής θα έπρεπε να περιγράψει μια μικρή ιστορία (ή ένα παραμύθι) σε κάποιον ακροατή, και αυτός με τη σειρά του θα έπρεπε στη συνέχεια να περιγράψει αυτά που άκουσε σε ένα τρίτο πρόσωπο. Κατά τη διάρκεια της περιγραφής αυτής της ιστορίας, υπήρχαν κάμερες και μικρόφωνα που τους κατέγραφαν. Καθώς δεν ήξεραν τον πραγματικό σκοπό αυτής της

διαδικασίας, όλοι τους κατά τη διάρκεια της αφήγησης χρησιμοποιούσαν ανεπαίσθητα κάποιες χειρονομίες, χωρίς όμως να γνωρίζουν ότι σκοπός αυτού του πειράματος ήταν η μελέτη των χειρονομιών που θα έκαναν. Μετά την καταγραφή των κινήσεών τους, ο McNeill έκανε ανάλυση τους και τις κατηγοριοποίησε σε τέσσερις κατηγορίες, ανάλογα με τον τύπο των χειρονομιών. Όλοι οι τύποι των κινήσεων ήταν συμβολικοί, δηλαδή αναφερόντουσαν στο γεγονός ότι το χέρι θα κινηθεί κάνοντας κάτι που θα προσπαθεί να υποδείξει αυτό που προσπαθεί να πει ο ομιλητής, χωρίς όμως η κίνηση αυτή να σχετίζεται με τη φυσική λειτουργία (semantic or pragmatic) που θα πρέπει να διατελλεί το χέρι. Πιο συγκεκριμένα, οι εικονικές κινήσεις/χειρονομίες είναι πιο περιεκτικές και προσπαθούσαν να τονίσουν αυτό που βρίσκει ο χρήστης σιωπηλό σε κάποια περίπτωση ή που θέλει να περιγράψει σε αυτόν που αναφέρεται. Με τις μεταφορικές χειρονομίες ο χρήστης πάλι προσπαθεί να περιγράψει μια εικόνα, αλλά πιο σύντομα και περιεκτικά (abstract). Με τις χειρονομίες τύπου beats ο χρήστης κάνει την κίνηση σε δύο φάσεις ρυθμικά, και η οποία κίνηση πολλές φορές τείνει να επαναληφθεί. Τέλος, με τις χειρονομίες τύπου deictics μπορούν να είναι είτε σύντομες (abstract) είτε αρκετά περιεκτικές (concrete). Σκοπός φυσικά αυτής της κατηγοριοποίησης και μελέτης ήταν να ενισχύσει τη θεωρία του ότι η ομιλία και οι χειρονομίες, παρόλο που προέρχονται από διαφορετικά κέντρα ελέγχου του εγκεφάλου, ουσιαστικά συσχετίζονται μεταξύ τους και αποτελούν ένα ολοκληρωμένο σύστημα έκφρασης. Αυτό που δεν μπόρεσε να αποδείξει είναι ότι οι χειρονομίες που κάνουν οι άνθρωποι κατά την ομιλία τους δεν έχουν “προφορικές” ιδιότητες, δηλαδή ιδιότητες που σχετίζονται με το περιεχόμενο των όσων λέει ο ομιλητής, και κατά συνέπεια δεν μπορούν να ταξινομηθούν ή να κατηγοριοποιηθούν ανάλογα με το περιεχόμενο του προφορικού λόγου.

Επομένως, οι χειρονομίες αποτελούν έναν πανίσχυρο τρόπο επικοινωνίας. Η βασική διαφορά ανάμεσα στα συστήματα του σήμερα και αυτά που αναπτύχθηκαν τα τελευταία 50 χρόνια είναι η ανάπτυξη των ισχυρών και φθηνών τεχνολογιών για τους αισθητήρες και την επεξεργασία, οι οποίες κάνουν πλέον δυνατή την εφαρμογή αυτών των συστημάτων σε φθηνές συσκευές μαζικής παραγωγής. Έχουμε ήδη δει μεγάλες προόδους στην χρήση τους. Οι χειρονομίες θα προτυποποιηθούν και έτσι θα υπάρχει μια βάση για να στηριχθούν όλες οι εφαρμογές που χρησιμοποιούν χειρονομίες με αποτέλεσμα να διευκολύνεται έτσι και οι χρήση τους, καθώς δεν θα απαιτείται η απομνημόνευση διαφορετικών χειρονομιών για κάθε συσκευή που χρησιμοποιούμε. Για παράδειγμα το κούνημα μια συσκευής αρχίζει να ερμηνεύεται παντού ως «δώσε κάτι εναλλακτικό». Το οριζόντιο κούνημα των δαχτύλων σημαίνει πάνε σε μια νέα σελίδα (π.χ. iPhone). Το τράβηγμα 2 δαχτύλων συνεπάγεται την μεγέθυνση μια εικόνας και το αντίθετο. Αυτές οι χειρονομίες ήδη χρησιμοποιούνται σε πολλά εμπορικά συστήματα που διαθέτουν διεπαφή χειρονομιών. Η μέχρι τώρα εμπειρία στον σχεδιασμό διεπαφών μπορεί να βοηθήσει σημαντικά στην προτυποποίηση των χειρονομιών, με αποτέλεσμα οι μελλοντικές εφαρμογές Αλληλεπίδρασης Ανθρώπου-Υπολογιστή να βασίζονται σε ένα κοινό μοντέλο χειρονομιών, δημιουργώντας τις προϋποθέσεις μιας τυποποίησης των κινήσεων/χειρονομιών των χεριών.



Εικόνα 1. Ταξινόμηση χειρονομιών. (Κουτσαμπάσης, Π. (2011) )

## A.2. Διεπαφές με χειρονομίες

Η Αλληλεπίδραση Ανθρώπου – Υπολογιστή με διεπαφές είναι μια καινούργια πρόκληση στην βιομηχανία τις τελευταίες δεκαετίες. Η πρόδος που έχει γίνει στο μέγεθος, την κατανάλωση ενέργειας και το κόστος των επεξεργαστών, των μνημών, των καμερών και άλλων αισθητηριακών συσκευών (πχ. στα γάντια ανίχνευσης χειρονομιών ή στα συστήματα/αισθητήρια οπτικής αναγνώρισης) έχει ανοίξει τον δρόμο στον έλεγχο τους με χειρονομίες και κινήσεις. Οι καινούργιες αλληλεπιδράσεις χαρακτηρίζονται ως «φυσικές» με τις αντίστοιχες «Φυσικές Διεπαφές Χρήστη».

Οι διεπαφές με χειρονομίες διακρίνονται στις μηχανικές και απτικές διεπαφές, στις διεπαφές «υπολογιστικής όρασης» και στις διεπαφές που βασίζονται σε εκφράσεις του προσώπου (Αβούρης, 2000). Ο τρόπος που επικοινωνούμε με αυτές τις διεπαφές χαρακτηρίζεται ως τεχνική αλληλεπίδρασης και αφορά την χειρονομία που κάνουμε και την απόκριση του συστήματος σε αυτή. Οι τεχνικές αλληλεπίδρασης που χρησιμοποιούνται στις διεπαφές με χειρονομίες σχετίζονται π.χ. με την εγγύτητα, την κίνηση του σώματος, το δείξιμο, τον κυματισμό του χεριού, την εισαγωγή των χεριών, την περιστροφή αντικειμένου, το πάτημα με το πόδι, το ταρακούνημα και την κλίση του σώματος.

Για την δημιουργία ενός λεξικού χειρονομιών, δηλαδή ενός συνόλου χειρονομιών οι οποίες μπορούν να χρησιμοποιηθούν σε μια τέτοια διεπαφή, πρέπει αρχικά να ελέγξουμε άλλες εφαρμογές και να πάρουμε ιδέες από αυτές αλλά και να συμβουλευτούμε την γνώμη χρηστών. Στη συνέχεια οι χειρονομίες που επιλέξαμε πρέπει να εξεταστούν από χρήστες έτσι ώστε να διαπιστώσουμε κατά πόσο είναι κατάλληλες για τη διεπαφή μας (Κουτσαμπάσης, 2011).

Η σχεδίαση των διεπαφών με χειρονομίες πρέπει να λαμβάνει πάντα υπόψη την φυσιολογία του ανθρώπινου σώματος έτσι ώστε να αποφεύγεται η δημιουργία δύσχρηστων και κουραστικών διεπαφών. Επίσης πρέπει να ακολουθούνται κάποιοι κανόνες που προάγουν την ευκολία χρήσης και την αποτελεσματικότητα των διεπαφών. Η σχεδίαση πρέπει να οδηγείται από την ιδέα να δημιουργήσουμε διεπαφές με τις οποίες ο χρήστης θα μπορεί να διαδρά αυθόρμητα χωρίς επεξηγήσεις.

Οι διεπαφές αυτές με χειρονομίες αποτελούν ένα μεγάλο πεδίο έρευνας και ανάπτυξης στην Αλληλεπίδραση Ανθρώπου – Υπολογιστή Ανθρώπου-Μηχανής, και περιλαμβάνουν Αλληλεπίδραση Ανθρώπου – Υπολογιστή με

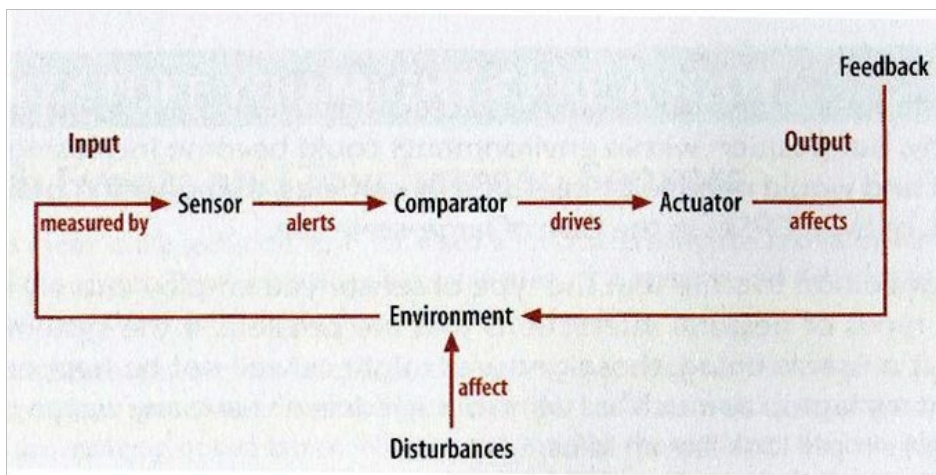
- (α) κινήσεις δακτύλων – ‘νοηματική’,
- (β) κινήσεις χεριών,
- (γ) κινήσεις- στάσεις σώματος (postures),
- (β) κινήσεις ποδιών (πίεση, προσανατολισμός, βηματισμός τρέξιμο, κλπ),
- (ε) εκφράσεις προσώπου,
- (στ) βλέμμα (gaze interaction),
- (ζ) γραφίδα (pen-based interfaces), κ.α.

Στη διεθνή βιβλιογραφία που σχετίζεται με την Αλληλεπίδραση Ανθρώπου – Υπολογιστή Ανθρώπου-Μηχανής (ή υπολογιστών) ο όρος χειρονομία έχει χρησιμοποιηθεί για πολλούς τύπους κινήσεων του χεριού για τον έλεγχο των διάφορων διαδικασιών των υπολογιστών. Γι’ αυτό ίσως και ο Sturman (1993), για να αποφύγει τη σύγχυση, ορίζει την είσοδο ολόκληρου του χεριού (whole-hand input) σαν την πλήρη και απευθείας χρήση των ιδιοτήτων του χεριού για τον έλεγχο των διαδικασιών των υπολογιστών, κάνοντας ταυτόχρονα και έναν πιο ακριβή προσδιορισμό ποιών τύπων οι κινήσεις των χεριών θα πρέπει να χρησιμοποιούνται (πχ. όχι μόνο η θέση, αλλά και το σχήμα του χεριού), καθώς επίσης και για πιο λόγο θα πρέπει να εφαρμόζονται.

Πολλές εφαρμογές υλοποιούν συνήθως τη χρήση συγκεκριμένων σημάτων των χεριών ή διάφορες συγκεκριμένες θέσεις του χεριού. Σε μια άλλη έρευνα, ο Rook (1995) παρουσιάζει τη χρήση σημάτων χεριών τα οποία δείχνουν σε ένα ρομπότ πως να εκτελεί κάποιες τυποποιημένες κινήσεις για να ολοκληρώσει μια διαδικασία. Η συγκεκριμένη δουλειά πλέον τίθεται υπό ερωτηματικό, καθώς ίσως να ήταν πιο εύκολο να υλοποιήσεις ένα αντίστοιχο σύστημα ελέγχου του ρομπότ χρησιμοποιώντας μόνον το πληκτρολόγιο. Το νόημα πάντως αυτής της δουλειάς (όπως και άλλων παραπλήσιων) είναι ότι αφήνει το χρήστη να δράσει πιο ελεύθερα και να ελέγξει τη μηχανή (εδώ ρομπότ) κάνοντας απλές κινήσεις των

χειριών. Γι'αυτό και αυτές οι έρευνες αποτέλεσαν το έναυσμα για να ξεκινήσει και να επεκταθεί η έρευνα στο συγκεκριμένο επιστημονικό χώρο (Mavridis et. al., 2012).

Επίσης, η εργασία του Myers (1998) κάνει μια ανασκόπηση σε συστήματα που κάνουν χρήση χειρονομιών τα οποία ήταν εμπορικά από το 1960. Στην πρωτοποριακή εργασία του Kueger (1983) σχετικά με την εικονική πραγματικότητα, γίνεται μια εισαγωγή στην Αλληλεπίδρασης Ανθρώπου – Υπολογιστή με την χρήση χειρονομιών πάνω σε μεγάλες προβαλλόμενες εικόνες. Τα συστήματα πολλαπλών σημείων επιλογής (multi-touch systems) άρχισαν να εμφανίζονται την δεκαετία του 1980 και όπως αναφέρεται στην εργασία του Buxton (2007) το πρώτο τέτοιο σύστημα δημιουργήθηκε το 1982. Εξειδικευμένοι αισθητήρες για τον εντοπισμό την ανθρώπινης θέσης και κίνησης εξετάζονται εδώ και πολύ καιρό από την βιομηχανία των παιχνιδιών. Επίσης μουσικά όργανα που χειρίζονται με χειρονομίες και πολύ-απτικές συσκευές εισόδου μουσικής αποτελούν πεδίο έρευνας και εμπορικής εφαρμογής εδώ και μισό αιώνα περίπου, καθώς το Theremin, ένα ηλεκτρονικό synthesizer που κάνει χρήση χειρονομιών προτυποποιήθηκε από τον δημιουργό του το 1928.



Εικόνα 2. Γενική περιγραφή ενός συστήματος διεπαφής με χειρονομίες.

### A.3. Αρχές σχεδίασης διεπαφών με χειρονομίες

Στα χρόνια που ακολούθησαν, έγιναν πολλές μελέτες για το πως θα πρέπει να σχεδιάζονται οι διάφορες διεπαφές που χρησιμοποιούνται στα συστήματα αλληλεπίδρασης με τον άνθρωπο. Με βάση τις μελέτες αυτές προέκυψαν τα παρακάτω μοντέλα και αρχές σχεδίασης διεπαφών (Κουτσαμπάσης, 2011).

#### -Το άγγιγμα του Μίδα

Η διεπαφή πρέπει να καταλαβαίνει πότε να αρχίζει και πότε να σταματά την ερμηνεία μιας χειρονομίας. Μπορούμε να καθορίσουμε κάποιες χειρονομίες ως εναρκτήριες και καταληκτικές για μια σύνοδο όπου το σύστημα θα αναγνωρίζει τις κινήσεις. Όπως ο Μίδας δεν ήθελε ότι άγγιζε να γίνεται χρυσός, έτσι κι εδώ δεν πρέπει όποτε κινούμαστε να ενεργοποιείται το σύστημα.

#### -Πολιτισμικά θέματα

Είναι γνωστό πως η μη προφορική επικοινωνία εξαρτάται από πολιτισμικά στοιχεία. Ο πολιτισμός μπορεί να επηρεάζει την σημασία της χειρονομίας, τον ρυθμό εκτέλεσης της και την ταχύτητά της. Όταν μια χειρονομία δεν ταιριάζει σε κάποιον πολιτισμό τότε μπορεί να προσαρμοστεί η διεπαφή ανάλογα με τον χρήστη και να επιλέγει ο ίδιος τις χειρονομίες που του ταιριάζουν για να χρησιμοποιεί.



### - Σειρά εκτέλεσης επιμέρους ενεργειών

Καθορίζεται το κάθε βήμα που πρέπει να εκτελέσει ο χρήστης στην διεπαφή για να φέρει σε πέρας μια συγκεκριμένη διεργασία.

Παράδειγμα:

1. Ο χρήστης κάνει μια χειρονομία για εισαγωγή και η χειρονομία εκτελείται στο σημείο τοποθέτησης του αντικειμένου.
2. Η διεπαφή επιστρέφει ένα σύνολο από αντικείμενα για εισαγωγή.
3. Ο χρήστης κάνει μια χειρονομία επιλογής και διαλέγει το αντικείμενο που θέλει.
4. Η διεπαφή αφαιρεί τη λίστα αντικειμένων και εισάγει το αντικείμενο που επιλέχθηκε.

### - Σχεδίαση με βάση το πλαίσιο χρήσης

Η εφαρμογή θα πρέπει να ερμηνεύει τις χειρονομίες σε σχέση με το πλαίσιο χρήσης.

1. Χωρικές ζώνες: Ο χώρος γύρω από τον χρήστη μπορεί να χωριστεί σε διάφορες ζώνες διάδρασης, όπως χώρος μενού, χώρος εργασίας, χώρος εντολών κ.α. Η θέση στην οποία γίνεται μια χειρονομία καθορίζει και το αποτέλεσμα της.
2. Τρόπος χρήσης: Μπορούν να επαναχρησιμοποιηθούν χειρονομίες με διαφορετικούς τρόπους. Όπως για παράδειγμα το ποντίκι στο Photoshop όπου ανάλογα με τον τρόπο που επιλέγουμε να χρησιμοποιηθεί μπορεί να ζωγραφίσει, να αποκόψει μέρος της εικόνας, κ.α.

### - Ανάδραση

Η χρήση χειρονομιών για την επικοινωνία με το σύστημα μπορεί να κάνει τον χρήστη να νιώσει αποκομμένος. Πρέπει επομένως ο χρήστης να τροφοδοτηθεί με ανάδραση η οποία θα τον πληροφορεί ότι το σύστημα καταλαβαίνει την παρουσία και τις κινήσεις του. Η ανάδραση του συστήματος προς τον χρήστη είναι συχνά καλύτερο να μην είναι μόνο οπτική αλλά και ακουστική και απτική.

### - Ολοκλήρωση και Αποσαφήνιση

Η σχεδίαση μιας αποκλειστικά και μόνο διεπαφής χειρονομιών είναι δύσκολη εργασία. Η ταξινόμια συνεργασίας όταν έχουμε πολλές τροπικότητες φαίνεται παρακάτω:

1. Μεταφορά: Πρώτα εκτελείται η μια τροπικότητα και μετά η άλλη.
2. Ισοδυναμία: Χρήση οποιασδήποτε τροπικότητας θα φέρει το ίδιο αποτέλεσμα.
3. Συμπληρωματικότητα: Βάλε αυτό εκεί (Λές «βάλε» δείχνοντας το αντικείμενο και μετακινείς το χέρι σου στην τοποθεσία και λες «εκεί» για να ολοκληρωθεί η διαδικασία).
4. Πλεονασμός: Πολλές ταυτόχρονες τροπικότητες.
5. Ειδικότητα: Μόνο μια τροπικότητα έχει πρόσβαση σε κάποια λειτουργία.

### Σημαντικές εφαρμογές και παραδείγματα

Δύο πολύ ενδιαφέρουσες εφαρμογές κιναισθητικού ελέγχου είναι οι κονσόλες παιχνιδιών Nintendo Wii και Xbox Kinect (Microsoft). Και οι δύο περιπτώσεις αφορούν ως επί το πλείστον εφαρμογές οικογενειακών αθλητικών παιχνιδιών και γυμναστικής όπου ο χρήστης κινεί τον εικονικό χαρακτήρα με την κίνηση των χεριών και του σώματος του. Στην πρώτη περίπτωση χρησιμοποιούνται τα Wii Remote (Wiimote) και Wii Fit. Το Wiimote (συντόμευση του Wii Remote) είναι ένα χειριστήριο με ιδιαίτερα προηγμένες δυνατότητες: περιλαμβάνει επιταχυνσίμετρο (accelerometer), υπέρυθη κάμερα, μεγάφωνο, αισθητήρες αναγνώρισης κίνησης και προσανατολισμού, δυνατότητα δείξιματος (pointing) και δόνησης, σύνδεση με Bluetooth και μπορεί να ενισχυθεί με επισυναπτόμενα στοιχεία (attachments). Το Wii Fit είναι μια πλατφόρμα με αισθητήρες πίεσης που υπολογίζει σε πραγματικό χρόνο την κατανομή του βάρους του χρήστη και την πίεση που ασκεί με το σώμα του στη πλατφόρμα και εκτιμά την στάση του σώματός του (posture). Στην δεύτερη περίπτωση, 'η διεπαφή εξαφανίζεται', αφού δεν χρησιμοποιείται κανένα χειριστήριο – η πλατφόρμα αναγνωρίζει μέσω της κάμερας και κατάλληλου λογισμικού σε πραγματικό χρόνο τις κινήσεις του σώματος. Περισσότερες λεπτομέρειες για τα συστήματα αυτά παρουσιάζονται σε επόμενη ενότητα.

## Βασικά τεχνολογικά ζητήματα και περιορισμοί

Η πλέον ενδιαφέρουσα προοπτική της κιναισθητικής αλληλεπίδρασης αφορά τον έλεγχο των εφαρμογών με ελάχιστα ή καθόλου χειριστήρια και χωρίς κάποιας μορφής φυσική σύνδεση με τον υπολογιστή (accessory-free). Οι τεχνολογίες που μπορούν να υποστηρίξουν την κιναισθητική Αλληλεπίδρασης Ανθρώπου – Υπολογιστή μπορούν να διακριθούν στις εξής:

1. Τεχνολογίες υπολογιστικής όρασης
2. Τεχνολογίες ασύρματων δικτύων
3. Τεχνολογίες αισθητήρων

Οι τεχνολογίες υπολογιστικής όρασης (computer vision) μπορούν να υποστηρίξουν κιναισθητική Αλληλεπίδρασης Ανθρώπου – Υπολογιστή χωρίς χειριστήρια. Ένα σύστημα υπολογιστικής όρασης για υποστήριξη κιναισθητικού ελέγχου αποτελείται από (α) μια ή περισσότερες κάμερες που καταγράφουν τις κινήσεις του χρήστη, (β) ενδεχομένως υποστηρικτικό φωτισμό (συχνά στο υπέρυθρο φάσμα), και κυρίως (γ) λογισμικό αναγνώρισης κινήσεων και χειρονομιών. Οι αλγόριθμοι υπολογιστικής όρασης γενικά αποτελούνται από τρία μέρη (Kortum, 2008):

(α) τμηματοποίηση (segmentation): εδώ η εικόνα διαχωρίζεται σε διακριτά μέρη με στόχο την αναγνώριση (στοιχείων) του ανθρωπίνου σώματος,

(β) καταγραφή (tracking) των κινήσεων (ή/και χειρονομιών) των μερών του ανθρωπίνου σώματος,

(γ) αντιστοιχίση (classification) των κινήσεων (ή/και χειρονομιών) με αυτές που έχουν νόημα για την εφαρμογή.

Βεβαίως, κάθε λογισμικό υπολογιστικής όρασης μπορεί να εστιάζει σε διαφορετικές πτυχές της κιναισθητικής αλληλεπίδρασης. Έτσι, π.χ. υπάρχουν λογισμικά που εστιάζουν στην ανιχνεύση χειρονομιών και κινήσεων χεριών όπως το g-speak, και άλλα που καταγράφουν όλο το σώμα σε δύο και τρεις διαστάσεις (3D tracking).

Όταν η χρήση χειριστηρίων ή άλλων φορητών συσκευών είναι απαραίτητη, τότε η επικοινωνία με τον υπολογιστή θα πρέπει να υποστηρίζεται από ασύρματες συνδέσεις όπως wi-fi, Bluetooth, RFID (Radio Frequency Identification), κλπ. Ο τύπος της ασύρματης σύνδεσης εξαρτάται από πολλούς παράγοντες όπως, π.χ. το εάν απαιτείται οπτική επαφή ή όχι (το Bluetooth απαιτεί οπτική επαφή), την απόσταση μεταξύ πομπού και δέκτη, το επίπεδο ασφαλείας των δεδομένων που μεταδίδονται, κ.α.

Οι πιο συχνοί τύποι αισθητήρων για διεπαφές χειρονομιών ανιχνεύουν:

- Πίεση – ανιχνεύουν ότι κάτι πιέστηκε/πατήθηκε (και ενδεχομένως να υπολογίζουν δύναμη, κλίση, κίνηση, κλπ.), π.χ. κουμπιά και διακόπτες (μηχανικοί).
- Φως – ανιχνεύουν την παρουσία πηγών φωτός (και ενδεχομένως και άλλα στοιχεία όπως ένταση), π.χ. φωτοανιχνευτές, αλλά και κάμερες.
- Εγγύτητα (proximity) – την παρουσία ενός αντικειμένου (ενδεχομένως να υπολογίζουν την απόσταση, κίνηση, κλπ), π.χ. ηλεκτρικού πεδίου, κάμερες, αλλά ακόμα και μικρόφωνα, κ.α.
- Ήχο – ανιχνεύουν την παρουσία ήχων, - μικρόφωνα
- Κίνηση – μπορεί να ανιχνεύουν και ταχύτητα, συνήθως με παλμούς μικροκυμάτων (οι οποίοι αντανακλώνται στο αντικείμενο και από την επιστροφή τους υπολογίζεται η κίνηση).
- Προσανατολισμό, όπως συστήματα πλοήγησης, επίσης σημαντικό για χειριστήρια (Wii)

Οι παραπάνω αισθητήρες, ανάλογα με την εφαρμογή μπορεί να βρίσκονται (α) στο περιβάλλον και να ανιχνεύουν την παρουσία, κίνηση, κατεύθυνση, κλπ, του χρήστη, (β) πάνω στο σώμα του χρήστη (π.χ. σε κάποιο γάντι ή ρούχο) ή/και (γ) σε χειριστήριο που κρατάει ο χρήστης, π.χ. wiimote.

## A.4. Τεχνικές αλληλεπίδρασης: δυνατότητες και περιορισμοί

Αν και οι εφαρμογές κιναισθητικού ελέγχου χαρακτηρίζονται από μεγάλη ποικιλία, κάποια υποδείγματα αλληλεπίδρασης (interaction patterns) που έχουν αναγνωριστεί περιλαμβάνουν (Saffer, 2009):

*Η εγγύτητα (απ)ενεργοποιεί (Proximity (de)activates)*

Στοιχεία αλληλεπίδρασης	Χρησιμότητα	Τεχνολογία	Περιορισμοί χρήσης
<ul style="list-style-type: none"> <li>• Οθόνη</li> <li>• Στοιχεία του περιβάλλοντος (που ελέγχονται από Η/Υ ή αυτοματισμούς): π.χ. φως, ήχος, νερό, κλιματισμός</li> </ul>	<ul style="list-style-type: none"> <li>• Για να αποφύγετε φυσική επαφή (αγγίγματα) του χρήστη.</li> <li>• Για οικολογική χρήση πόρων (π.χ. νερό σε δημόσιες τουλέτες)</li> <li>• Όταν θέλετε να ξαφνιάσετε το χρήστη!</li> </ul>	<ul style="list-style-type: none"> <li>• Αισθητήρες κίνησης (π.χ. υπέρυθρου φωτός ή πλέγματος (tripwire))</li> <li>• Αισθητήρες ραδιοσυχνότητας (RFID) (ο χρήστης πρέπει να έχει πομπό (RFID tag))</li> <li>• Ψηφιακή κάμερα + σχετικό λογισμικό</li> <li>• Μικρόφωνο!</li> </ul>	<ul style="list-style-type: none"> <li>• Μπορεί να ενεργοποιηθεί κατά λάθος</li> <li>• Να σχεδιάζετε με δικλείδες ασφαλείας</li> </ul>

*Η κίνηση του σώματος ενεργοποιεί (Move body to activate)*

Στοιχεία αλληλεπίδρασης	Χρησιμότητα	Τεχνολογία	Περιορισμοί χρήσης
<ul style="list-style-type: none"> <li>• Συνδέστε τη δραστηριότητα του χρήστη με κάτι που θα ήθελε να συμβεί, π.χ.</li> <li>• Πλησιάζει πόρτα - άνοιγμα</li> <li>• Περιπατάει σε δρόμο - εμφάνιση χάρτη σε οθόνη</li> </ul>	<ul style="list-style-type: none"> <li>• Να αποφύγουμε άγγιγμα και άσκηση φυσικής δύναμης</li> <li>• Να προσφέρουμε χρήσιμη πληροφορία</li> </ul>	<ul style="list-style-type: none"> <li>• Ανίχνευση κίνησης και κατεύθυνσης:</li> <li>• Ψηφιακές κάμερες με κατάλληλο λογισμικό</li> <li>• Υπέρυθρο φως</li> <li>• Επιταχυνσίμετρο (accelerator)</li> </ul>	<ul style="list-style-type: none"> <li>• Μπορεί να ενεργοποιηθεί κατά λάθος.</li> <li>• Προσφέρετε ενδείξεις για να προετοιμάσετε το χρήστη.</li> </ul>

*Το δείξιμο επιλέγει/ενεργοποιεί (Point to select/activate)*

Στοιχεία αλληλεπίδρασης	Χρησιμότητα	Τεχνολογία	Περιορισμοί χρήσης
Ψηφιακά αντικείμενα (σε μεσαίες/μεγάλες οθόνες ή προβαλλόμενα)	Χειρισμός (επιλογή, μετακίνηση, κλπ) ψηφιακών αντικειμένων	(ανίχνευση κατεύθυνσης και υπολογισμός αντικειμένου) <ul style="list-style-type: none"> <li>• Κάμερα και κατάλληλο λογισμικό (σε κάποιες περιπτώσεις απαιτεί βαθμονόμηση)</li> <li>• Χειριστήριο</li> <li>• Γάντι που υπολογίζει / δείχνει θέση και προσανατολισμό</li> </ul>	Το χέρι δεν είναι σταθερό! Χρειάζονται <ul style="list-style-type: none"> <li>• μεγάλοι στόχοι</li> <li>• μικρή διάρκεια παραμονής / επιλογής στόχου</li> <li>• Αποφυγή λαθών, και επαναφορά.</li> </ul>

Παράδειγμα: η χρήση του χειριστηρίου Wii mote για επιλογή και πλοήγηση.

*Ο κυματισμός του χεριού ενεργοποιεί (Wave to activate)*

Στοιχεία αλληλεπίδρασης	Χρησιμότητα	Τεχνολογία	Περιορισμοί χρήσης
<ul style="list-style-type: none"> <li>• Στοιχεία περιβάλλοντος όπως φως, νερό, κ.α.</li> </ul>	<ul style="list-style-type: none"> <li>• Είναι φυσική κίνηση</li> <li>• Είναι δύσκολο να γίνει κατά λάθος</li> </ul>	<ul style="list-style-type: none"> <li>• Ψηφιακή κάμερα + κατάλληλο λογισμικό</li> <li>• Υπέρυθρος αισθητήρας</li> <li>• Αισθητήρας θερμοκρασίας</li> </ul>	Πρέπει να σταματάει αυτόματα

*Η περιστροφή αλλάζει την κατάσταση ενός αντικειμένου (Rotate to change state)*

Στοιχεία αλληλεπίδρασης	Χρησιμότητα	Τεχνολογία	Περιορισμοί χρήσης
<ul style="list-style-type: none"> <li>• Φορητή Συσκευή</li> </ul>	<ul style="list-style-type: none"> <li>• Για να αλλάξετε την όψη (οριζόντια-κατακόρυφη) δημιουργώντας ωφέλιμο χώρο.</li> <li>• Αντικείμενα με διπτές λειτουργίες που αλλάζουν χρήση ανάλογα με τον προσανατολισμό (π.χ. κινητό και φωτογραφική μηχανή)</li> </ul>	<ul style="list-style-type: none"> <li>• Γυροσκόπιο</li> <li>• Επιταχυνσίμετρο</li> </ul>	Μπορεί να ενεργοποιηθεί κατά λάθος Δυνατότητα άμεσης ρύθμισης (configuration)

Παράδειγμα: περιστρέφοντας το κινητό μας τηλέφωνο, περιστρέφεται και η οθόνη του.

*Το πάτημα με το πόδι ενεργοποιεί (Step to activate)*

Στοιχεία αλληλεπίδρασης	Χρησιμότητα	Τεχνολογία	Περιορισμοί χρήσης
<ul style="list-style-type: none"> <li>• Επιφάνειες ή συσκευές στις οποίες ο χρήστης μπορεί να πατήσει</li> </ul>	<ul style="list-style-type: none"> <li>• Όταν θέλετε να ενεργοποιηθεί μια δράση του συστήματος με το πάτημα του ποδιού του χρήστη</li> <li>• Όταν τα χέρια είναι κατειλημμένα</li> </ul>	Αισθητήρες πίεσης	Μπορεί να ελεργοποιηθεί κατά λάθος

Παράδειγμα: η πλατφόρμα παιχνιδιών και γυμναστικής Nintendo Wii Fit βασίζεται στην καταγραφή της κατανομής του βάρους του σώματος του χρήστη.

*Το ταρακούνημα ενεργοποιεί (Shake to activate)*

Στοιχεία αλληλεπίδρασης	Χρησιμότητα	Τεχνολογία	Περιορισμοί χρήσης
<ul style="list-style-type: none"> <li>Φορητές συσκευές</li> <li>Ήχος, φωτισμός, οθόνη</li> </ul>	<ul style="list-style-type: none"> <li>Όταν θέλετε να αλλάξετε την κατάσταση ενός αντικειμένου που κρατάτε με το χέρι σας.</li> <li>Είναι πιο εύκολο από το να δείτε το αντικείμενο και να πατήσετε κάποιο κουμπί...</li> </ul>	Επιταχυνσίμετρο	Μπορεί να ελεγχτοποιηθεί κατά λάθος

*Η κλίση προκαλεί κίνηση (Tilt to move)*

Στοιχεία αλληλεπίδρασης	Χρησιμότητα	Τεχνολογία	Περιορισμοί χρήσης
<ul style="list-style-type: none"> <li>Το σώμα ή η οπτική του χρήστη στο χώρο</li> <li>Συσκευές μεταφοράς και χειρισμού</li> </ul>	Έλεγχος κατεύθυνσης σε έναν 3D χώρο – πραγματικό ή εικονικό.	<ul style="list-style-type: none"> <li>Επιταχυνσίμετρα σε πολλούς άξονες,</li> <li>ή/και Γυροσκόπια,</li> <li>ή/και πολλαπλοί αισθητήρες πίεσης</li> </ul>	Κινήσεις όπως σκύψιμο, κλίση προς τα πλάγια ή πίσω, κ.α. μπορεί να μην είναι εύκολες ή ακόμα και επικίνδυνες για ηλικιωμένους ή πολύ μικρά παιδιά

## B. Συστήματα Διάδρασης με Χειρονομίες (εφαρμογές σε οικιακό περιβάλλον)

### B.1. Εισαγωγή

Η αναγνώριση χειρονομιών αποτελεί μια κατηγορία τεχνικών, που βασίζονται στην κίνηση - μετακίνηση των ανθρώπινων άνω άκρων (βραχίονας, πήχης, καρπός) ή απλά σε μια συγκεκριμένη διαμόρφωση του ανθρώπινου χεριού. Οι χειρονομίες είναι μέσα μη λεκτικής αλληλεπίδρασης μεταξύ των ανθρώπων (non - verbal communication). Κυμαίνονται από τις απλές ενέργειες κατά τις οποίες χρησιμοποιούμε το χέρι μας για να δείξουμε ή να μετακινήσουμε αντικείμενα, μέχρι πιο σύνθετες ενέργειες, που εκφράζουν συναισθήματα και μας επιτρέπουν να επικοινωνούμε με τους άλλους. Φέρουν δηλαδή ένα σημαντικό μέρος πληροφοριακού περιεχομένου.

Η αναγνώριση και ερμηνεία των χειρονομιών απαιτεί από τη μηχανή την ικανότητα να μετρήσει τις δυναμικές ή στατικές διαμορφώσεις του ανθρώπινου χεριού, του βραχίονα ή ακόμα και άλλων μερών του ανθρώπινου σώματος, τα οποία συμμετέχουν στην κίνηση. Οι πρώτες συσκευές οι οποίες έκαναν αυτή την εργασία, μετρούσαν απλά τη γωνία των αρθρώσεων του χεριού (γωνία μεταξύ πήχη-βραχίονα-καρπού) καθώς και τη χωρική θέση του χεριού. Χαρακτηριστικά παραδείγματα αυτής της κατηγορίας είναι οι λεγόμενες συσκευές βασισμένες στη χρήση γαντιών (glove - based devices) (Quam, 1990). Οι διεπαφές χειρονομιών βασισμένες στη χρήση γαντιών απαιτούν από το χρήστη να φοράει μια συσκευή εισόδου και να φέρει γενικά το φορτίο των καλωδίων που συνδέουν τη συσκευή με τον υπολογιστή. Αυτό όμως, εμποδίζει την ευκολία και τη φυσικότητα με τις οποίες ο χρήστης μπορεί να αλληλεπιδρά με το υπολογιστικό περιβάλλον. Ακόμα και αν η χρήση τέτοιων συγκεκριμένων συσκευών μπορεί να δικαιολογηθεί από την εξειδικευμένη περιοχή εφαρμογής, π.χ. προσομοίωση χειρουργικής επέμβασης σε περιβάλλον εικονικής πραγματικότητας, η περίπτωση τέτοιων συσκευών εισόδου, απέχει πολύ από τη φυσικότητα αλληλεπίδρασης που είναι ο επιδιωκόμενος σκοπός. Αυτό έχει οδηγήσει την έρευνα προς περισσότερο 'φυσικές' διεπαφές χειρονομιών. Τα προβλήματα που εμφανίζονται με τη χρήση γαντιών και άλλων παρόμοιων συσκευών, μπορούν να παρακαμφθούν με την εφαρμογή τεχνικών χρήσης video. Η προσέγγιση αυτή χρησιμοποιεί ένα σύνολο από βιντεοκάμερες, οι οποίες καταγράφουν το προς μελέτη περιβάλλον - αντικείμενο, έναν υπολογιστή και τεχνικές δανεισμένες από την αναγνώριση προτύπων για την ανάλυση και ερμηνεία των χειρονομιών.

Μέχρι σήμερα το μεγαλύτερο μέρος της εργασίας, το οποίο βασίζεται στην οπτική ερμηνεία των χειρονομιών στο πλαίσιο της HCI, έχει στραφεί στην αναγνώριση των στατικών χειρονομιών ή των στάσεων - διαμορφώσεων των χεριών. Για το σκοπό αυτό έχει χρησιμοποιηθεί μια μεγάλη ποικιλία μοντέλων, τα περισσότερα από τα οποία είναι δανεισμένα από το πεδίο των τεχνικών αναγνώρισης αντικειμένων. Οι εικόνες των χεριών, γεωμετρικές ροπές ολόκληρων εικόνων ή συγκεκριμένων περιοχών ενδιαφέροντος (region of interest) (μέσες τιμές και αποκλίσεις κ.α), περιγράμματα, σκιαγραφήσεις και τρισδιάστατα

σκελετικά μοντέλα χεριών (3D skeletal models) είναι μερικά παραδείγματα. Βέβαια, εκτός από τις στατικές χειρονομίες υπάρχουν και οι δυναμικές (dynamic hand gestures), οι οποίες αποτελούν μια χωροχρονική διαμόρφωση του χεριού (space - time gestures, spatio - temporal gestures). Αποτελούν ένα προκλητικά ενδιαφέρον πεδίο, με τελικό σκοπό την ενσωμάτωση των δυναμικών χαρακτηριστικών των χειρονομιών στην HCI.

Η βασική λογική είναι ότι, οι δυναμικές χειρονομίες είναι ενέργειες και η κίνηση των χεριών μεταβιβάζει πληροφορία, όπως ακριβώς η στάση τους. Διάφορες προσεγγίσεις έχουν προταθεί για την ανάλυση της δυναμικής χειρονομίας, οι οποίες κυμαίνονται από την ολική ανάλυση της κίνησης των χεριών ως την ανεξάρτητη ανάλυση των άκρων των δακτύλων. Η σημερινή τάση στο χώρο της αναγνώρισης χειρονομιών αποτελείται από τη συγχώνευση μεθόδων των στατικών και δυναμικών χειρονομιών, με τελικό στόχο την κατασκευή στιβαρών μοντέλων αναγνώρισης, ανεξάρτητα από την ύπαρξη ή όχι κίνησης σε οποιοδήποτε περιβάλλον.

## **B.2. Ιστορική Αναδρομή**

Πολλές και ποικίλες είναι οι προσεγγίσεις των ερευνητών στην αναγνώριση χειρονομιών. Οι Yamato, Ohya και Ishii (1992) ήταν από τους πρώτους που χρησιμοποίησαν τα διακριτά μοντέλα hidden Markov για την αναγνώριση έξι διαφορετικών χτυπημάτων tennis. Πιο συγκεκριμένα, χώρισαν κάθε εικόνα σε πλέγματα και πήραν ως χαρακτηριστικό γνώρισμα κάθε εικόνας τον αριθμό των εικονοστοιχείων σε κάθε πλέγμα. Παρά τη φτώχη αυτή περιγραφή της κίνησης, εμφάνισαν ιδιαίτερα καλά αποτελέσματα.

Οι Starner και Pentland (1995) χρησιμοποίησαν συνεχή μοντέλα Markov για την αναγνώριση περιορισμένου αριθμού κινήσεων νοηματικής, με ή χωρίς γραμματική. Μοντελοποίησαν με κανονική τ.μ. τα σύμβολα εξόδου. Το χαρακτηριστικό διάνυσμα ήταν το κέντρο της παλάμης των δύο χεριών (x,y συντεταγμένες) καθώς και η μεταξύ τους απόσταση και γωνία του άξονα σύνδεσης. Για την κατάτμηση της εικόνας χρησιμοποίησαν χρωματιστά γάντια, τα οποία ανιχνεύονταν από ένα σύστημα καμερών και φιλτράρονταν από την υπόλοιπη εικόνα.

Οι Nam και Wohn (1995) προσέγγισαν με ιδιαίτερο τρόπο το πρόβλημα, θεώροντας ως βασικά συστατικά της χειρονομίας την θέση του χεριού, τον προσανατολισμό της παλάμης και την κίνηση ολόκληρων των άνω άκρων. Όρισαν την πρωταρχική χειρονομία ως την κίνηση των άκρων που παραμένει σταθερή στα δύο πρώτα χαρακτηριστικά. Από τις συνεχόμενες κινήσεις των μουσικών για τη διεύθυνση ορχήστρας, χρησιμοποιώντας τεχνικές κλίσης (gradient techniques), 'έμαθαν' αυτές τις χειρονομίες. Με ένα δίκτυο μοντέλων hidden Markov πέτυχαν υψηλά ποσοστά αναγνώρισης. Μειονέκτημα αποτελεί το γεγονός ότι περιορίζονται σε μικρό αριθμό χειρονομιών, που εμφανίζουν έντονη περιοδικότητα.

Η μοναδική ίσως προσέγγιση με χρήση ασαφούς λογικής παρουσιάστηκε από τους Wilson και Bobick, (1997 & 1995), Μετά από πολλές εκτελέσεις όρισαν την πρωτεύουσα τροχία (2D) μιας χειρονομίας, στην οποίας διατήρησαν τα χρονικά χαρακτηριστικά και τη σειρά εξέλιξης της. Όρισαν την χειρονομία ως ένα σύνολο καταστάσεων και μια συγκεκριμένη χειρονομία ως αλληλουχία τέτοιων καταστάσεων. Η κατάταξη επιτυγχάνεται με χρήση ασαφών τελεστών.

Άλλες υλοποιήσεις, που στηρίζονται σε κλασικές τεχνικές dynamic time wrapping εφάρμοσαν οι Darell και Pentland (1992 & 1993) οι οποίοι συγκρίνουν τις εικόνες εισόδου με προεπιλεγμένες εικόνες που αποτελούν τη μοντελοποίηση μιας χειρονομίας. Χρησιμοποιούν ως μέτρο κατάταξης τη συσχέτιση (correlation) των εικόνων. Η μέθοδος αυτή εμφανίζει ευαισθησία στην εμφάνιση του χεριού και χρειάζεται μεγάλο πλήθος εκτελέσεων για τον ορισμό της βέλτιστης ακολουθίας - προτύπου.

Μια ενδιαφέρουσα επίλυση στο πρόβλημα της απόρριψης λανθασμένων χειρονομιών έδωσαν οι Hyeon και Jin (1999), οι οποίοι εκμεταλλεύτηκαν μια ιδιότητα των μοντέλων Markov: κάθε κατάσταση αποτελεί υποπρότυπο μιας χειρονομίας. Κατασκεύασαν ένα μοντέλο συνδέοντας όλες τις δυνατές καταστάσεις που μπορούσε να βρεθεί αυτό, το οποίο αποτέλεσε ένα είδος καταφλίου για την απόρριψη ή αποδοχή μιας χειρονομίας. Ως χαρακτηριστικά γνωρίσματα επέλεξαν τη γωνία του κέντρου βάρους της παλάμης μεταξύ δύο διαδοχικών εικόνων.

Μια απλή και αποτελεσματική προσέγγιση αποτελεί η εφαρμογή μάσκας κίνησης και χρώματος για τον εντοπισμό των χεριών από τους Karpouzis, Raouzaίου, Drosopoulos, Ioannou, Balomenos, Tsapatsoulis και Koliias (2004). Οι παραπάνω χρησιμοποιούν την πληροφορία της χειρονομίας, την οποία συνδυάζουν με αναγνώριση συναιθήματος και φωνής για πολυτροπική επικοινωνία ανθρώπου - υπολογιστή.

Τέλος, πολλές παραλλαγές των μοντέλων hidden Markov έχουν εμφανιστεί. Κλασική παραλλαγή αποτελούν τα παράλληλα μοντέλα Markov από τους Vogler και Metaxas (1999). Μοντελοποιείται κάθε χέρι ξεχωριστά με ένα μοντέλο Markov και η άθροιση των παραπάνω πληροφοριών χρησιμοποιείται για αναγνώριση νοηματικής γλώσσας.

	<b>Λεξιλόγιο Χειρονομιών</b>	<b>Ελεγχόμενες Λειτουργίες</b>	<b>Υλικό</b>	<b>Χαρακτηριστικά</b>	<b>Αναγνώριση Προτύπων</b>
Yamato, Ohya και Ishii (1992)	-	Human actions and hand tracking	-	Low-level image features, training	Symbols
Starner και Pentland (1995)	sentence level American Sign Language (ASL) without explicitly modeling the fingers	Hand tracking	solidly-colored cloth gloves, camera, Silicon Graphics Indigo 2 with Galileo video board, Silicon Graphics 200MHz Indy workstation	hands tracked by their color, real time, training	hand shape, orientation, and trajectory
Nam και Wohn (1995)	Specific gestures	Human actions and hand tracking	Mouse input device, magnetic sensor, video camera,	Time-augmented configuration space, training	Hand states and gestures, motion trajectory
Wilson και Bobick, (1997)	Specific gestures, states of gestures	time-invariant but order-preserving gesture tracking	mouse input device, magnetic spatial position and orientation sensor, video camera,	computes prototype gestures, arc length parameterization of data points, training	gestures from an unsegmented, continuous stream of sensor data, trajectories
Darell και Pentland (1992)	View models, gesture patterns	Human hand gestures, 3D object models	General purpose computer, cognex vision processor, sun workstation	Learning, tracking, training, real-time	Model object and behavior, use of view models and stereotypical patterns (e.g. gestures)
Karpouzis et. al. (2004)	various gesture classes (e.g. clapping, etc.)	Emotions (facial expressions combined with hand gestures)	web-camera,	Gesture tracking and recognition, facial expression analysis, hand detection and	facial expressions and hand gestures, gesture sequences,

				tracking,	emotions (joy, sadness, etc.)
--	--	--	--	-----------	-------------------------------

Στις επόμενες παραγράφους, θα γίνει αναφορά στις μεθόδους σχεδίασης και αξιολόγησης τέτοιων συστημάτων, καθώς και λεπτομερείς ανάλυση στο λογισμικό και το υλικό που χρησιμοποιείτε στις μέρες μας για την ανάπτυξη των πλέον σύγχρονων τεχνικών αλληλεπίδρασης ανθρώπου-υπολογιστή.

### B.3. Μέθοδοι σχεδίασης και αξιολόγησης

Στη διαδικασία ανάπτυξης μιας κιναισθητικής διεπαφής σε ένα σύστημα αλληλεπίδρασης ανθρώπου-υπολογιστή, συνήθως χρησιμοποιούνται κάποιες γενικές μέθοδοι αλληλεπίδρασης και διαδραστικής σχεδίασης, οι οποίες μπορούν να διακριθούν (Κουτσαμπάσης, 2011) στις: (α) Αναζήτηση και Έρευνα (research & inquiry), σχεδίαση (design) και αξιολόγηση (evaluation). Οι περισσότερες μέθοδοι στις μέρες μας έχουν αναπτυχθεί με βάση αυτά τα κριτήρια και βήματα (βλ. παρακάτω), καθώς ο στόχος των περισσότερων συστημάτων αλληλεπίδρασης ανθρώπου-υπολογιστή είναι ανθρωποκεντρικός και το σύστημα θα πρέπει να ανταπεξέρχεται ανάλογα σε διάφορες “μη τυποποιημένες” χειρονομίες.

Γι’ αυτό το λόγο, η έμφαση κυρίως δίνεται στις χειρονομίες και τις κινήσεις του σώματος, και το πιο σημαντικό θέμα είναι η ανακάλυψη και ο έλεγχος των κατάλληλων χειρονομιών σε σχέση με τους στόχους του χρήστη και τις δράσεις (ή ακόμα και την ανάδραση) του ίδιου του συστήματος. Τελικός στόχος είναι η σωστή σχεδίαση των τεχνικών αλληλεπίδρασης (interaction techniques) που θα χρησιμοποιηθούν στο σύστημα, και προς αυτήν την κατεύθυνση οι Nielsen et al (2008) προτείνουν την παρακάτω διαδικασία εύρεσης και ελέγχου χειρονομιών:

Βήμα 1. Εντοπισμός των λειτουργιών του συστήματος (λειτουργικοί στόχοι). Σε αρκετές περιπτώσεις το σύστημα μπορεί να υπάρχει σε περιβάλλον γραφείου (desktop), π.χ. να πρόκειται για μια σχεδιαστική εφαρμογή, έναν εικονικό κόσμο, κλπ. Θα πρέπει να επιλεγθούν οι σημαντικές χειρονομίες για το συγκεκριμένο πλαίσιο χρήσης που ενδιαφέρει το χρήστη (πιθανότατα ένα υποσύνολο μιας υπάρχουσας εφαρμογής, ενδεχομένως με την προσθήκη κάποιων νέων λειτουργιών).

Βήμα 2. Εντοπισμός των λογικών χειρονομιών (αναζήτηση χειρονομιών). Το ‘λογικό’ ορίζεται σε σχέση με αυτό που περιμένουν οι χρήστες ότι θα ισχύει. Άρα, θα πρέπει να γίνουν δοκιμές με κάποιους χρήστες πριν γίνει η τελική επιλογή των χαρακτηριστικών χειρονομιών που θα αντιπροσωπεύουν το σύστημα. Για παράδειγμα, μπορεί να ζητηθεί από τους χρήστες να:

- (α) ταξινομήσουν πιθανές χειρονομίες σε κάρτες,
- (β) προτείνουν αυθόρμητα μια χειρονομία,
- (γ) βιντεοσκοπηθεί μια συνέντευξη μαζί τους, ώστε να καταγραφούν καλύτερα οι χειρονομίες (αυθόρμητες) που θα κάνουν.

Επίσης, εδώ είναι σημαντικό να ληφθεί υπόψη ποιες χειρονομίες υποστηρίζονται από την τεχνολογική πλατφόρμα που θα υποστηρίξει την εφαρμογή. Για παράδειγμα μπορεί να μην υπάρχει η δυνατότητα να δημιουργηθούν νέες χειρονομίες στο υπάρχον σύστημα, λόγω περιορισμένων δυνατοτήτων στο λογισμικό.

Βήμα 3. Ανάλυση των δεδομένων (προκαταρκτική σχεδίαση: χειρονομίες και σενάρια χρήσης). Αυτό το βήμα περιλαμβάνει τη μέτρηση των συχνοτήτων των χειρονομιών που προτάθηκαν, τη διάρκεια τους, της εκτίμησης της καταλληλότητας τους σε σχέση με το πλαίσιο χρήσης, θέματα κουλτούρας, κ.α. Εδώ ουσιαστικά χρησιμοποιείται κάποιου είδους σχεδιαστική σκέψη για να παρθούν κάποιες αποφάσεις όσο το δυνατόν πιο τεκμηριωμένα σε σχέση με τα δεδομένα. Επίσης, σε αυτό το στάδιο θα πρέπει να φτιαχτούν κάποια σενάρια χρήσης, γενικευμένα στην αρχή και αναλυτικότερα στη συνέχεια, να ληφθεί υπόψη η αλληλουχία των χειρονομιών, κ.α.

Βήμα 4. Έλεγχος των χειρονομιών (αναλυτική σχεδίαση: ευχρηστία και τεχνικές αλληλεπίδρασης). Σε αυτό το βήμα υπάρχουν κάποιες δοκιμές που πρέπει να γίνουν για την αναλυτική σχεδίαση του συστήματος. Αυτό που προτείνεται από τον Kortum (2008) είναι οι δοκιμές να γίνουν με απλά πρωτότυπα σε χαρτί (paperprototyping) ή σε προγράμματα κατασκευής πρωτοτύπων. Τα στάδια είναι τα εξής:

(α) Επιλογή της χειρονομίας: Εδώ θα πρέπει ο «εκπαιδευτής» να δείξει στους χρήστες μια χειρονομία και να τους ζητήσει να μαντέψουν τι θα συμβεί σε συγκεκριμένα σημεία της αλληλεπίδρασης (κατά τη διάρκεια διάφορων σεναρίων χρήσης).

(β) Σκέψη της χειρονομίας: Εδώ θα πρέπει να ζητηθεί από τους χρήστες να ξεκινήσουν να αλληλεπιδρούν με το σύστημα χρησιμοποιώντας κάποιες χειρονομίες. Ο «εκπαιδευτής» θα πρέπει να δώσει έμφαση στο γεγονός ότι οι χρήστες μπορεί να μην θυμούνται ή να μην εντοπίσουν εύκολα τις σωστές χειρονομίες (όπου σωστές είναι αυτές που έχουν ήδη προκαθοριστεί), αλλά θα πρέπει παρ'όλα αυτά να σημειώσει και τις 'λάθος' χειρονομίες, γιατί στο τέλος μπορεί τελικά να χρησιμοποιηθούν αυτές για πρακτικούς λόγους. Επιπλέον, θα πρέπει να σημειωθούν οι χρόνοι ενθύμησης των χειρονομιών και να εκτιμηθεί η ευκολία με την οποία μπορούν τελικά οι χρήστες να θυμηθούν τι πρέπει να κάνουν. Τέλος, θα πρέπει να γίνει και εντοπισμός των ενδεχόμενων σημείων στα οποία πρέπει να ενεργοποιηθεί βοήθεια, κ.α.

(γ) Ένταση και φυσική δύναμη: Σε αυτό το στάδιο ζητείται από τους χρήστες να εκτιμήσουν πως οι ίδιοι αντιλαμβάνονται την ευκολία της κάθε χειρονομίας σε σχέση με την προσπάθεια που πρέπει να καταβάλουν. Εδώ έχει επίσης σημασία να δοθεί βαρύτητα στο τι συμβαίνει σε επαναλαμβανόμενες λειτουργίες και σε 'μακρές' αλληλεπιδράσεις με το σύστημα. Επίσης, θα πρέπει να ομαδοποιηθούν τα αποτελέσματα σε σχέση με τις ομάδες που συμμετείχαν στις δοκιμές, για παράδειγμα σε αυτούς που μπορεί να κουράζονται πιο εύκολα όπως οι ηλικιωμένοι, κτλ.

Βήμα 5. Βελτίωση των χειρονομιών με έλεγχο επί του τελικού συστήματος (τεχνικός έλεγχος και έλεγχος ευχρηστίας με χρήστες). Στο τελευταίο βήμα, και αφού υλοποιηθεί το σύστημα, είναι απαραίτητο να γίνουν και άλλοι έλεγχοι με χρήστες. Οι έλεγχοι σε αυτό το στάδιο θα αφορούν συνολικά το σύστημα, αλλά ειδικά για τις χειρονομίες τα σημαντικά θέματα είναι: (α) η διάρκεια, (β) ο χρονισμός και (γ) ο προσανατολισμός της κάθε χειρονομίας. Εφόσον έχουν γίνει δοκιμές πιο πριν και έχει ακολουθηθεί μια ανθρωποκεντρική σχεδιαστική διαδικασία όπως αυτή που περιγράφεται στα παραπάνω βήματα, οι τεχνικοί έλεγχοι αναμένεται να έχουν ως βασικό αποτέλεσμα βελτιστοποιήσεις στη χρήση του συστήματος.

#### **B.4. Συστήματα Αλληλεπίδρασης με Ανθρώπους και Διεπαφές που Βασίζονται στις Χειρονομίες**

Αφού περιγράφηκε το γενικό πλαίσιο που διέπει τα συστήματα αλληλεπίδρασης ανθρώπου-υπολογιστή, έγινε μια σύντομη ιστορική αναδρομή και περιγράφηκε ο τρόπος σχεδιασμού που ακολουθείται σε αυτά τα συστήματα, στις επόμενες παραγράφους γίνεται μια πιο λεπτομερής αναφορά σε σύγχρονες μεθόδους και συστήματα αλληλεπίδρασης ανθρώπου-υπολογιστή, τα οποία πλέον τείνουν όλο και περισσότερο να πλησιάζουν να ενταχθούν στην καθημερινότητά μας. Επίσης γίνεται αναφορά και στις διεπαφές των συστημάτων με την αντίστοιχη τεχνολογία που χρησιμοποιούν, ενώ σε επόμενη ενότητα γίνεται η σύγκριση/σχολιασμός τους όσον αφορά την αποτελεσματικότητα και την ακρίβεια με την οποία αναγνωρίζονται οι χειρονομίες των χρηστών.

##### *Microsoft Kinect Sensor*

Μια από τις πλέον πρόσφατες μηχανές ανίχνευσης κίνησης και αλληλεπίδρασης με τις ανθρώπινες χειρονομίες είναι το *Microsoft Kinect Sensor*. Το Kinect είναι μια συσκευή ανίχνευσης κίνησης της Microsoft για το Xbox 360 και τα Windows PCs. Είναι στην ουσία μια περιφερειακή συσκευή εισόδου που διαθέτει φακούς-κάμερες και ηχητικούς αισθητήρες ώστε να επιτρέπει στους χρήστες να ελέγχουν και να αλληλεπιδρούν εξ αποστάσεως χωρίς να είναι απαραίτητη η επαφή τους με κάποιο άλλο χειριστήριο,



ενσύρματο ή ασύρματο. Η αλληλεπίδραση πραγματοποιείται χρησιμοποιώντας χειρονομίες και προφορικές εντολές μέσω του φυσικού περιβάλλοντος του χρήστη.



Εικόνα 3. Microsoft Kinect.

Το Kinect (Xbox 360 Handbücher und Spezifikationen, 2010) βασίζεται σε τεχνολογίες λογισμικού που αναπτύσσονται από τη Rare, θυγατρική εταιρεία της Microsoft Game Studios, και την τεχνολογία της ισραηλινής PrimeSens πάνω στις κάμερες και τις συσκευές λήψης βίντεο. Η PrimeSens ανέπτυξε ένα σύστημα που μπορεί να αντιληφθεί χειρονομίες και καθιστά δυνατό τον έλεγχο της συσκευής χρησιμοποιώντας μια υπέρυθη κάμερα, δύο φακούς και ένα ειδικό μικροτσίπ για την παρακολούθηση της κίνησης των αντικειμένων και των ατόμων σε τρεις διαστάσεις. Αυτό το σύστημα τρισδιάστατης σάρωσης, που ονομάζεται Light Coding, χρησιμοποιεί μια παραλλαγή τρισδιάστατης εικόνας η οποία μπορεί να αναπαρασταθεί σε τρισδιάστατο χώρο. Το σύστημα διαθέτει έναν RGB φακό, έναν αισθητήρα βάθους που είναι συνδυασμός δύο φακών/καμερών και ένα πολλαπλό μικρόφωνο, τα οποία λειτουργούν με το συγκεκριμένο λογισμικό που περιλαμβάνει το Microsoft Kinect SDK και παρέχουν πλήρη τρισδιάστατη καταγραφή της κίνησης, αναγνώριση προσώπου και δυνατότητες αναγνώρισης φωνής.

Ο αισθητήρας βάθους αποτελείται από έναν υπέρυθρο προβολέα λέιζερ σε συνδυασμό με ένα αισθητήρα CMOS, ο οποίος καταγράφει δεδομένα τρισδιάστατου βίντεο κάτω από οποιεσδήποτε συνθήκες φωτισμού. Σύμφωνα με τις πληροφορίες που παρέχονται, το Kinect είναι σε θέση να εντοπίζει ταυτόχρονα έως έξι άτομα, μεταξύ των οποίων δύο ενεργούς παίκτες των οποίων αναλύει την κίνηση με δυνατότητα αντίληψης 20 αρθρώσεων ανά παίκτη.



Εικόνα 4. Τρόπος λειτουργίας Kinect.

Ο αισθητήρας Kinect εξάγει βίντεο με ρυθμό πλαισίων 30 Hz. Η ροή βίντεο χρησιμοποιεί το σύστημα χρωμάτων RGB 8-bit και ανάλυση 640 × 480 pixel με χρωματικό φίλτρο Bayer, ενώ η μονόχρωμη αισθητήρια ροή βίντεο χρησιμοποιεί ανάλυση 640 × 480 pixel με βάθος 11-bit, το οποίο παρέχει 2.048 επίπεδα ευαισθησίας. Ο αισθητήρας Kinect έχει ένα πρακτικό όριο που κυμαίνεται από 1,2 έως 3,5 μέτρα απόσταση. Η περιοχή που καλύπτει το οπτικό πεδίο του Kinect είναι περίπου 6 τ.μ., αν και ο αισθητήρας μπορεί να διατηρήσει την εστίαση παρακολούθησης σε ένα διευρυμένο φάσμα έως και 6 μέτρων. Το πολλαπλό μικρόφωνο διαθέτει τέσσερις μικροφωνικές συσκευές και καθεμία λειτουργεί με κανάλι των 16-bit ήχου με ρυθμό δειγματοληψίας 16 kHz.

Το λογισμικό στο οποίο βασίζεται το Kinect είναι το Kinect Windows software development kit (SDK) και το XNA Framework. Το πρώτο δίνει τη δυνατότητα στους προγραμματιστές να χρησιμοποιήσουν τη C++, C#, ή τη Visual Basic για να δημιουργήσουν εφαρμογές που υποστηρίζουν κινήσεις χεριών και αναγνώριση φωνής χρησιμοποιώντας το Kinect με τους αισθητήρες του και έναν υπολογιστή ή μια ενσωματωμένη συσκευή.

Από την άλλη, το Microsoft XNA είναι ένα σύνολο εργαλείων διαχείρισης περιβάλλοντος πραγματικού χρόνου κατασκευασμένο από τη Microsoft που διευκολύνει την ανάπτυξη και τη διαχείριση ηλεκτρονικών παιχνιδιών. Το XNA σήμερα περιλαμβάνει ολόκληρα τμήματα ανάπτυξης της Microsoft Games, όπως το Xbox Development Kit και το XNA Game Studio. Το XNA Framework είναι βασισμένο στις εγγενείς εφαρμογές .NET Compact Framework 2.0 για ανάπτυξη παιχνιδιών για Xbox 360 και .NET Framework 2.0 για Windows. Περιλαμβάνει ένα εκτεταμένο σύνολο βιβλιοθηκών ειδικών για την ανάπτυξη παιχνιδιών, για να επιτύχει τη μεγιστοποίηση χρήσης κοινού κώδικα σε όλες τις πλατφόρμες. Είναι φιλικό και απλό προς το δημιουργό, πράγμα το οποίο το καθιστά αποδοτικό για να δημιουργεί στα παιχνίδια ένα καλοστημένο περιβάλλον λειτουργίας. Αυτό είναι και ένα από τα βασικά πλεονεκτήματά του έναντι των υπόλοιπων αντίστοιχων συσκευών. Παιχνίδια που εκτελούνται στο framework του XNA είναι τεχνικά δυνατό να γραφτούν σε οποιαδήποτε .NET γλώσσα, αλλά μόνο το IDE της C# του XNA Game Studio Express και όλες οι εκδόσεις του Visual Studio 2008 και 2010 το υποστηρίζουν επίσημα. Το XNA έχει ενσωματωμένο ένα μεγάλο αριθμό εργαλείων, όπως το XACT που είναι συμβατό με κάθε είδους πλατφόρμα, ώστε να ενισχύσει τη δημιουργία περιεχομένου στην ανάπτυξη παιχνιδιών. Το XNA υποστηρίζει τόσο τη δημιουργία δισδιάστατου όσο και τρισδιάστατου περιβάλλοντος, που μπορούν να αποτελούν το χώρο δράσης κάποιου παιχνιδιού.

### *Samsung Smart Interaction*

Μια άλλη πρόσφατη τεχνολογία που αναπτύχθηκε από τη Samsung που επιτρέπει την έξυπνη Αλληλεπίδρασης Ανθρώπου – Υπολογιστή Ανθρώπου-Μηχανής είναι το SSI (*Samsung Smart Interaction*). (Samsung, 2012). Αυτό αποτελεί μια καινοτομία που παρουσιάστηκε το 2012 και που σκοπό έχει να παράσχει ένα νέο σύστημα ελέγχου φωνής και χειρονομιών για τηλεοράσεις. Χρησιμοποιεί μια ενσωματωμένη κάμερα και μικρόφωνο για να μπορεί να ελέγχετε η λειτουργία μόνο με την ομιλία και με χειρονομίες. Αυτή η τεχνολογία είναι διαθέσιμη στη σειρά τηλεοράσεων πλάσματος PNE8000, καθώς και στις νέες τηλεοράσεις που υποστηρίζουν την τεχνολογία LED. Προς το παρόν, κανένας άλλος κατασκευαστής δε προσφέρει τέτοιες λειτουργίες στις τηλεοράσεις του.



Εικόνα 5. Το σύστημα SSI (Samsung Smart Interaction).

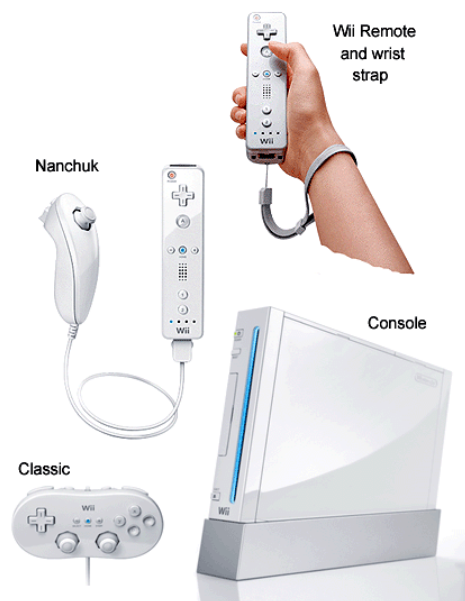
Το Smart Interaction που προσφέρεται μέσα από αυτό το σύστημα, πέρα από τα προηγούμενα χαρακτηριστικά, ενσωματώνει και ένα Bluetooth-to-IR blaster και δύο διαφορετικά χειριστήρια: ένα χειριστήριο με πολλά κουμπιά και μια επιφάνεια αλληλεπίδρασης, τα οποία μπορούν να χρησιμοποιηθούν και για το χειρισμό της τηλεόρασης. Επίσης, η αναγνώριση φωνής και το πρόγραμμα ελέγχου χειρονομιών είναι αρκετά ευέλικτο, αν και προς το παρόν παρέχει περιορισμένη δυνατότητα χειρισμού. Πιο συγκεκριμένα, ο έλεγχος φωνής δουλεύει με κάποιες φωνητικές εντολές ή με επιλογές από το μενού, αλλά υπάρχει μεγάλη καθυστέρηση στην ανταπόκριση του συστήματος, με αποτέλεσμα οι χρήστες, μετά τις δοκιμές που έγιναν, να προτιμούν το μενού από τις φωνητικές εντολές. Υπάρχουν επίσης και κάποια άλλα “νηπιακά” προβλήματα με το σύστημα ελέγχου φωνής, τα οποία αποτρέπουν το χρήστη από το να κάνει ευρεία χρήση του.

Όσον αφορά το σύστημα ελέγχου χειρονομιών, δουλεύει αρκετά ικονοποιητικά όσον αφορά την απόδοση, αλλά παρατηρούνται κάποια προβλήματα στην ανοχή της εικόνας όταν υπάρχουν σκιάσεις. Παρόλο που η φωτεινότητα της εικόνας είναι πάρα πολύ καλή, κάποιες φορές οι χρήστες κατά τη διάρκεια των δοκιμών των φωνητικών εντολών (Katzmaier, 2012) δεν μπορούσαν καν να ενεργοποιήσουν το σύστημα ελέγχου φωνής. Οι δοκιμές που έγιναν αφορούσαν την ενεργοποίηση των μενού, την κίνηση του χεριού για την ενεργοποίηση μιας επιλογή από το μενού, διαδοχικά κλικ για τη δοκιμή κάποιων λειτουργιών, περιήγηση στο διαδίκτυο, κ.α. Τα αποτελέσματα έδειξαν ότι παρόλο που σε κάποιες περιπτώσεις η ανταπόκριση ήταν πάρα πολύ καλή, σαν γενική εντύπωση δεν άφησε τις καλύτερες εντυπώσεις στους χρήστες. Το ίδιο συνέβει και με τον φωνητικό έλεγχο, ο οποίος ήταν αρκετά πιο αργός στην ανταπόκριση σε σχέση με το σύστημα ελέγχου χειρονομιών.

Όσον αφορά το υλικό, ο ασύρματος blaster που χρησιμοποιείται από τη συσκευή δέχεται εντολές από το Bluetooth μέσω της τηλεόρασης και τις μεταφράζει σε υπέρυθρα κύματα εντολών για τη συσκευή ή για το Blue-ray player. Συνδυάζεται με τα προηγούμενα συστήματα ελέγχου χειρονομιών και φωνής, αλλά παρέχει περιορισμένες δυνατότητες που δε σε επιτρέπουν να χρησιμοποιήσεις πλήρως τις απομακρυσμένες συσκευές. Μετά από αρκετά πειράματα, το συμπέρασμα που προκύπτει είναι ότι παρόλο που σε σύστημα παρέχει τεχνολογικές καινοτομίες που δεν υπάρχουν σε άλλα συστήματα, το τελικό αποτέλεσμα δεν είναι ακόμα ικανοποιητικό και οι περισσότεροι χρήστες δε θα το προτιμούσαν για προσωπική χρήση.

#### *Nintendo Wii*

Ένα άλλο παρόμοιο σύστημα που έχει παρουσιαστεί είναι το Wii της Nintendo, το οποίο υπόσχεται ότι στο μέλλον οι τηλεοράσεις θα ελέγχονται πλήρως με χειρονομίες, με βάση τις καινοτομίες που εισαγάγει. Σε σύγκριση με τους ανταγωνιστές της, το Playstation 3 και το Xbox 360, υστερεί από άποψη γραφικών και άλλων τεχνικών χαρακτηριστικών, όμως το δυνατό σημείο της κονσόλας είναι ο χειρισμός, ο οποίος γίνεται με έναν αρκετά επαναστατικό τρόπο από αυτούς έχει συνηθίσει ο χρήστης μέχρι στιγμής.



Εικόνα 6. Το Wii της Nintendo.

Πιο συγκεκριμένα, ενώ μέχρι στιγμής έχουμε συνηθίσει στα κοινά χειριστήρια παιχνιδιών, το Wii κάνει τη διαφορά. Ο χειρισμός γίνεται μέσω δύο τμημάτων: με το Remote, το οποίο μοιάζει με τηλεχειριστήριο και το Nunchuk, το οποίο είναι ένα κλασικό joystick σε μικρότερο μέγεθος. Αυτά τα δύο τμήματα συνδέονται με ένα καλώδιο μεταξύ τους. Το Wii Remote ενσωματώνει αισθητήρες κίνησης, οι οποίοι λαμβάνουν τις κινήσεις που κάνει ο παίκτης με τα χέρια του και μεταφέρονται στο παιχνίδι. Για παράδειγμα σε ένα παιχνίδι τένις μπορείς να χρησιμοποιήσεις το Remote σαν ρακέτα, σε ένα παιχνίδι ξιφομαχιών σαν σπαθί κτλ. Μερικά παιχνίδια όπως το Fling Smash, Wii Sports Resort, Red Steel 2 προϋποθέτουν μεγαλύτερη ακρίβεια και για αυτό χρειάζονται την χρήση ενός περιφερειακού που ονομάζεται Wii Motion Plus. Το Wii Motion Plus μπαίνει πίσω από το Wii Remote και προσφέρει 1:1 κίνηση. Το καινούριο μοντέλο του Wii Remote είναι το Wii Remote Plus και ενώ έχει ίδιο μέγεθος με το απλό Wii Remote έχει είδη ενσωματωμένο μέσα του το Wii Motion Plus.

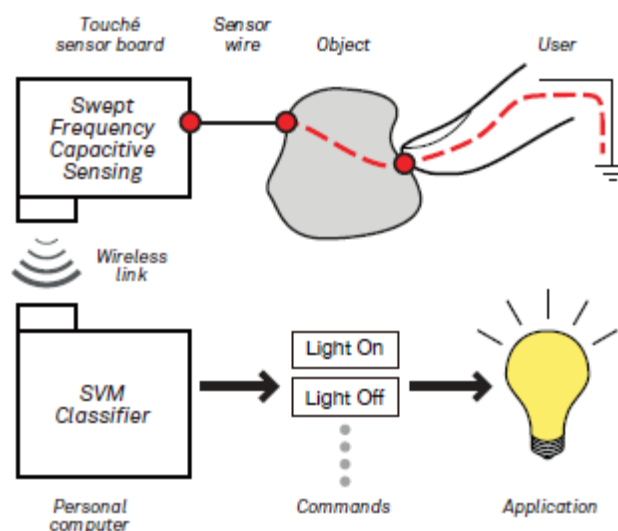
### *Disney Touche*

Μια άλλη καινοτόμα τεχνολογία που παρουσιάστηκε το 2012 είναι το Touché, που αναπτύχθηκε από τη Disney Research Group σε συνεργασία με το Πανεπιστήμιο του Τόκιο και το Πανεπιστήμιο Carnegie Mellon στην Αμερική. Αυτή η τεχνολογία προτείνει μια πρωτότυπη τεχνική εναλλαγής συχνότητας χωρητικής αντίληψης που μπορεί να ανιχνεύσει περισσότερες από μια μορφές αφής, αλλά ταυτόχρονα μπορεί να αναγνωρίζει πολύπλοκες συνθέσεις των χεριών και του σώματος κατά τη διάρκεια της αλληλεπίδρασης με την αφή. Αυτό δίνει τη δυνατότητα να ενισχυθεί σημαντικά η Αλληλεπίδραση Ανθρώπου – Υπολογιστή αφής σε ένα ευρύ φάσμα εφαρμογών, από την ενίσχυση στις συμβατικές οθόνες αφής για το σχεδιασμό σεναρίων αλληλεπίδρασης, με το μοναδικό περιβάλλον και τη ξεχωριστή χρήση υλικών. Για παράδειγμα, σε έρευνες τους χρησιμοποιήσαν πολύπλοκα σενάρια ευαισθησίας στην αφή και τη χειρονομία όχι μόνο σε υπολογιστικές συσκευές αλλά και σε αντικείμενα καθημερινής χρήσης. Επίσης έγινε προσπάθεια να χρησιμοποιηθεί η παραπάνω τεχνολογία και για το ανθρώπινο σώμα σε Αλληλεπίδραση Ανθρώπου – Υπολογιστή με τα υγρά. Με βάση την έρευνα, προέκυψε ότι ο χειρισμός των αντικειμένων και των υλικών με την ευαισθησία της αφής ήταν αρκετά απλός και αποτελεσματικός. Για παράδειγμα, ένα μόνο ειδικό καλώδιο είναι αρκετό για να μετατρέψει τα αντικείμενα και το περιβάλλον αλληλεπίδρασης σε αρκετά ευαίσθητα στις χειρονομίες και την αφή.



Εικόνα 7. Αναγνώριση αντικειμένου μέσω της αφής.

Όσον αφορά το σύστημα, προτείνεται μια πρωτότυπη διαδικασία αίσθησης μέσω της χωρητικής αφής, που ονομάζεται SFCS (*Swept Frequency Capacitive Sensing*). Σε έναν κανονικό χωρητικό αισθητήρα το αντικείμενο αφής ανιχνεύεται από ένα ηλεκτρικό σήμα σε συγκεκριμένη συχνότητα και στη συνέχεια το κύκλωμα ελέγχου ανιχνεύει το επιστρεφόμενο σήμα και προσδιορίζει το συμβάν αφής που προηγήθηκε, ανιχνεύοντας τις αλλαγές στο σήμα που προκλήθηκαν από τις ηλεκτρικές ιδιότητες του ανθρώπινου χεριού που άγγιξε το αντικείμενο. Με την προτεινόμενη τεχνολογία, παρακολουθείτε η απόκριση στο χωρικό άγγιγμα του χεριού κατά μήκος ενός εύρους συχνοτήτων. Τα αντικείμενα που επηρεάζονται από ένα ηλεκτρικό σήμα αποκρίνονται διαφορετικά σε διαφορετικές συχνότητες, και γι' αυτό οι αλλαγές στο επιστρεφόμενο σήμα θα είναι διαφορετικής συχνότητας. Επομένως, αντί να μετρείται ένα μόνο γεγονός αφής, μετρείται ένα πολυδιάστατο πλήθος σημείων αφής. Στη συνέχεια χρησιμοποιούν ένα σύστημα ταξινόμησης και εκπαίδευσης/μάθησης για να δείξουν ότι μπορούν να εξάγουν αξιόπιστα δεδομένα αλληλεπίδρασης (βλέπε Εικόνα 8) και να ανιχνεύσουν όχι μόνο ότι ένα συμβάν συνέβη, αλλά και πως συνέβη. Και το πιο σημαντικό είναι ότι αυτός ο αισθητήρας είναι ουσιαστικά ένα ηλεκτρόδιο (καλώδιο) που μπορεί να ανήκει και στο ίδιο το αντικείμενο παρατήρησεις.



Εικόνα 8. Αρχιτεκτονική Touché.

Το υλικό του Touché αποτελείται από ένα αντικείμενο αλληλεπίδρασης το οποίο ενσωματώνεται στον αισθητήρα του μέσω ενός καλωδίου. Για την επεξεργασία χρησιμοποιείται ένας μικροεπεξεργαστής ARM και μια γεννήτρια σήματος. Μετά την ανίχνευση του είδους αφής που έχει παρατηρηθεί, το σήμα στέλνεται μέσω Bluetooth σε έναν υπολογιστή, ο οποίος με κατάλληλο λογισμικό το επεξεργάζεται. Περισσότερες

πληροφορίες για την αρχιτεκτονική και το υλικό περιγράφονται στη δουλειά των Sato, Poupyrev and Harrison (2012).

Τα πειράματα και οι εφαρμογές που εξετάστηκαν σε αυτό το σύστημα ήταν τα εξής: έλεγχος ευαισθησίας για το άγγιγμα καθημερινών αντικειμένων, αίσθηση διάφορων χειρονομιών του ανθρώπου, αίσθηση διάφορων στάσεων του ανθρώπινου σώματος, βελτίωση των παραδοσιακών διεπαφών αφής και αίσθηση της αλληλεπίδρασης με ασυνήθη αντικείμενα (πχ. υγρά). Στα πειράματα συμμετείχαν 2 γκρουπ των 12 ατόμων και κάθε έρευνα διεξήχθη ανεξάρτητα από την άλλη. Η διαδικασία των πειραμάτων περιελάμβανε ένα στάδιο εκπαίδευσης των συμμετεχόντων σε κάποιες «τυποποιημένες» χειρονομίες για εξεικόωση, και το στάδιο της δοκιμής, όπου εκεί γινόταν οι λήψη των δεδομένων από τα προκαθορισμένα πέντε σενάρια πάνω στα οποία δούλευαν οι συμμετέχοντες και η επεξεργασία τους από το σύστημα. Τα συμπεράσματα που προέκυψαν ήταν πολύ ενθαρρυντικά και οι συμμετέχοντες έμειναν αρκετά ικανοποιημένοι από την απόκριση του συστήματος. Επίσης διαπιστώθηκε ότι ήταν δύσκολο να ανιχνευτούν με ακρίβεια οι πιο χαρακτηριστικές συχνότητες που επηρεάζουν τα αποτελέσματα όσον αφορά το υλικό, την αλληλεπίδραση, την εφαρμογή, το χρήστη και το περιεχόμενο, αλλά αυτό αφέθηκε σαν μελλοντική δουλειά.

### *Χρωματιστό γάντι των Wand and Popovic*

Μια άλλη ενδιαφέρουσα εμπορική τεχνολογία παρουσιάζεται στο (Wang and Popovic, 2009) από το MIT της Αμερικής, στην οποία προτείνεται ένας νέος, εύχρηστος και φθηνός οικονομικά τρόπος αλληλεπίδρασης του χρήστη με ένα σύστημα μέσω χειρονομιών. Στο σύστημα αυτό χρησιμοποιείται μια κάμερα για να εντοπίσει το χέρι το οποίο φοράει ένα κανονικό γάντι, στο οποίο εντυπώνεται ένα συγκεκριμένο χρωματιστό μοτίβο, το οποίο είναι σχεδιασμένο ώστε να απλοποιεί την εκτίμηση της θέσης του χεριού και να κάνει πιο αποτελεσματική την ανίχνευση του χεριού σε ρυθμούς πραγματικού χρόνου.

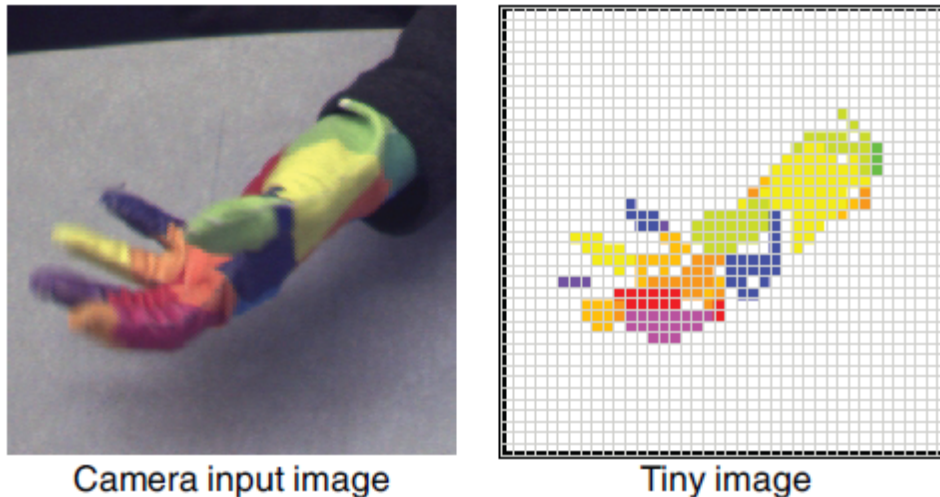


Εικόνα 9. Σύστημα αλληλεπίδρασης με χρωματιστό γάντι.

Πιο συγκεκριμένα, δημιούργησαν ένα γάντι το οποίο αποτελεί τη διεπαφή με το χρήστη και το οποίο ανάλογα με τη θέση που βρίσκεται ψάχνει στη βάση δεδομένων για να βρει την αντίστοιχη εικόνα. Αυτή η βάση δεδομένων δημιουργήθηκε δειγματοληπτώντας και καταγράφοντας διάφορες φυσικές θέσεις των χεριών και αρχειοθετώντας τις με κάποιο δείκτη. Το γάντι αποτελείται από 20 χρωματιστά κομμάτια τυχαία κατανομημένα και με 10 διαφορετικά χρώματα. Ο σκοπός είναι τα χρώματα αυτά να αναγνωρίζονται εύκολα από την κάμερα. Αυτό φαίνεται και από τα αποτελέσματα της συγκεκριμένης μεθόδου, όπου η μέθοδος εκτίμησης της θέσης του χεριού είναι πολύ αποδοτική στο να βρίσκει τη σωστή θέση σε κάθε πλαίσιο εικόνας.

Για τη δημιουργία της βάσης δεδομένων, μετέτρεψαν πρώτα τις εικόνες εισόδου της κάμερας σε μικρές εικόνες, τις φίλτράρισαν και στη συνέχεια ταξινόμησαν το κάθε εικονοστοιχείο είτε ως στοιχείο του φόντου, είτε ως ένα από τα 10 χρώματα που υπάρχουν στο γάντι. Στη συνέχεια, και αφού είχε γίνει ο διαχωρισμός των εικονοστοιχείων που ανήκουν ή όχι στο γάντι, έκοψαν και απομόνωσαν τα εικονοστοιχεία που ανήκαν μόνο στο γάντι. Στη συνέχεια, για να μπορέσουν να προσδώσουν στον αλγόριθμο εκτίμησης της θέσης του χεριού πραγματικούς χρόνους εκτέλεσης, χρησιμοποίησαν βελτιστοποιημένες αναζητήσεις στη βάση δεδομένων χρησιμοποιώντας ένα look up table με την τεχνική του κοντινού γείτονα. Χρησιμοποιήθηκε επίσης αντιστροφή κινηματική για να βελτιώσει την ακρίβεια των αποτελεσμάτων δημιουργώντας διάφορους περιορισμούς προβολής του χεριού, καθώς και ένας προσωρινός όρος εξομάλυνσης στην διαδικασία αντιστροφής κινηματικής για να μειώσει το jitter που υπάρχει στον οπτικό άξονα της κάμερας.

Το πειραματικό σύστημα που χρησιμοποιήσαν αποτελείται από μια κάμερα με φακό 4 χλστ. και μέγεθος εικόνας 640x480 εικονοστοιχεία. Χρησιμοποιήθηκε το Matlab και κάποιες βιβλιοθήκες του για να κάνουν calibration στην κάμερα, καθώς και ένα μοντέλο 26 βαθμών ελευθερίας τρισδιάστατου χεριού. Όλος ο αλγόριθμος υλοποιήθηκε σε κώδικα σε ένα σύστημα απλού τετραπύρηνου επεξεργαστή στα 2.4GHz. Η βάση δεδομένων αποτελούνταν από 100000 δείγματα, ενώ χρησιμοποιήθηκαν συγκεκριμένες τεχνικές βελτιώσεις του χρόνου αναζήτησης στη βάση δεδομένων, όπως η απόσταση Hamming. Επειδή όμως η μέθοδος αυτή δεν παρέχει μεγάλη ακρίβεια, χρησιμοποιήθηκε επίσης η αντίστροφη κινηματική για να βαθμολογήσει τις διαφορές μεταξύ της εικόνας της βάσης δεδομένων και της πραγματικής εικόνας του χεριού.

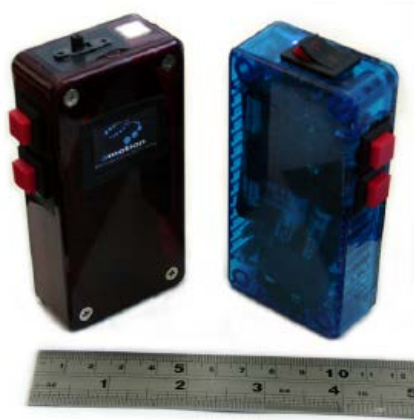


Εικόνα 10. Εκτίμηση της θέσης του χεριού.

Τα πειραματικά αποτελέσματα του παραπάνω συστήματος έγιναν για διάφορες αλληλουχίες δοκιμών διάρκειας 10 δευτερολέπτων το καθένα, αποτελούμενα από διαφορετικές θέσεις τοποθέτησης του χεριού. Με βάση τα πειράματα, προκύπτει ότι ο αλγόριθμος αυτός μπορεί να αναγνωρίσει αξιόπιστα και γρήγορα αρκετά περίεργες θέσεις του χεριού με γρήγορες κινήσεις και με κάποιο βαθμό επικάλυψης. Επίσης, με την εισαγωγή της αντίστροφης κινηματικής ελαχιστοποιήθηκαν τα σφάλματα επαναπροβολής του χεριού, ενώ αντίθετα το jitter δεν αφαιρέθηκε σε αρκετά μεγάλο ποσοστό. Επίσης, στη δουλειά αυτήν παρουσιάζονται και τρεις εφαρμογές όπου εφαρμόστηκε η συγκεκριμένη τεχνική με επιτυχία, αποδεικνύοντας ότι είναι μια αρκετά υποσχόμενη μέθοδος για το μέλλον.

### 3Motion

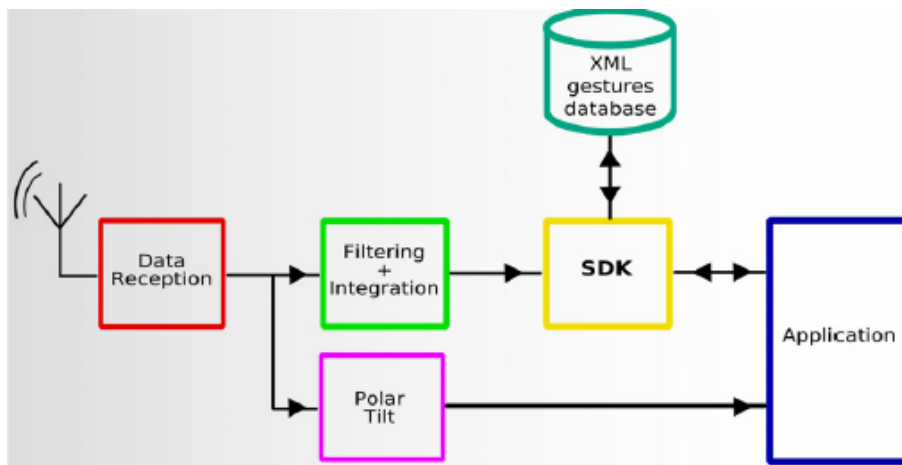
Μια άλλη ενδιαφέρουσα τεχνική παρουσιάζεται στο (Keir et. al., 2006) με το όνομα 3motion<sup>TM</sup>, η οποία αναφέρεται σε ένα σύστημα τρισδιάστατης αλληλεπίδρασης χειρονομιών αποτελούμενο από χαμηλού κόστους υλικά και από ένα γενικής χρήσης λογισμικό. Αυτό το σύστημα προτείνεται ότι είναι καλό να χρησιμοποιηθεί στις περιπτώσεις όπου τα κλασικά συστήματα παρακολούθησης είναι πολύ δαπανηρά ή μη πρακτικά εξαιτίας των απαιτήσεων τους. Το σύστημα αποτελείται από ένα επιταχυνσιόμετρο τριών αξόνων το οποίο μεταδίδει διαρκώς δεδομένα σε μια συσκευή μέσω Bluetooth. Το λογισμικό λαμβάνει το σήμα και τα δεδομένα και προσπαθεί να τα ταιριάξει με τα στοιχεία που έχει σε μια βάση δεδομένων με τρισδιάστατες χειρονομίες, για να ενεργοποιήσει την κατάλληλη λειτουργία.



Εικόνα 11. Το σύστημα 3motion™.

Πιο συγκεκριμένα, το σύστημα που παρουσιάζεται αποτελείται από τρία μέρη, από την κινητή συσκευή ανίχνευσης του χεριού, το λογισμικό αναγνώρισης χειρονομιών και την εφαρμογή. Στην παρακάτω εικόνα φαίνεται το μπλοκ διάγραμμα αυτού του συστήματος. Ο ηλεκτής κινήσεων του χεριού αποτελείται από ένα επιταχυνσιόμετρο που μεταδίδει τις μετρήσεις σε πραγματικό χρόνο σε μια πλατφόρμα. Αυτή η πλατφόρμα μπορεί να είναι ένας υπολογιστής, ένα κινητό τηλέφωνο, μια παιχνιδιομηχανή, κ.α. Στην πλατφόρμα εκτελείται ένα πρόγραμμα ανασύνθεσης της απεσταλμένης πληροφορίας, η οποία συγκρίνεται με τα δεδομένα που υπάρχουν ήδη σε μια βάση δεδομένων XML για να βρεθεί η αντίστοιχη κίνηση που έγινε. Αυτό το σύστημα είναι ανεξάρτητο από τη συσκευή που θα χρησιμοποιηθεί (πχ. μπορεί να γίνει χρήση του ποντικιού ή άλλων χωρικών ανιχνευτών) και αυτό αποτελεί ένα από τα μεγαλύτερα πλεονεκτήματά του.

Το υλικό του συστήματος είναι αρκετά προσιτό σε κόστος λόγω του εμπορικού προσανατολισμού του συστήματος, και αποτελείται από ένα σύνολο αισθητηρίων και διεπαφών. Πιο αναλυτικά, αποτελείται από τρεις αισθητήρες αναγνώρισης της γωνίας με χρήση γυροσκοπίου, μαγνητόμετρα για τον προσδιορισμό της περιστροφικής κίνησης και επιταχυνσιόμετρα για τον προσδιορισμό της ταχύτητας. Τα σήματα από αυτά τα αισθητήρια εισάγονται στο σύστημα μέσω ενός PCB chip, τα οποία επεξεργάζονται και συγκρίνονται με άλλα καταχωρημένα σήματα για τον προσδιορισμό της κίνησης και της θέσης.



Εικόνα 12. Σχεδιάγραμμα του συστήματος 3motion™.

Το υλικό του συστήματος αποτελείται από chip της STmicroelectronics, το οποίο παρέχει τρεις αναλογικές εξόδους για τα αισθητήρια και ασύρματη σύνδεση για επικοινωνία με τη συσκευή του χειριστήριου. Το λογισμικό που χρησιμοποιείται για τη διαχείριση αυτής της συσκευής είναι προγραμματισμένο για να κάνει και σωστή διαχείριση της κατανάλωσης ενέργειας, ενώ με την υποστήριξη του Bluetooth είναι δυνατή η επικοινωνία μεταξύ της συσκευής ανίχνευσης της κίνησης και της συσκευής ελέγχου (πχ. Η/Υ, κινητό τηλέφωνο, κτλ). Η τροφοδοσία γίνεται από μπαταρίες ή από ρεύμα, ενώ η απόδοση του συστήματος αναφέρεται σαν αξιόπιστη και συνεχής, χωρίς διακοπές στην επικοινωνία μεταξύ των συσκευών.



Το λογισμικό που χρησιμοποιείται για την αναγνώριση και σύγκριση των χειρονομιών χρησιμοποιεί έναν μη παραμετρικό αλγόριθμο εύρεσης και αντιστοίχισης καμπυλών που έχει χρησιμοποιηθεί προηγουμένως και σε άλλες εφαρμογές. Για κάθε καμπύλη, ο αλγόριθμος υπολογίζει τον προσανατολισμό της κίνησης, την θέση και το μέγεθος της. Πιο μπροστά βέβαια, έχει προηγηθεί η συλλογή και ταξινόμηση κάποιων χειρονομιών που έχουν αποθηκευτεί σε μια βάση δεδομένων για τη διαδικασία της αντιστοίχισης. Το λογισμικό είναι ένα SDK που περιγράφεται στο (Keir et. al., 2006) και υλοποιήθηκε στη C++, χρησιμοποιώντας και εξάγωντας τα αποτελέσματα σε XML μορφή για μεγαλύτερη αναγνωσιμότητα, αξιοπιστία και αξιολόγηση.

Το πειραματικό μέρος του συστήματος περιελάμβανε δοκιμές σε διάφορες εφαρμογές, εκ των οποίων τρεις περιγράφονται στο (Keir et. al., 2006). Δοκιμάστηκε το σύστημα σε ένα σενάριο αλληλεπίδρασης με ένα παιχνίδι τρισδιάστατων γραφικών, όπου ο χρήστης έπρεπε να αμυνθεί ή να επιτεθεί με κάποιες κινήσεις, σε ένα άλλο σενάριο περισσότερο φυσικών κινήσεων και χειρονομιών όπου ο χρήστης έπρεπε να αντιδράει σε ενέργειες που γινόντουσαν από το περιβάλλον δοκιμής (για παράδειγμα ένας χαρακτήρας του περιβάλλοντος πετούσε ένα αντικείμενο και ο χρήστης έπρεπε να αντιδράσει αναλόγως), ενώ το τρίτο σενάριο προέβλεπε τη χρήση κινητού τηλεφώνου με λογισμικό Symbian σε συσκευή ελέγχου, όπου το κινητό τηλέφωνο λειτουργούσε σαν το μαστούνι του γκολφ και ο χρήστης έπρεπε με βάση το σενάριο να κάνει συγκεκριμένες κινήσεις και βολές. Το πρόγραμμα για αυτήν την περίπτωση γράφτηκε σε C++ χρησιμοποιώντας τη βιβλιοθήκη γραφικών GapiDraw. Τα συμπεράσματα, με βάση τα πειράματα, έδειξαν ότι το σύστημα αυτό είναι αρκετά αξιόπιστο, έχει γρήγορες ανταποκρίσεις και απαιτεί μικρό χρονικό διάστημα εκπαίδευσης του χρήστη για να αρχίσει να το χρησιμοποιεί.



Εικόνα 13. Παραδείγματα χρήσης του 3motion™.

#### *Το σύστημα των Manresa, Varona, Mas and Perales*

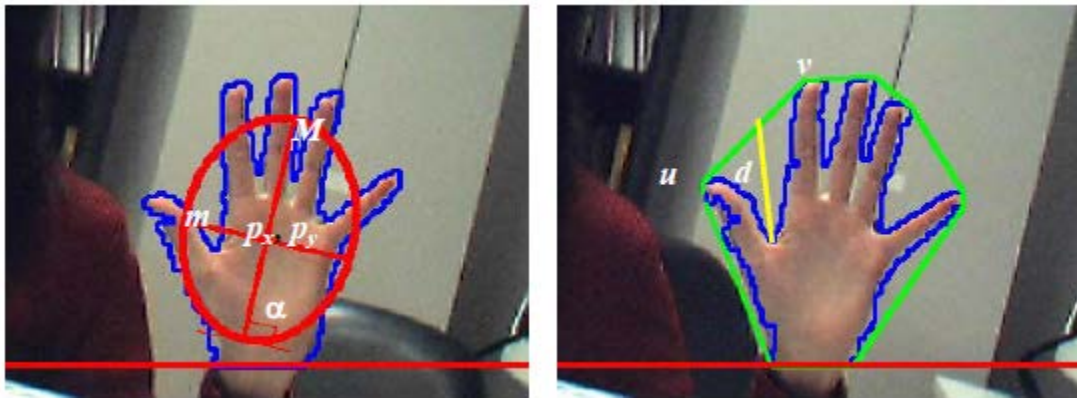
Μια άλλη δουλειά που στοχεύει να χρησιμοποιηθεί για τον έλεγχο των βιντεοπαιχνιδιών παρουσιάζεται στο (Manresa, Varona, Mas and Perales, 2005) για τον έλεγχο και την αναγνώριση χειρονομιών. Ο σκοπός της εργασίας είναι ο περιορισμός των αποκρίσεων σε πραγματικό χρόνο και ο έλεγχος των μη περιορισμένων περιβάλλοντων. Ο αλγόριθμος που παρουσιάζεται βασίζεται σε τρία βήματα, στην τμηματοποίηση του χεριού, τον εντοπισμό και την παρακολούθηση του, και τέλος στην αναγνώριση της χειρονομίας του χεριού.

Για το μέρος της τμηματοποίησης επιλέχθηκε να χρησιμοποιηθεί διαχωρισμός με βάση το χρώμα του χεριού, καθώς αυτό αποτελεί ένα αρκετά ισχυρό χαρακτηριστικό αναγνώρισης. Για τον εντοπισμό του χεριού και τον περιορισμό των σφαλμάτων από την προηγούμενη διαδικασία εισήχθη ο έλεγχος της ταχύτητας του μοντέλου και η ονοματοποίηση των εικονοστοιχείων της εικόνας. Από τη προηγούμενη διαδικασία εξάγονται αρκετά χαρακτηριστικά του χεριού τα οποία τροφοδοτούνται σε έναν ταξινομητή ο οποίος αναγνωρίζει την εκάστοτε χειρονομία και την ταξινομεί με βάση τις τέσσερις κλάσεις χειρονομιών και σε μια από τις τέσσερις διαφορετικές κατευθύνσεις κίνησης.

Πιο συγκεκριμένα, για το σύστημα αυτό χρησιμοποιήθηκε μια κάμερα του εμπορίου για την αναγνώριση των χειρονομιών στα τρισδιάστατα παιχνίδια. Ο αλγόριθμος παρέχει ταχύτητες πραγματικού χρόνου, αναγνωρίζοντας το χέρι και απομονώνοντας το από το φόντο. Αυτό γίνεται με βάση το χρώμα του δέρματος του χεριού, ενώ λαμβάνοντας υπόψιν ότι το πρόσωπο έχει επίσης το ίδιο χρώμα με το χέρι, οι συγγραφείς

ισχυρίζονται ότι μπορούν να επεκτείνουν την εφαρμογή αυτή και στην αναγνώριση διάφορων εκφράσεων του προσώπου. Για την τμηματοποίηση του χεριού χρησιμοποιήθηκε μια χαμηλού υπολογιστικού κόστους μέθοδος βασισμένη σε ένα πιθανοτικό μοντέλο της κατανομής των χρωμάτων των εικονοστοιχείων του χεριού. Για να γίνει αυτό ο χρήστης βάζει το χέρι του μπροστά σε ένα πλαίσιο για να γίνει η αναγνώριση του χρώματος του δέρματος (και αυτό αποτελεί και έναν περιορισμό του συστήματος), ενώ στη συνέχεια τα χρώματα των εικονοστοιχείων του δέρματος του μετατρέπονται σε κατάλληλες τιμές στο HSL χρωματικό μοντέλο για να μπορέσουν να επεξεργαστούν στο επόμενο επίπεδο.

Για τον εντοπισμό και την ανίχνευση του χεριού χρησιμοποιήθηκε μια απλή κάμερα USB και ένας αλγόριθμος εντοπισμού και διαρκείς ανανέωσης της θέσης του χεριού μετρώντας την απόσταση που διανύει κάθε φορά (σε κάθε πλαίσιο) το χέρι. Τέλος, για την αναγνώριση της χειρονομίας χρησιμοποιήθηκαν τέσσερα μοντέλα χειρονομιών και τέσσερις γνωστές κατευθύνσεις στις οποίες θα μπορούσε να κινηθεί. Για την κάθε χειρονομία προσδόθηκε και μια εντολή στο σύστημα (πχ. όταν το χέρι θα είναι ανοιχτό, θα σημαίνει ότι πρέπει να εκτελεστεί η εντολή Start). Επίσης ελήφθη υπόψιν και η περιστροφή του χεριού με αντίστοιχες εντολές. Για την αναγνώριση των χειρονομιών χρησιμοποιήθηκε το contour και το convex hull του χεριού με τα οποία αναγνωρίστηκαν τα εικονοστοιχεία που αντιστοιχούν στο ίδιο μέρος του χεριού σε διαδοχικά πλαίσια εικόνων.



Εικόνα 14. Εντοπισμός του χεριού.

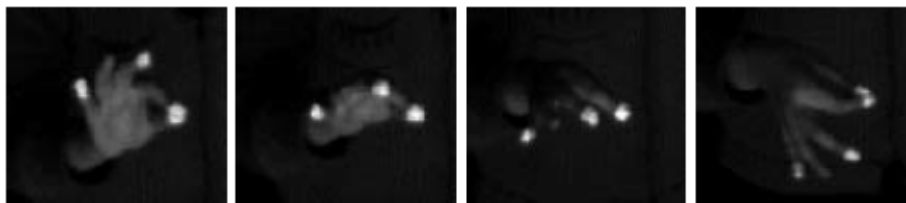
Η πειραματική πλατφόρμα που χρησιμοποιήθηκε βασίστηκε στη C++ και στις βιβλιοθήκες της OpenCV, ενώ το υλικό του συστήματος περιλαμβάνει την κάμερα και έναν υπολογιστή με επεξεργαστή στα 1.8 GHz. Για τα πειράματα χρησιμοποιήθηκαν 24 χρήστες, πραγματοποιώντας ο καθένας 40 διαφορετικές κινήσεις, επομένως στο σύστημα εισήχθησαν συνολικά 960 διαφορετικές χειρονομίες, με διαφορετικά χρώματα χεριών και διαφορετικό τρόπο εκτέλεσης της κάθε χειρονομίας. Μετά την ανάλυση των αποτελεσμάτων προέκυψε ότι το 99% των περιπτώσεων αναγνωρίστηκε με ακρίβεια, ενώ επίσης και η απόκριση του συστήματος ήταν αρκετά ικανοποιητική.



Εικόνα 15. Αναγνώριση διαφόρων χειρονομιών.

Μια άλλη ενδιαφέρουσα δουλειά παρουσιάζεται στο (Kim and Fellner, 2004), η οποία διαφοροποιείται από τις υπόλοιπες ως προς τον τρόπο που αλληλεπιδρά με το χρήστη. Στη μέθοδο αυτήν χρησιμοποιείται ένας τοίχος οπίσθιας προβολής ο οποίος αλληλεπιδρά με διάφορες χειρονομίες των χεριών και επιτρέπει το χειρισμό αντικειμένων όπως για παράδειγμα την μετατροπή, την περιστροφή, τη σύμκρυνση και τη μεγένθυση τους, καθώς και διάφορες διαδικασίες επιλογής τους μέσω ενός οπτικού συστήματος παρακολούθησης και εντοπισμού του χεριού. Για το λόγο αυτό χρησιμοποιούν μια νέα τεχνική εντοπισμού των άκρων των δακτύλων με κάποιους δείκτες που αποτελούνται από άσπρο χαρτί σε συνδυασμό με μαύρο φωτισμό.

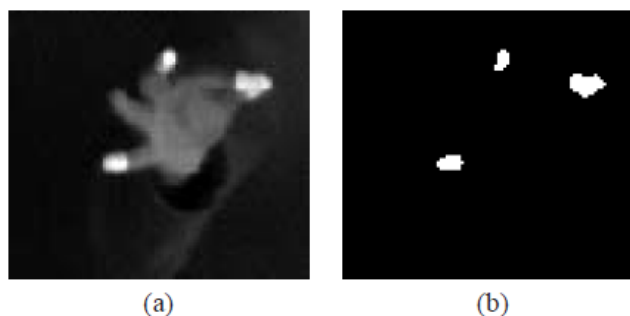
Πιο συγκεκριμένα, για το σύστημα αυτό χρησιμοποιείται ένα μεγάλο σύνολο στοιχείων αλληλεπίδρασης που συνδυάζεται με μια μεγάλη οθόνη προβολής για να δημιουργήσει ένα περιβάλλον αλληλεπίδρασης με όσο το δυνατόν πιο ρεαλιστικά χαρακτηριστικά. Για να μην περιοριστούν τα χαρακτηριστικά και η ποιότητα της εικόνας της κάμερας, χρησιμοποιήθηκε μια πηγή μαύρου φωτισμού, το οποίο αποτελεί και ένα από τα υπερίωδη μήκη κύματος τα οποία δεν μπορούν να γίνουν αντιληπτά από το ανθρώπινο μάτι. Αντίθετα, μπορούν και αλληλεπιδρούν με κάποια άλλα υλικά και αυτός ήταν και ο αντικειμενικός σκοπός αυτής της μεθόδου. Δηλαδή να χρησιμοποιηθούν δείκτες (markers) στα δάκτυλα των χεριών από υλικά που θα αλληλεπιδρούν με την υπερίωδη ακτινοβολία του μαύρου φωτός (πχ. άσπρο χαρτί κολλημένο στα δάκτυλα των χεριών), αναδεικνύοντας εντονότερα τα χαρακτηριστικά του χεριού και τις αντίστοιχες κινήσεις του. Εδώ θα πρέπει να προσεχτεί το γεγονός ότι κάποια χρώματα ρούχων (πχ. άσπρες μπλούζες) αλληλεπιδρούν εξίσου με το υπερίωδες φως, οπότε θα πρέπει οι χρήστες να φορούν ρούχα διαφορετικού χρώματος από αυτά που αναγνωρίζονται από το σύστημα (περιορισμός του συστήματος).



Εικόνα 16. Οι άκρες των δακτύλων και οι markers.

Στις προδιαγραφές του συστήματος αναφέρεται ότι θα πρέπει να θέτει ελάχιστους περιορισμούς στο περιβάλλον εφαρμογής και στις κινήσεις του χρήστη, θα πρέπει να μπορεί να εφαρμοστεί σε διάφορα περιβάλλοντα και να επιτρέπει τους χρήστες να αλληλεπιδρούν μαζί του χωρίς προηγούμενη ειδική εκπαίδευση. Σαν διεπαφή αλληλεπίδρασης επιλέχθηκε το γυμνό χέρι (δηλαδή χωρίς κάποιο γάντι), ενώ ο εντοπισμός και η ανίχνευση του χεριού θα γίνεται από ένα οπτικό σύστημα εντοπισμού. Το σύστημα αυτό μπορεί να εφαρμοστεί και σε άλλα περιβάλλοντα, αρκεί να γίνει αντίστοιχη αλλαγή στους markers και στην πηγή φωτισμού.

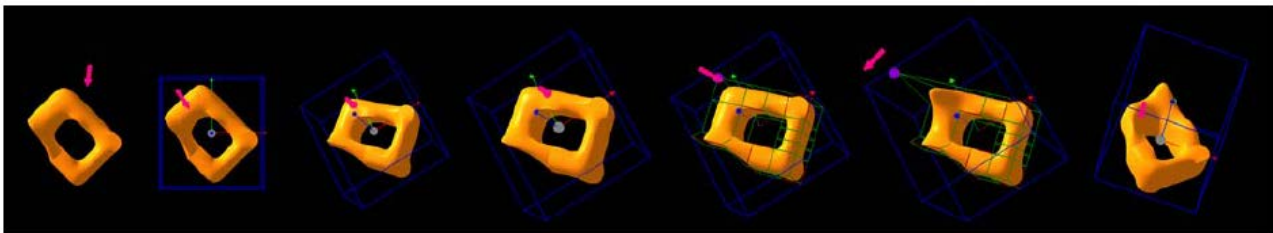
Επομένως, μετά την ανάδειξη των χαρακτηριστικών των δακτύλων μέσω αυτού του ιδιαίτερου συστήματος φωτισμού, το σύστημα εντοπισμού χρησιμοποιεί την κάμερα για να εντοπίσει αυτά τα χαρακτηριστικά, να τα τμηματοποιήσει χρησιμοποιώντας έναν αλγόριθμο κατωφλίου, και στη συνέχεια να κάνει αντιστοίχιση της εκάστοτε χειρονομίας με έναν μετρικό δείκτη που θα υποδεικνύει την τρισδιάστατη θέση του χεριού. Τέλος, ένας αλγόριθμος εντοπισμού σημείων στην εικόνα μειώνει τις πλεονάζουσες πληροφορίες που υπάρχουν στην εικόνα για να επιτευχθούν ταχύτητες πραγματικού χρόνου, ενώ με ένα φίλτρο Kalman εκτιμάται η κίνηση του χεριού μέσω της ανίχνευσης των markers που υπάρχουν στα δάκτυλα.



Εικόνα 17. Αναγνώριση (a) και τμηματοποίηση (b) των μέρων του χεριού που έχουν τους markers.

Για τα πειράματα του παραπάνω συστήματος χρησιμοποιήθηκε ένα σύστημα που αποτελούταν από έναν σταθμό εργασίας (workstation) με μια κάρτα ανίχνευσης εικόνων πλαισίου και δύο κάμερες σταθμικής σάρωσης. Το σύστημα μπορεί να λειτουργεί σε πραγματικό χρόνο με 21-30 πλαίσια/δευτερόλεπτο και περιορίζεται μόνον από τα χαρακτηριστικά των καμερών.

Οι δοκιμές έγιναν μέσω της αλληλεπίδρασης τρισδιάστατων αντικειμένων, τα οποία ο χρήστης μπορούσε να επεξεργαστεί με συγκεκριμένες κινήσεις, όπως φαίνεται στην παρακάτω εικόνα. Παρόλο που το σύστημα αλληλεπιδρούσε καλά στις περισσότερες των περιπτώσεων, κάποιες φορές η αλλαγή στη φάση της χειρονομίας δημιουργούσε κάποια προβλήματα, καθώς και ο περιορισμένος αριθμός markers στα χέρια λόγω απόκρυψης/επικάλυψης (occlusion). Επίσης, κατά τη διάρκεια των περιστροφών των αντικειμένων παρατηρήθηκαν επίσης κάποια προβλήματα, κυρίως όμως λόγω των μεγάλων γωνιών περιστροφής.



(a)



(b)

Εικόνα 18. Χειρισμός αντικειμένων κατά την πειραματική διαδικασία (α) τα αντικείμενα αλληλεπίδρασης και (b) οι διάφορες λειτουργίες που δοκιμάστηκαν, όπως διακοπή, εντοπισμός(αντικειμένου), περιστροφή, μεγένθυση, κατάδειξη, μετατροπή και περιστροφή.

Σε γενικές γραμμές, το εν λόγω σύστημα παρουσιάζει καλή απόκριση σε πραγματικό χρόνο, έχει ευρύ πεδίο εφαρμογής λόγω των ελάχιστων περιορισμών που έχει, αλλά θα πρέπει να δοθεί προσοχή σε κάποιες περιπτώσεις όπου το σύστημα δεν βγάζει ακριβή αποτελέσματα ή δημιουργείται αβεβαιότητα ως προς το σωστό εντοπισμό των markers των χεριών για την ταξινόμηση, κατηγοριοποίηση και ερμηνεία των αντίστοιχων χειρονομιών.

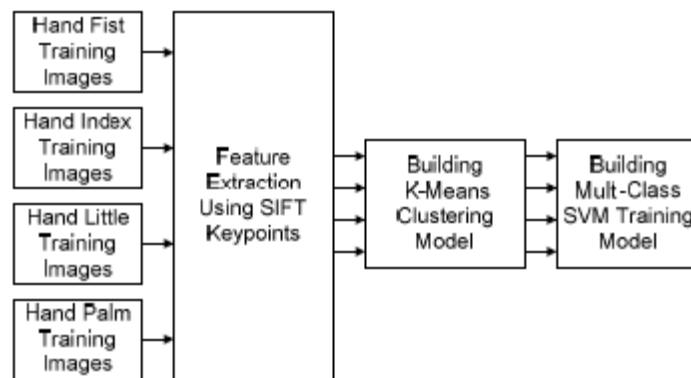
*Το βασισμένο σε SVM σύστημα των Dardas and Alhaj*

Μια άλλη ενδιαφέρουσα μελέτη παρουσιάζεται στο (Dardas and Alhaj, 2011) αναφορικά με ένα σύστημα που υλοποιεί μια διεπαφή αναγνώρισης κινήσεων του χεριού που δημιουργεί συγκεκριμένες εντολές για να ελέγχει αντικείμενα απευθείας σε ένα παιχνίδι (περιβάλλον δοκιμής). Το προτεινόμενο σύστημα χρησιμοποιεί κάποια εξειδικευμένα χαρακτηριστικά (Bag-of) και ένα SVM (Support Vector Machine) για να δημιουργήσει ένα φιλικό και εύχρηστο περιβάλλον ελέγχου του παιχνιδιού.

Πρωτό όμως γίνει αναφορά στον τρόπο λειτουργίας του συστήματος, θα πρέπει να αναφέρουμε τους τρεις λόγους για τους οποίους οι συγγραφείς θεωρούν ότι οι εφαρμογές αναγνώρισης κινήσεων/χειρονομιών των χεριών υπερέχουν σε πλεονεκτήματα έναντι των υπολοίπων εφαρμογών αλληλεπίδρασης ανθρώπου-υπολογιστή. Καταρχάς γιατί μπορούν να ανακτούν πληροφορίες χωρίς να έρχονται οι χρήστες σε άμεση επαφή με το ίδιο το σύστημα (πχ. στην περίπτωση χρήσης του συστήματος για χειρουργική επέμβαση). Δεύτερον γιατί μπορούν να βοηθούν τους ανάπηρους ώστε να αλληλεπιδρούν με τις διάφορες οικιακές συσκευές τους χωρίς να χρειάζεται να μετακινούνται, και τρίτον, γιατί μπορούν οι χρήστες να περιηγούνται σε μεγάλο όγκο δεδομένων και να διαχειρίζονται καλύτερα τις τρισδιάστατες πληροφορίες που υπάρχουν σ' αυτόν.

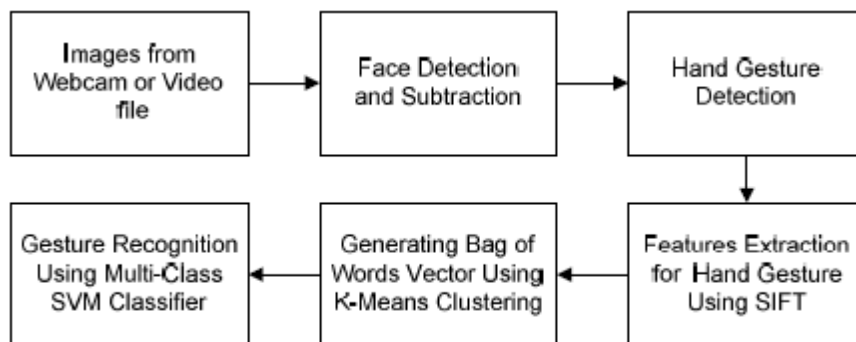
Συνεπώς, στα πλαίσια ανάπτυξης αυτού του συστήματος, αναπτύχθηκε ένα παιχνίδι προσομοίωσης σε Java 3D για να προσομοιώσει το πειραματικό περιβάλλον. Το σύστημα αναγνώρισης χειρονομιών χωρίζεται σε δύο στάδια, το στάδιο εκπαίδευσης και το στάδιο δοκιμής. Τα τμηματοποιημένα και ομαδοποιημένα δεδομένα του SVM δημιουργούνται στο στάδιο της εκπαίδευσης και χρησιμοποιούνται στο στάδιο της δοκιμής για την αναγνώριση των χειρονομιών.

Για την αναγνώριση των χειρονομιών ελήφθησαν 100 εκπαιδευτικές εικόνες από κάθε χέρι από διαφορετικούς ανθρώπους και με διαφορετικές περιβαλλοντικές συνθήκες και εισήχθησαν στον ταξινομητή του SVM. Τα σημεία κλειδιά της κάθε χειρονομίας αναγνωρίζονται και εξάγονται μέσω του αλγόριθμου SIFT σε πραγματικό χρόνο. Επίσης, χρησιμοποιώντας έναν αλγόριθμο εξαγωγής χαρακτηριστικών το σύστημα μαθαίνει ένα «εικονικό λεξιλόγιο» χρησιμοποιώντας τον αλγόριθμο τμηματοποίησης K-means, βρίσκοντας ταυτόχρονα και τις συχνότητες εμφάνισης της κάθε «εικονικής λέξης».



Εικόνα 19. Στάδιο εκπαίδευσης.

Στο στάδιο της δοκιμής, αφού το σύστημα λάβει τις εικόνες από την κάμερα, αναγνωρίζει πρώτ' απ' όλα το πρόσωπο του χρήστη και το διαχωρίζει ώστε να μπορέσει να αναγνωρίσει το χρώμα του δέρματος. Στη συνέχεια η εικόνα τμηματοποιείται και για κάθε μικρό μέρος της εξάγονται τα σημεία κλειδιά τα οποία εμπεριέχονται στις χειρονομίες. Τα σημεία αυτά εισάγονται στον ταξινομητή/τμηματοποιητή για να βρεθεί η αντίστοιχη «λέξη» και τέλος μέσω του SVM να αναγνωριστεί η αντίστοιχη χειρονομία.



Εικόνα 20. Στάδιο δοκιμής.

Στα πειράματα που έγιναν για την παραπάνω μέθοδο χρησιμοποιήθηκαν μόνον τέσσερις χειρονομίες για την ταξινόμηση (φαίνονται στην παρακάτω εικόνα) και μια κάμερα του εμπορίου με επαρκή αριθμό πλαισίων ανά δευτερόλεπτο για ανάλυση και επεξεργασία σε πραγματικό χρόνο. Οι χειρονομίες περιελάμβαναν διαφορετικά μεγέθη χειρονομιών (μεγένθυση-σμίκρυνση), περιστροφή και διαφορετικές συνθήκες φωτισμού, ενώ συνολικά ελέγχθησαν 4000 διαφορετικές εικόνες.

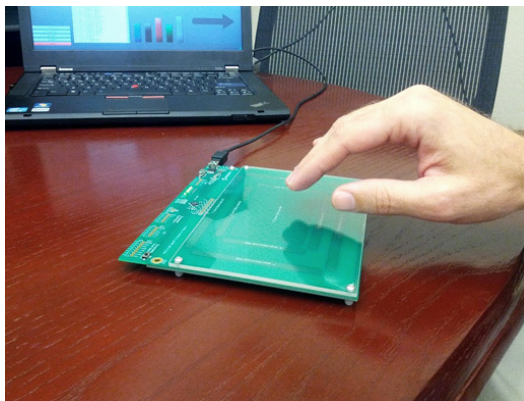


Εικόνα 21. Οι τέσσερις χειρονομίες που μπορεί να ανιχνεύσει το σύστημα.

Τα αποτελέσματα έδειξαν ότι ο χρόνος λήψης και επεξεργασίας/αναγνώρισης των χειρονομιών είναι αρκετά γρήγορος για εφαρμογές πραγματικού χρόνου, ενώ δοκιμάστηκαν διάφοροι τρόποι δημιουργίας εντολών μέσω των χειρονομιών, όπως η αναγνώριση χειρονομιών μεταξύ των θέσεων του χεριού, με βάση τη διεύθυνση της κίνησης του χεριού, καθώς και η απόσταση του χεριού από την κάμερα. Στο περιβάλλον προσομοίωσης, ο χρήστης μπορούσε να δημιουργήσει εντολές στο πληκτρολόγιο με βάση την αντίστοιχη χειρονομία και να κινήσει ένα ελικόπτερο, ενώ για τη διασύνδεση των αρχείων με το σύστημα χρησιμοποιήθηκε C# και βιβλιοθήκες τις C/C++. Σαν τελικό συμπέρασμα οι χρήστες έμειναν αρκετά ικανοποιημένοι από την απόκριση του συστήματος και την ακρίβεια των κινήσεων, ενώ για να μπορέσει να επεκταθεί η εφαρμογή του, θα πρέπει να υλοποιηθούν διαφορετικά περιβάλλοντα προσομοιώσεις με διαφορετικές απαιτήσεις και φυσικά να χρησιμοποιηθούν περισσότερες χειρονομίες για τον έλεγχο των αντικειμένων.

#### *Το βασισμένο σε ηλεκτρικά πεδία σύστημα του Leber*

Στο (Leber, 2012) παρουσιάζεται ένας νέος τρόπος αλληλεπίδρασης ανθρώπου-υπολογιστή, ο οποίος ενσωματώνει μια νέα τεχνική μέτρησης τρισδιάστατων μεγεθών μέσω ενός ελεκτή που εκμεταλλεύεται κατάλληλα τα ηλεκτρικά πεδία. Αυτός ο ελεκτή μπορεί να αλληλεπιδρά με διάφορες κινητές συσκευές (πχ. κινητά τηλέφωνα) οι οποίες στις μέρες μας έχουν κάποιου είδους οπτικό αισθητήρα (πχ. κάμερα) ώστε να ερμηνεύσουν κατάλληλα κάποιες χειρονομίες και κινήσεις του χεριού. Ο περιορισμός στη συγκεκριμένη τεχνολογία είναι οι κινήσεις που μπορεί να κάνει το χέρι πάνω στην επιφάνεια αλληλεπίδρασης και το εύρος της αναγνωρίσιμης περιοχής.



Εικόνα 22. Το σύστημα διεπαφής.

Παρόλο τους περιορισμούς που έχει, οι κατασκευαστές ισχυρίζονται ότι παρουσιάζει κάποια σημαντικά πλεονεκτήματα έναντι των άλλων συστημάτων οπτικής αναγνώρισης χειρονομιών που βασίζονται στις κάμερες, όπως για παράδειγμα η χαμηλή κατανάλωση ενέργειας ή ακόμα και το είδος της εφαρμογής που θα μπορούσε να χρησιμοποιηθεί (για παράδειγμα ισχυρίζονται ότι με κατάλληλες τροποποιήσεις θα μπορούσε να ενσωματωθεί σε πλήθος κινητών συσκευών όπως τα κινητά τηλέφωνα).

Η λειτουργία του συστήματος βασίζεται στη μεταφορά ενός ηλεκτρικού σήματος, με το οποίο γίνεται ο υπολογισμός της τρισδιάστατης θέσης του χεριού με βάση την επίδραση που προκαλεί το χέρι στο σήμα. Ο ελεκτή αυτός μπορεί να λειτουργεί κάτω από συνθήκες χαμηλού φωτισμού, καθώς επίσης και όταν ο χρήστης βρίσκεται αρκετά κοντά στον αισθητήρα (σε αντίθεση με κάποια άλλα συστήματα που έχουν περιορισμούς στο εύρος του πεδίου αλληλεπίδρασης). Επίσης, λόγω της κατασκευής του, μπορεί να εφαρμοστεί σε πλήθος άλλων εφαρμογών, όπως σε αυτοκίνητα, πληκτρολόγια, διακόπτες, κτλ., δηλαδή και σε συσκευές που δεν έχουν ενσωματωμένη κάμερα.

Στην παρούσα φάση ο ελεγκτής μπορεί να αναγνωρίζει μέχρι και 10 προκαθορισμένες χειρονομίες, αλλά μπορεί να προγραμματιστεί και να προσαρμοστεί στις ανάγκες του χρήστη. Η ακρίβεια του συστήματος είναι εφάμιλλη με αυτήν του ποντικιού, αλλά ακόμα υπάρχουν κάποιιοι περιορισμοί ως προς τον τρόπο αναγνώρισης των χειρονομιών (πχ. περιορισμός στον τρόπο τοποθέτησης του χεριού) που το εμποδίζουν να αναγνωρίσει σωστά κάποιες από τις κινήσεις του χρήστη.

#### *Τηλεοράσεις με αλληλεπίδραση χειρονομιών*

Στον χώρο των τηλεοράσεων, η Hitachi στο (Saenz, 2009) παρουσιάζει μια τηλεόραση που θα μπορεί να αναγνωρίζει τις χειρονομίες των θεατών και αντί για τηλεχειριστήριο, θα μπορεί ο χρήστης να κουνά το χέρι του για να αλλάξει κανάλια ή να αυξομειώσει τον ήχο. Πιο συγκεκριμένα, μια κίνηση του χεριού θα σημαίνει ότι θέλει να εκτελέσει μια εντολή και έτσι θα ενεργοποιείται το αντίστοιχο σύστημα της τηλεόρασης. Στη συνέχεια, ανάλογα με την κίνηση που θα κάνει θα εκτελείται και η αντίστοιχη εντολή. Για παράδειγμα, μια κίνηση του χεριού προς τα εμπρός μπορεί να υποδεικνύει ότι ο χρήστης θέλει να επιλέξει το συγκεκριμένο κανάλι, ή μια περιστροφική κίνηση ότι θέλει να αυξήσει τον ήχο. Για το σκοπό αυτόν χρησιμοποιήθηκε ειδικό λογισμικό ενσωματωμένο στην τηλεόραση, ενώ η εταιρεία Canasta που ανέπτυξε το υλικό χρησιμοποίησε δύο κάμερες για να δημιουργήσει την τρισδιάστατη (στερεοσκοπική) εικόνα του ακροατηρίου.

Η Toshiba βέβαια έχει ήδη παρουσιάσει το δικό της σύστημα τηλεόρασης το 2008 (Miller, 2008) το οποίο αλληλεπιδρά και αυτό με τον χρήστη μέσω χειρονομιών, και το οποίο βασίζεται και σε πολλαπλές κάμερες για τον εντοπισμό και τη λήψη του περιβάλλοντος αλληλεπίδρασης, αλλά παρουσιάζει μια διαφορά ως προς τα προηγούμενα συστήματα. Ουσιαστικά υλοποιεί ένα σύστημα με κέρσορα στην οθόνη, και ο χρήστης μέσω του χεριού του μπορεί να ελέγχει (κάνοντας τις αντίστοιχες χειρονομίες) των κέρσορα στην τηλεόραση για να εκτελέσει διάφορες εντολές. Μπορεί επίσης να σταματήσει τη ροή της τηλεόρασης μόνο με μία κίνηση του χεριού, ενώ μπορεί ακόμη και να περιστρέψει τρισδιάστατες εικόνες που εμφανίζονται στην οθόνη. Τέλος, μια ακόμα εφαρμογή που υπάρχει σε αυτό το σύστημα είναι η δυνατότητα αναγνώρισης εικόνων και η εκτέλεση προκαθορισμένων εντολών/λειτουργιών ανάλογα με την εικόνα που αναγνωρίστηκε. Για παράδειγμα, μπορεί μέσω των αντίστοιχων επιλογών που παρέχονται, ο χρήστης δείχνοντας μια πιστωτική κάρτα στην τηλεόραση (ουσιαστικά στο οπτικό σύστημα της) να γίνεται αναγνώριση της και να εκτελείται η αντίστοιχη εντολή που έχει προκαθορίσει (πχ. για παραγγελία μέσω ίντερνετ).

Επειδή στον χώρο της τηλεόρασης τα τελευταία χρόνια γίνεται μεγάλη πρόοδος και εξέλιξη στις τεχνικές αλληλεπίδρασης του ανθρώπου με αυτές, ο Saenz (2010) παρουσιάζει μια σύνοψη των τεχνολογιών αυτών που συμπεριλαμβάνουν Αλληλεπίδρασης Ανθρώπου – Υπολογιστή χρήστη και υπολογιστή (εδώ τηλεόρασης), αναφερόμενος κυρίως στις τρεις εξέχουσες εταιρείες του χώρου, όπως η Microsoft, η Toshiba και η Hitachi και στα αντίστοιχα προϊόντα τους.

#### *Build-it*

Στην εργασία στο (Rauterberg, Bichsel, Meier and Fjeld, 1997) παρουσιάζεται μια μέθοδος που προτείνει ένα νέο είδος μηχανικών κινήσεων που θα πρέπει να γίνονται από τους χρήστες μέσω χειρονομιών που αλληλεπιδρούν με τα διάφορα συστήματα HCI για να επιτύχουν κάποιους στόχους ή να εκτελέσουν συγκεκριμένες εντολές. Οι κινήσεις αυτές θα πρέπει να είναι λειτουργικές, και όχι ανατομικές ή μηχανικές και συστήνουν ότι για να επιτευχθεί σωστή διάδραση ανθρώπου-υπολογιστή οι χρήστες θα πρέπει να συμπεριφέρονται με φυσικό τρόπο, κινώντας ταυτόχρονα όλο το σώμα τους (χέρια, πόδια, πρόσωπο, κτλ), και γ'αυτό το λόγο χρησιμοποιούν ένα ευφύες σύστημα αναγνώρισης προτύπων (NUI), το οποίο υποστηρίζει τη διάχυση των πραγματικών και των εικονικών αντικειμένων, μέσω μιας γενικής σχεδιαστικής στρατηγικής που βασίζεται στην επαυξημένη πραγματικότητα. Αυτή μπορεί και αναγνωρίζει οπτικές (πχ. εικόνες), ακουστικές (πχ. φωνητικές εντολές) και άλλες μορφές εισόδου (πχ. θέση σώματος) από το χρήστη, ενώ μπορεί επίσης να αναγνωρίζει διάφορα φυσικά αντικείμενα στο χώρο ελέγχου του. Η χωρική θέση του χρήστη παρακολουθείται από μια ή περισσότερες κάμερες και ο ήχος καταγράφεται με μικρόφωνα, με αποτέλεσμα να δημιουργείται ένα εσωτερικό τρισδιάστατο μοντέλο του χρήστη. Η είσοδος του συστήματος μπορεί να δεχτεί σήματα από περισσότερους από έναν χρήστες, ενώ η έξοδος του υποστηρίζει τριών ειδών τύπους, την οθόνη, τη συσκευή προβολής που τονίζει το περιβάλλον εργασίας και ένα πολυκαναλικό

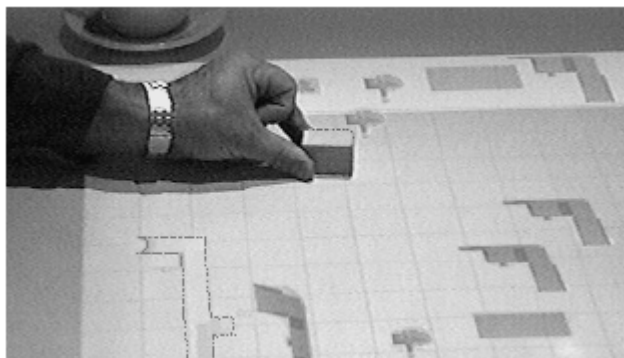
σύστημα ηχείων. Μέσω αυτών των αισθητηρίων οι χρήστες μπορούν να δίνουν εντολές στο σύστημα κάνοντας κάποιες πιο «φυσικές» και όχι τυποποιημένες χειρονομίες και να εκτελούν τις εντολές πάνω σε κάποια αντικείμενα, συνδυάζοντας φωνητικά και κινητικά ερεθίσματα.



Εικόνα 23. Χώρος εργασίας του συστήματος Build-it.

Το μοντέλο που υλοποιήθηκε ονομάζεται Build-it, και αποτελείται από ένα δωμάτιο δοκιμών, μαζί με κάποια φυσικά και εικονικά αντικείμενα. Τα χαρακτηριστικά του βασίζονται σε μια οθόνη προβολής, μια ASK 960 LCD οθόνη, κάμερα CCD και έναν επεξεργαστή γραφικών. Το λογισμικό αποτελείται από δύο ανεξάρτητες διαδικασίες που τρέχουν παράλληλα και συνδέονται μέσω ενός καναλιού επικοινωνίας, και αφορούν την ανάλυση των εικόνων βίντεο και την αναγνώριση/ερμηνεία των κινήσεων. Γι' αυτό το σκοπό χρησιμοποιήθηκε το προγραμματιστικό περιβάλλον του MET++, ενώ υποστηρίζει και CAD μοντέλα.

Όσον αφορά το πειραματικό μέρος, το σύστημα υποστηρίζει διάφορες λειτουργίες όπως για παράδειγμα η επιλογή ενός εικονικού αντικειμένου σε ένα εικονικό μαγαζί, τη μετακίνηση τους, την περιστροφή τους, κ.α. Στην παρακάτω εικόνα φαίνεται ένα τέτοιο παράδειγμα χρήσης του συστήματος. Οι δοκιμές έγιναν από μηχανικούς και διευθυντές και ήταν εμπειρικές. Με βάση τις δοκιμές, φάνηκε ότι το σύστημα ανταποκρίνεται σωστά στα περισσότερα γνωρίσματα και ερεθίσματα που δέχεται από τους χρήστες και είναι εύκολο στη μάθηση. Για παράδειγμα, οι περισσότεροι από τους χρήστες μπόρεσαν να φτιάξουν μια εικονική κατασκευή σε λιγότερο από 30 δευτερόλεπτα.



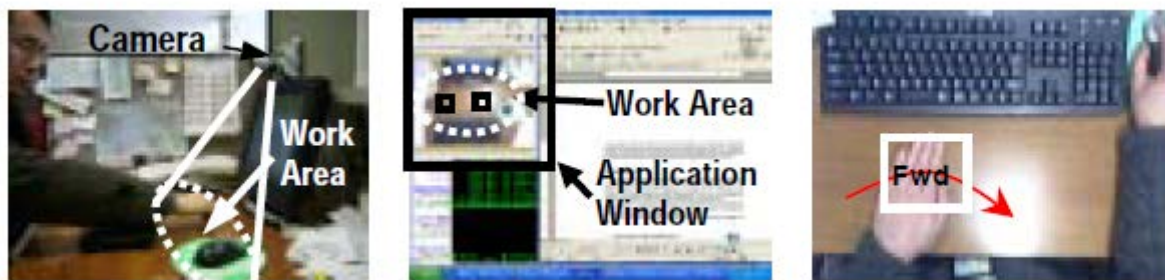
Εικόνα 24. Το μενού αντικειμένων (άσπρο), η πάνω όψη (γκρι) και το χέρι του χρήστη που μετακινεί το αντικείμενο αλληλεπίδρασης (τουβλάκι).

#### *Το βασισμένο σε hotspots σύστημα των Jaimes and Jianyi*

Σε μια άλλη εργασία των (Jaimes and Jianyi, 2005) παρουσιάζεται ένα πρωτότυπο και προσαρμόσιμο σύστημα διάδρασης ανθρώπου-υπολογιστή βασισμένο σε κάμερες, το οποίο χρησιμοποιεί κάποια σημεία ενδιαφέροντος (hotspots) για τη δισδιάστατη διάδραση με χειρονομίες.



Πιο συγκεκριμένα, μια κάμερα κοιτάει στην επιφάνεια εργασίας και η εικόνα που λαμβάνεται από την κάμερα εμφανίζεται στην οθόνη του χρήστη. Ο χρήστης ορίζει μια περιοχή αλληλεπίδρασης μέσα στην επιφάνεια εργασίας, τοποθετώντας την κάμερα έτσι ώστε να καταδεικνύει τη συγκεκριμένη περιοχή. Επίσης, ο χρήστης ορίζει νέες εντολές με την παραμετροποίηση των περιοχών hotspot, και τις εκτελεί κινώντας το χέρι του προς την ανάλογη περιοχή. Μια περιοχή ενδιαφέροντος ενεργοποιείται όταν το χέρι του χρήστη περάσει μέσα από μια ορθογώνια περιοχή (παράθυρο σάρωσης) που ορίζεται από την εικόνα της κάμερας. Για παράδειγμα, ένα σημείο ενδιαφέροντος ενεργοποιείται όταν ο χρήστης μετακινεί το χέρι του από τα δεξιά προς τα αριστερά, που εισέρχεται και εξέρχεται από την παραπάνω ορθογώνια περιοχή. Μια τέτοια εφαρμογή είναι η μετακίνηση σε μια άλλη ιστοσελίδα ή σε άλλη σελίδα ενός κειμένου, όπου για να μετακινηθεί για παράδειγμα σε μια προηγούμενη σελίδα σε ένα κείμενο, ο χρήστης θα πρέπει να κινήσει το χέρι του από τα αριστερά προς τα δεξιά κατά μήκος της περιοχής hotspot που έχει ορίσει προηγουμένως. Το συγκεκριμένο σύστημα έχει αρκετές εφαρμογές, καθώς η περιοχή ενδιαφέροντος μπορεί να προσαρμοστεί ανάλογα με την εφαρμογή, ενώ μπορούν επίσης να οριστούν και διαφορετικά hotspots με διαφορετική λειτουργία το καθένα. Για παράδειγμα, μπορούν να οριστούν διάφορα hotspots που το καθένα να αντιστοιχεί σε ξεχωριστό μουσικό ήχο. Άρα, σε μια οθόνη με πολλαπλές περιοχές hotspot, ο χρήστης μπορεί κουνώντας τα χέρια να παίζει διαφορετικές μελωδίες. Η ανίχνευση των χειρονομιών γίνεται με την αναγνώριση και ενεργοποίηση μιας σειράς από προκαθορισμένα σημεία ενδιαφέροντος. Για παράδειγμα, μια κίνηση από τα δεξιά προς τα αριστερά, μέσα στο παράθυρο σάρωσης ενεργοποιεί το αντίστοιχο hotspot δεξιά-αριστερά, εκτελώντας και την αντίστοιχη εντολή.



Εικόνα 25. Λειτουργία του συστήματος. Η κάμερα δείχνει το φυσικό περιβάλλον εργασίας στην επιφάνεια εργασίας (αριστερά). Η εικόνα της διεπαφής (κέντρο) από το χρήστη μπορεί να αναπροσαρμοστεί αναλόγως και δείχνει τις αντίστοιχες εικόνες που μπορεί να λάβει, καθώς και τα αντίστοιχα hotspots (δεξιά).

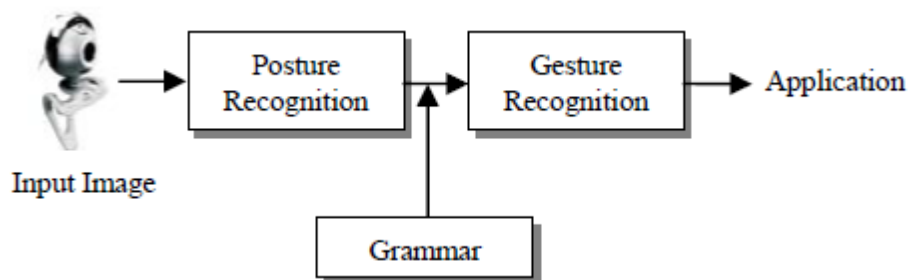
Το σύστημά αποτελείται από μια κάμερα που τοποθετείται πάνω από την επιφάνεια εργασίας για να αναγνωρίζει τις περιοχές ενδιαφέροντος, ενώ από το σύστημα εκτελείται ένας αλγόριθμος αναγνώρισης χειρονομιών για την εκτέλεση της αντίστοιχης εντολής. Η λειτουργία του βασίζεται στα εξής βήματα: φιλτράρισμα της εικόνας και περιορισμός στο μέρος του δέρματος του χεριού, ανίχνευση χρώματος δέρματος, ενεργοποίηση περιοχής ενδιαφέροντος και ανίχνευση/αναγνώριση χειρονομίας. Το σύστημα αυτό οι συγγραφείς ισχυρίζονται ότι είναι εξαιρετικά ευέλικτο και αποδοτικό, διότι επιτρέπει στο χρήστη να προσαρμόσει τη διεπαφή με νέες εντολές που μπορούν να αποδοθούν σε συγκεκριμένα σημεία ή σύνθετα σημεία βασισμένα σε διάφορες ακολουθίες. Το φυσικό περιβάλλον εργασίας στην επιφάνεια εργασίας μπορεί να οριστεί απλά τοποθετώντας σε αντίστοιχο σημείο την κάμερα, επεκτείνοντας έτσι τις δυνατότητες του συστήματος. Επίσης, το σύστημα επιτρέπει να οριστούν και να προσαρμοστούν οι περιοχές ενδιαφέροντος, ανάλογα με τα χαρακτηριστικά της κάμερα και την περιοχή που μπορεί να σαρώσει, ενώ μπορούν να αποθηκευτούν ξεχωριστά layouts για διάφορα hotspots, ανάλογα με την εφαρμογή που θέλει να εκτελέσει ο χρήστης. Το σύστημα μπορεί να δουλέψει με απλές κάμερες του εμπορίου σε πραγματικό χρόνο και να χρησιμοποιήσει τεχνικές μάθησης μηχανής για να αυτοματοποιήσει την ανίχνευση του δέρματος του χρήστη για να βελτιστοποιήσει την αναγνώριση των χειρονομιών του χεριού του.

Οι δοκιμές του παραπάνω συστήματος έγιναν χρησιμοποιώντας 12 εκπαιδευτικά μοντέλα και ελήφθησαν 32000 pixel. Το στάδιο της εκπαίδευσης του συστήματος αποτελούσε το 66% της όλης διαδικασίας, ενώ το υπόλοιπο ήταν οι δοκιμές. Σε σχέση με τις άλλες μεθόδους που έγινε η σύγκριση των αποτελεσμάτων, φαίνεται ότι το σύστημα είναι αρκετά ανταγωνιστικό, αν και τα αποτελέσματα δεν μπορούν να συγκριθούν απευθείας, καθώς η ακρίβεια του συστήματος εξαρτάται από μη συγκρίσιμα κριτήρια, όπως η ταχύτητα που έγιναν οι χειρονομίες και η ακρίβεια της εκάστοτε κίνησης. Παρ'όλ'αυτά, στις περιπτώσεις που η ταχύτητα

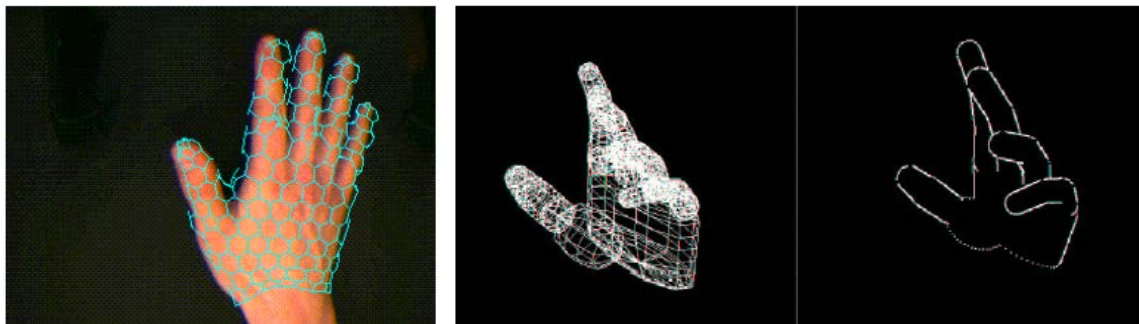
της χειρονομίας ήταν σε λογικά πλαίσια, το σύστημα παρουσίασε 100% ακρίβεια στην αναγνώριση και εκτέλεση της αντίστοιχης εντολής.

### Γάντια δεδομένων vs. Οπτικής Αναγνώρισης

Με βάση τα παραπάνω, γίνεται εύκολα αντιληπτό ότι το πιο σημαντικό πρόβλημα στα συστήματα διάδρασης με χειρονομίες είναι η υλοποίηση της αναγνώρισης των χειρονομιών από τους υπολογιστές. Γι' αυτό και πολλές μέθοδοι προτιμούν να περιορίζουν τους περιορισμούς που υπεισέρχονται από αυτή τη διαδικασία χρησιμοποιώντας κυρίως Αλληλεπίδρασης Ανθρώπου – Υπολογιστή με ένα «γάντι δεδομένων» και "Οπτική αναγνώριση", όπως αναφέρεται στο (Garg, Aggarwal and Sofat, 2009). Και αυτό γιατί το γάντι μπορεί να προσαρμοστεί ανάλογα (πχ. με κατάλληλο χρωματισμό ή αισθητήρια) ώστε να αναγνωρίζεται ευκολότερα από το σύστημα αλληλεπίδρασης. Το γάντι στη δική τους περίπτωση είναι εξοπλισμένο με ένα μεγάλο αριθμό αισθητήρων και συνδέεται με τον υπολογιστή από μια ηλεκτρική γραμμή. Ο χρήστης μπορεί απλά να σταθεί μπροστά από την οθόνη με το γάντι και να κουνήσει το χέρι του. Το γάντι συλλέγει τις κινήσεις από το χέρι και τα δάκτυλα και μεταδίδει την πληροφορία στον υπολογιστή.



Εικόνα 26. Η διαδικασία αναγνώρισης χειρονομιών.



Εικόνα 27. Η αναγνώριση και η τρισδιάστατη απεικόνιση του χεριού.

Παρά το γεγονός ότι αποδείχθηκε ότι τα γάντια παρέχουν πολύ ακριβή αποτελέσματα, οι συσκευές αυτές είναι αρκετά ακριβές και παρουσιάζουν πολλές φορές δυσάρεστη εμπειρία για τους χρήστες. Για παράδειγμα, το καλώδιο επικοινωνίας περιορίζει το πεδίο των δραστηριοτήτων του χρήστη ενώ το μέγεθος του γαντιού πολλές φορές δεν ταιριάζει σε διαφορετικούς χρήστες.

Σε αντίθεση, η οπτική Αλληλεπίδρασης Ανθρώπου – Υπολογιστή έχει αποκτήσει μεγαλύτερη σημασία τα τελευταία χρόνια γιατί δεν εμφανίζει τους παραπάνω περιορισμούς. Τα συστήματα αυτά μπορούν να έχουν μία ή περισσότερες οπτικές εισόδους και να αναγνωρίζουν τις χειρονομίες των ανθρώπων με μια ή περισσότερες κάμερες, χωρίς τη χρήση οποιονδήποτε επιπλέον συσκευών. Με άλλα λόγια, οι χρήστες δεν χρειάζεται να αγγίζουν τίποτα. Αυτό το σύστημα μπορεί να ξεπεράσει κάποιους περιορισμούς σε σχέση με το γάντι δεδομένων και ο χρήστης μπορεί να αισθάνεται εντελώς ελεύθερος με το περιβάλλον εργασίας αλληλεπίδρασης.

### Αναγνώριση Χειρονομιών μέσω Xbox

Παράδειγμα τέτοιας τεχνικής παρουσιάζεται στην εργασία των (Oskeoi and Huosheng, 2009), όπου οι συγγραφείς αναπτύσσουν νέα τεχνική διάδρασης ανθρώπου-υπολογιστή βασισμένη σε οπτικά σήματα για το Xbox. Πιο συγκεκριμένα, εφαρμόζεται ένας αλγόριθμος ανίχνευσης χαρακτηριστικών για να αναγνωρίζει

τις κινήσεις του κεφαλιού και των χεριών του χρήστη, και να τις μεταφραζει σε εντολές για το παιχνίδι. Για το λόγο αυτό χρησιμοποιείται η πυραμιδοειδής υλοποίηση της ανίχνευσης χαρακτηριστικών των Lucas Kanade, για να εντοπίσει τις οπτικές ροές πληροφορίας σε μια σειρά από πλαίσια εικόνων. Για το σχεδιασμό του συστήματος ελήφθη υπόψιν ότι θα πρέπει να λειτουργεί παράλληλα με τους ελεγκτές της κονσόλας του παιχνιδιού, και πιο συγκεκριμένα για τον έλεγχο ενός αυτοκινήτου μέσω του Xbox. Το σύστημα ανιχνεύει τις κινήσεις του κεφαλιού και των χεριών και παρακάμπτει κάποιες εντολές του χειριστηρίου του αυτοκινήτου, χρησιμοποιώντας τις εντολές από τις κινήσεις του χρήστη. Δηλαδή το σύστημα τρέχει παράλληλα με το κανονικό λογισμικό και το παρακάμπτει όπου χρειάζεται. Ένα τέτοιο παράδειγμα φαίνεται στην παρακάτω εικόνα, όπου διακρίνονται οι διάφορες κινήσεις του κεφαλιού του χρήστη (επιλέχθηκε το κεφάλι γιατί η αναγνώριση της κίνησης είναι πιο απαιτητική από αυτήν του χεριού) και η αντίστοιχη ερμηνεία του από το σύστημα (μετατόπιση της θέσης της πράσινης κουκίδας).



Εικόνα 28. Υλικό που απαιτείται για το σύστημα.



Εικόνα 29. Εντοπισμός και ανίχνευση των χαρακτηριστικών του προσώπου και εκτέλεση της αντίστοιχης εντολής με βάση τη χειρονομία του κεφαλιού (το ίδιο συμβαίνει με το χέρι). Η πράσινη κουκίδα δείχνει την μετατόπιση, η οποία αντιστοιχεί και στην αντίστοιχη εκτέλεση εντολής.

Η ανάπτυξη του συστήματος έγινε με Java σε κανονικό υπολογιστή, ενώ χρησιμοποιήθηκαν βιβλιοθήκες από τη C++, όπως η OpenCV. Επίσης χρησιμοποιήθηκε το Xbox Java SDK για να μπορέσει να γίνει ο έλεγχος των ελεγκτών της παιχνιδομηχανής. Τα πειράματα έγιναν με τρεις χρήστες ηλικίας 15, 21 και 25

ετών. Η διαδικασία ήταν η εξής. Κάθισαν ο καθένας ξεχωριστά μπροστά στην κάμερα και σε μια οθόνη που ήταν συνδεδεμένη με το Xbox, και προσπάθησαν να ελέγξουν την κατεύθυνση του αυτοκινήτου με τις κινήσεις του κεφαλιού και των χεριών τους. Για το πείραμα δεν χρησιμοποιήθηκε κάποιου είδους εκπαίδευση για το σύστημα, και τα αποτελέσματα έδειξαν ότι με βάση αυτό το σύστημα, η βαθμολόγηση για το σωστό έλεγχο του αυτοκινήτου ήταν σχεδόν η μισή από αυτήν που είχαν οι χρήστες όταν χρησιμοποιούσαν τα χειριστήρια του Xbox. Στην αναφορά τους παρουσιάζεται επίσης και στατιστική ανάλυση των αποτελεσμάτων των χρηστών, όπου φαίνεται ότι η χρήση του οπτικού συστήματος αναγνώρισης χειρονομιών προσθέτει ένα επιπλέον φορτίο 50% παραπάνω απ'ότι με τα χειριστήρια του Xbox.

#### *Πειραματική πλατφόρμα αναγνώρισης χειρονομιών για πολεμικά παιχνίδια*

Σε μια άλλη μέθοδο των (Teixeira & et. al., 2006) παρουσιάζεται ένα τρισδιάστατο παιχνίδι γραφικών με θέμα τις πολεμικές τέχνες, το οποίο υποστηρίζει τη διάδραση μέσω χειρονομιών. Στην πραγματικότητα, αυτή η εφαρμογή υλοποιήθηκε για να δοκιμάσει και να αναλύσει τη διάδραση με χειρονομίες στα πλαίσια ενός παιχνιδιού που χρειάζεται γρήγορους, σύντομους και αξιόπιστους χρόνους απόκρισης. Στην εργασία αναλύονται διάφοροι παράγοντες που επηρεάζουν την εφαρμογή, όπως οι επιπτώσεις της διάδρασης με χειρονομίες που προκαλείται στο ρυθμό πλαισίων, στον χρόνο απόκρισης ελέγχου των κινήσεων που κάνουν οι χρήστες, κ.α. Σε αυτήν την εφαρμογή δοκιμάζονται κυρίως μη συμβατικοί μέθοδοι διάδρασης με χειρονομίες, οι οποίες δεν θα πρέπει να βασίζονται σε τυποποιημένες κινήσεις πάλης, όπως αυτές στις οποίες βασίζονται οι εμπορικές εφαρμογές και τα διάφορα παιχνίδια. Οι χαρακτήρες που χρησιμοποιήθηκαν ήταν χορευτές (και όχι πυγμάχοι) μέσα σε ένα εικονικό περιβάλλον. Το παιχνίδι αναγκάζει τους χρήστες να κάνουν διάφορες κινήσεις, από χορό μέχρι πάλη, για να μπερδέψει αλλά και να αξιολογήσει αυτό το σύστημα διάδρασης.



Εικόνα 30. Το περιβάλλον αλληλεπίδρασης του παιχνιδιού (αριστερά) και η Αλληλεπίδρασης Ανθρώπου – Υπολογιστή του χρήστη με αυτό μέσω των αντίστοιχων χειρονομιών (δεξιά).

Για την υλοποίηση του συστήματος δημιουργήθηκε μια πλατφόρμα, η CIDA, η οποία χωρίζεται σε δύο οντότητες και μπορεί να εκτελεστεί μέχρι και σε τρεις υπολογιστές. Για την υλοποίηση χρησιμοποιήθηκαν C++, OpenGL και Direct3d, μαζί με διάφορα plugins. Για την υλοποίηση της διεπαφής αλληλεπίδρασης του χρήστη με το σύστημα μέσω χειρονομιών χρησιμοποιήθηκε το ARToolKit, με το οποίο ορίστηκαν αρχικά δύο δείκτες (G και F, όπως φαίνονται και στην παραπάνω εικόνα) με τους οποίους ανιχνεύονται οι κινήσεις του χρήστη, σαν να γίνονταν από ένα joystick. Στο περιβάλλον αλληλεπίδρασης υπάρχουν επίσης έξι κουμπιά, τα οποία ο χρήστης μπορεί να ενεργοποιήσει χρησιμοποιώντας και πάλι τις αντίστοιχες χειρονομίες με τους δείκτες. Για το υλικό χρησιμοποιήθηκε ένας απλός επεξεργαστής P4 στα 3.0GHz και μια απλή κάρτα γραφικών του εμπορίου.

Τα πειραματικά αποτελέσματα μετά τις δοκιμές ήταν ικανοποιητικά, και έδειξαν ότι εκτός από το ότι ο ρυθμός επεξεργασίας των πλαισίων εικόνων ανά δευτερόλεπτο ήταν αρκετά καλός (εκτός από την περίπτωση με τη μεγαλύτερη δυνατή ανάλυση), στις υπόλοιπες περιπτώσεις ο χρόνος απόκρισης ήταν εκτός των προσδοκιών και προκαλούσε μειωμένη εμπειρία διαδραστικότητας με το χρήστη σε σχέση με τη χρηστικότητα και την εμπειρία πραγματικού χρόνου που θα έπρεπε να υπάρχει στη συγκεκριμένη εφαρμογή. Συμπερασματικά αναφέρεται ότι το σύστημα έχει πολλές προοπτικές, και γι'αυτό και οι συγγραφείς αναφέρονται και στα επόμενα βήματα που θα κάνουν προς αυτήν την κατεύθυνση.

## *Σύστημα αλληλεπίδρασης για ιατρικά δεδομένα*

Μια διαφορετική εργασία παρουσιάζεται από τους (Van Beurden and IJsselsteijn, 2010) στην οποία περιγράφεται μια μελέτη για το εύρος και την διαφορετικότητα των φυσικών χειρονομιών όταν υπάρχει διαδραστική Αλληλεπίδρασης Ανθρώπου – Υπολογιστή με ιατρικές εικόνες, χρησιμοποιώντας παραδοσιακά μη στερεοσκοπικά αλλά και στερεοσκοπικά μοντέλα παρουσίασης. Αυτή η μελέτη στοχεύει στο σχεδιασμό αλγόριθμων υπολογιστικής όρασης για να βοηθήσει την φυσική διαδραστικότητα με χειρονομίες σε ιατρικό περιβάλλον. Πιο συγκεκριμένα, μελετάει το εύρος και την ατομική μεταβλητότητα των χειρονομιών που χρησιμοποιούνται όταν επιτρέπονται μη περιορισμένες αλληλεπιδράσεις (για παράδειγμα όταν ο χρήστης προτρέπεται να κάνει ελεύθερες και φυσικές κινήσεις για να αλληλεπιδράσει με το σύστημα, και όχι τυποποιημένες). Η μέθοδος που χρησιμοποιήθηκε προσπαθεί να χειριστεί αυτά τα μοντέλα παρουσίασης (στερεοσκοπικό και μη) σε ένα πειραματικό σχεδιασμό μεταξύ αντικειμένων. Σε κάθε μοντέλο εφαρμόζονται επτά διαφορετικές λειτουργίες (τοποθέτηση, επιλογή, ενεργοποίηση, περιστροφή, σύμκρυνση, μεγέθυνση και απενεργοποίηση) και σκοπός είναι οι συμμετέχοντες στο πείραμα (24 άτομα μεταξύ 20 και 34 ετών) να εφαρμόσουν τις παραπάνω λειτουργίες σε ένα υποθετικό σενάριο, ώστε να εξεταστεί η επίδραση της στερεοσκοπικής αντίληψης στη διάδραση με χειρονομίες. Στους συμμετέχοντες προβάλλονταν τέσσερις εικόνες και αυτοί έπρεπε να τις μετακινήσουν προς συγκεκριμένες κατευθύνσεις, με βάση κάποια διαδικασία. Οι εικόνες που προβάλλονταν ήταν είτε στερεοσκοπικές, είτε μονοσκοπικές.

Για την ανάλυση των αποτελεσμάτων χρησιμοποιήθηκε ειδικό σύστημα ανάλυσης βίντεο. Αρχικά έγινε η κατηγοριοποίηση των αποτελεσμάτων, εικόνων και χειρονομιών των συμμετεχόντων (για παράδειγμα η χειρονομία του «διπλό κλικ» τυποιοιήθηκε και κατηγοριοποιήθηκε με βάση το γεγονός ότι συνήθως γίνεται με ένα δάκτυλο που κάνει δύο κινήσεις, κ.ο.κ.), και τα αποτελέσματα έδειξαν ότι για την τοποθέτηση και την επιλογή όμοιων χειρονομιών, χρησιμοποιήθηκαν το ίδιο οι στερεοσκοπικές και οι μονοσκοπικές απεικονίσεις. Όμως, για τις υπόλοιπες λειτουργίες παρατηρήθηκε μεγαλύτερη διαφοροποίηση μεταξύ των χειρονομιών, από το οποίο προκύπτει ότι για κάποιες διαδράσεις με χειρονομίες (πχ. τοποθέτηση και επιλογή), οι χειρονομίες είναι σχεδόν πάντα όμοιες μεταξύ των ατόμων που χρησιμοποιούν το σύστημα, ενώ σε κάποιες άλλες λειτουργίες διαφοροποιούνται αρκετά μεταξύ των χρηστών. Από τις παραπάνω παρατηρήσεις προκύπτει ότι η υλοποίηση ενός γενικού συστήματος διάδρασης με χειρονομίες για αυτού του είδους τις λειτουργίες, δε θα μπορούσε να εφαρμοστεί εύκολα σε μεγάλο εύρος χρηστών, λόγω των μεγάλων διαφοροποιήσεων που υπάρχουν μεταξύ του τρόπου εκτέλεσης των λειτουργιών.

## *Αλληλεπίδραση ανθρώπου-υπολογιστή και υπολογιστική όραση*

Μια διαφορετική δημοσίευση παρουσιάζεται στο (Hassanpour and Shahbahrani, 2010), όπου δίνεται μια αναλυτική περιγραφή του τομέα της υπολογιστικής όρασης σε σχέση με την Αλληλεπίδρασης Ανθρώπου – Υπολογιστή της με τις χειρονομίες σε εφαρμογές διάδρασης ανθρώπου-υπολογιστή. Πιο συγκεκριμένα, περιγράφονται διάφορες μέθοδοι στον τομέα αυτόν και γίνεται μια ταξινόμηση τους, με βάση την εφαρμογή για την οποία αναπτύχθηκαν. Οι μέθοδοι που παρουσιάζονται είναι τρεις, μια δισδιάστατη χωρίς να χρησιμοποιούνται συγκεκριμένα σχηματικά μοντέλα χειρονομιών και κινήσεων, μια δεύτερη μέθοδος δισδιάστατη με προκαθορισμένα σχηματικά μοντέλα χειρονομιών, και τέλος μια τρίτη μέθοδος τρισδιάστατη με χρήση καμερών και στερεοσκοπικών εικόνων.

Επίσης, γίνεται μια εκτενής αναφορά στα διάφορα μοντέλα χειρονομιών που έχουν χρησιμοποιηθεί κατά καιρούς στα συστήματα αλληλεπίδρασης ανθρώπων-υπολογιστή, ενώ γίνεται και αναφορά στις διάφορες φυσικές ή τεχνητές κινήσεις (και τα χαρακτηριστικά τους) που θα πρέπει να εκτελούνται από τους χρήστες για την ακριβή αναγνώριση των χειρονομιών. Η ταξινόμηση που γίνεται στις χειρονομίες, με βάση αυτήν τη μελέτη, βασίζεται σε τρεις κατηγορίες, τις χειρονομίες επικοινωνίας, τις χειρονομίες χειρισμού και τις χειρονομίες ελέγχου. Οι πρώτες χρησιμοποιούνται για να εκφράσουν κυρίως μια ιδέα, οι δεύτερες για την Αλληλεπίδρασης Ανθρώπου – Υπολογιστή του ανθρώπου με κάποια αντικείμενα σε ένα περιβάλλον εργασίας, και οι τρίτες αποτελούν το γκρουπ των χειρονομιών που χρησιμοποιούνται για να ελέγξουν ένα σύστημα ή να καταδείξουν ένα σημείο ή αντικείμενο. Με βάση τα παραπάνω, μεγάλη σημασία παίζει το μοντέλο του χεριού που θα χρησιμοποιηθεί, γι' αυτό και γίνεται αναφορά σε ξεχωριστή ενότητα για το πως θα πρέπει να ερμηνεύεται η εκάστοτε χειρονομία από το σύστημα, με βάση τη θέση που θα έχει η παλάμη και το χέρι κατά τη διάρκεια της κίνησης. Γι' αυτό και παρουσιάζεται μια ομαδοποίηση των διαφόρων μεθόδων αναγνώρισης χειρονομιών, οι οποίες διακρίνονται στα μοντέλα χεριών με χαρακτηριστικά υψηλού

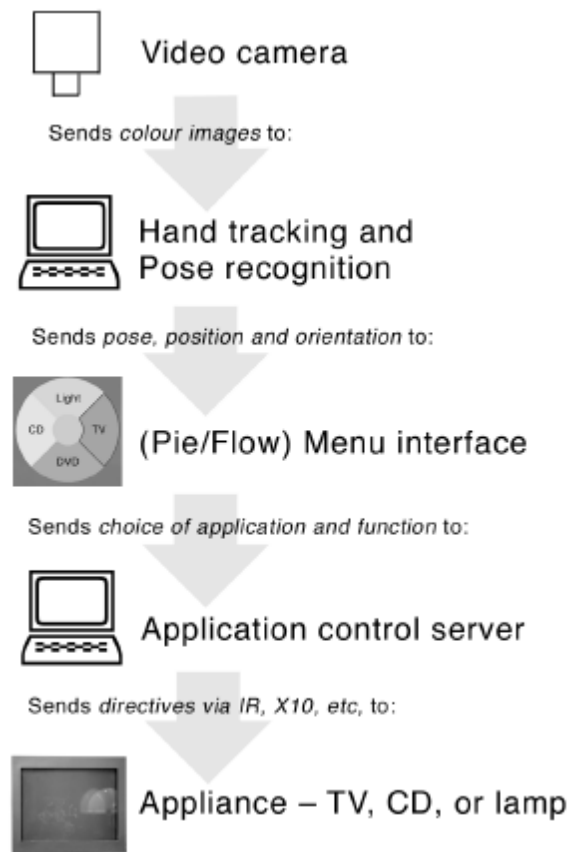
επιπέδου (πχ. για μεθόδους που χρησιμοποιούν τρισδιάστατα μοντέλα), στα οπτικά μοντέλα χεριών που βασίζονται στο συνδυασμό πληροφοριών από δισδιάστατες εικόνες του χεριού, και στα μοντέλα χεριών με χαμηλού επιπέδου χαρακτηριστικά, που βασίζονται στη χρήση ειδικού εξοπλισμού (πχ. γάντια) για την αναγνώριση του χεριού και επομένως δεν είναι αναγκαία η λήψη λεπτομερών πληροφοριών και μεγάλου όγκου δεδομένων από την ανάλυση των εικόνων των χεριών για να γίνει η αναγνώριση τους.

Στα συμπεράσματα αυτής της μελέτης αναφέρεται ότι ακόμα υπάρχουν αρκετοί περιορισμοί στις περισσότερες μεθόδους, ώστε να κάνουμε την υπόθεση ότι μπορούν αυτές να ενσωματωθούν στην καθημερινότητα μας, ενώ κάποιοι περιορισμοί, όπως η συνύπαρξη πολλών χεριών σε ένα σύστημα αλληλεπίδρασης ή η επικάλυψη (occlusion) μέρους του χεριού από κάποιο άλλο αντικείμενο, εμποδίζει τις περισσότερες μεθόδους στο να αναγνωρίσουν με ακρίβεια την εκάστοτε χειρονομία. Επίσης αναφέρει ότι η τάση στις μέρες μας είναι η χρήση καμερών για τα συστήματα αλληλεπίδρασης, ενώ για την αποφυγή των παραπάνω περιορισμών θα πρέπει να γίνει εκμετάλλευση των εικόνων δύο ή περισσότερων καμερών και η δημιουργία τρισδιάστατων μοντέλων ώστε να ελαχιστοποιηθούν αυτά τα προβλήματα.

### *Σύστημα ιεραρχικής σηματοδότησης μενού εντολών*

Στην μέθοδο των (Lenman, Bretzner and Thuresson, 2002) δίνεται μια περιγραφή των υπάρχουσών μεθόδων στον τομέα της διάδρασης ανθρώπου-υπολογιστή, με έμφαση στις τεχνολογίες που βασίζονται στις χειρονομίες χεριών. Ο τελικός σκοπός τους είναι να αναπτύξουν ένα σύστημα για να δοκιμάσουν το σενάριο όπου η ιεραρχική σηματοδότηση των μενού (hierarchical marking menus) μπορεί να χρησιμοποιηθεί για την ανάπτυξη κάποιου συνόλου εντολών με βάση τις χειρονομίες. Το βασικό σενάριο που δοκιμάζουν είναι ο απομακρυσμένος έλεγχος κάποιων ηλεκτρικών συσκευών που υπάρχουν στο σπίτι (όπως τηλεόραση, βίντεο, κτλ), ενώ σκοπεύουν να επεκτείνουν το σενάριο και σε εξωτερικά περιβάλλοντα με απρόβλεπτους παράγοντες. Για παράδειγμα, για τον έλεγχο του μενού της τηλεόρασης, αρχικά στην Αλληλεπίδρασης Ανθρώπου – Υπολογιστή του χρήστη με το μενού θα υπάρχει κάποιου είδους μήνυμα ανάδρασης από το σύστημα, ώστε να ξέρει ο χρήστης ότι επέλεξε και εκτέλεσε σωστά την εντολή, ενώ σκοπός είναι ο χρήστης να φτάσει στο σημείο να εκτελεί εντολές για τις διάφορες συσκευές, γνωρίζοντας εκ των προτέρων το αποτέλεσμα που θα έχουν.

Το σύστημα τους βασίζεται σε δύο διαφορετικά είδη μενού εντολών, από τα οποία μπορούν να αναγνωριστούν διάφορες εντολές με βάση τις χειρονομίες των χεριών. Το σύστημα όρασης που χρησιμοποιήσαν υποστηρίζει χρωματικές και σχηματικές διαφορές, ενώ για την εύρεση της θέσης των χεριών χρησιμοποιεί συνδυασμό ανίχνευσης πολυεπίπεδων χρωματικών χαρακτηριστικών, μαζί με ποιοτικές εσωσχεσιακές διαφοροποιήσεις της κλίμακας, της θέσης και της κατεύθυνσης. Με βάση το μοντέλο που σχηματίζεται από τα προηγούμενα, δημιουργείται ένα σχήμα της θέσης του χεριού, και σε κάθε εικόνα αναγνωρίζονται τα πολυεπίπεδα χαρακτηριστικά του. Στη συνέχεια αρχίζει ο εντοπισμός και η παρακολούθηση του χεριού, ενώ εφαρμόζοντας και μια παράλληλη διαδικασία φιλτραρίσματος, είναι δυνατός και ο προσδιορισμός της θέσης του χεριού, της κατάστασης του, του προσανατολισμού του και της κλίμακας του. Για το γρηγορότερο εντοπισμό του χεριού χρησιμοποιείται ανίχνευση του χρώματος του δέρματος, ενώ τα αντικείμενα που δεν αντιστοιχούν σε χρώματα παραπλήσια του δέρματος του χεριού φιλτράρονται και απομωώνονται. Παρακάτω φαίνεται και το σχηματικό διάγραμμα της λειτουργίας αυτού του συστήματος.



Εικόνα 31. Σχηματικό διάγραμμα της λειτουργίας του συστήματος.

Ο εξοπλισμός του πειραματικού περιβάλλοντος που χρησιμοποιήθηκε περιλαμβάνει ένα workstation 530 της DELL σε περιβάλλον Linux, μια οθόνη 19'', ένα framegrabber (καταγραφέα στιγμιοτύπων), διάφορους αισθητήρες και κάποιες συσκευές για τις οποίες θα εκτελείται ο έλεγχος και η Αλληλεπίδρασης Ανθρώπου – Υπολογιστή με το χρήστη (πχ. τηλεόραση). Όσον αφορά τη λειτουργία του μενού, επειδή το σύστημα είναι ακόμα σε στάδιο ανάπτυξης, έχει υλοποιηθεί περιορισμένο πλήθος επιλογών για την Αλληλεπίδρασης Ανθρώπου – Υπολογιστή με τις συσκευές, όπως το άνοιγμα και το κλείσιμο της τηλεόρασης, οι επιλογές επόμενο, προηγούμενο, κ.α.

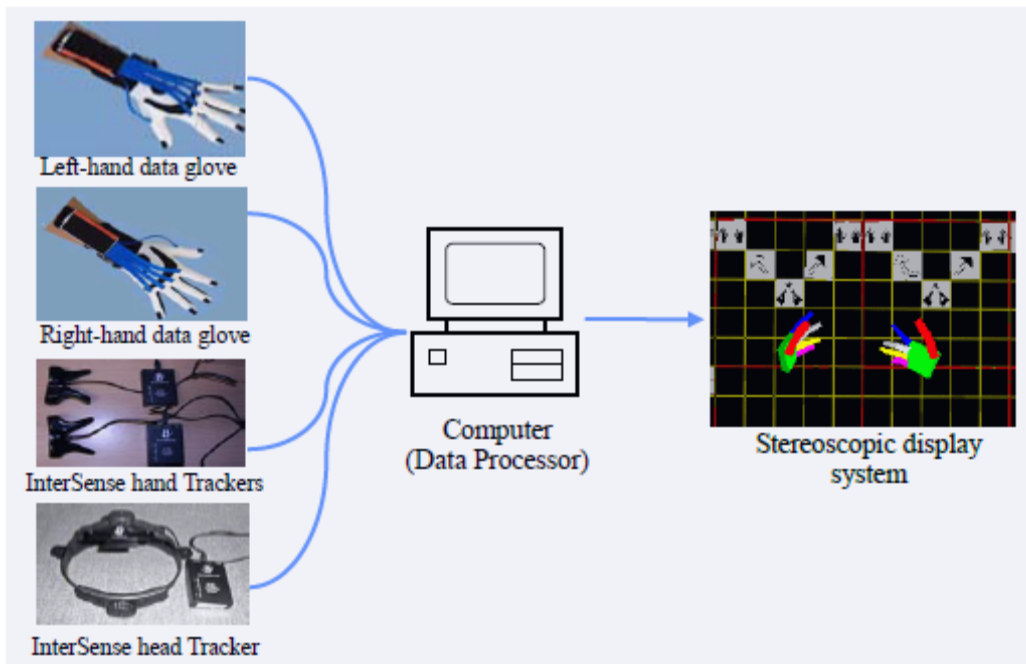
Από τα πειραματικά αποτελέσματα προκύπτει οι χρήστες χρειάζεται να εκτελέσουν αρχικά κάποιες εντολές σε επίπεδο εκπαίδευσης για να μπορέσουν να συνηθίσουν τη λειτουργία του. Επίσης, είναι αρκετά πιο σύνθετο σε σχέση με το προηγούμενο μοντέλο που είχαν αναπτύξει, ενώ η χρήση των μενού παρουσιάζει κάποιες δυσκολίες και ο χρήστης θα πρέπει να ασχοληθεί και να μάθει τις συγκεκριμένες λειτουργίες. Επίσης, οι χρήστες θα πρέπει να προσέχουν ώστε να μην κάνουν συγκεκριμένες κινήσεις με τα χέρια τους, λόγω των περιορισμών του συστήματος στην αναγνώριση των χειρονομιών (πχ. το σύστημα έχει περιορισμένο αριθμό χρήσιμων χειρονομιών που αντιστοιχούν σε εντολές). Συμπερασματικά, ισχυρίζονται ότι το σύστημα τους δεν είναι κατάλληλο ακόμα για ευρεία χρήση καθώς μένουν ακόμα πολλά που πρέπει να υλοποιηθούν (πχ. καλύτερη συσχέτιση των χειρονομιών με τις αντίστοιχες λειτουργίες, ή ο περιορισμός και η αποφυγή συγκεκριμένων χειρονομιών που σχετίζονται με κοινωνικές ή φυλετικές ομάδες), αλλά έχει πολλές προοπτικές εξέλιξης, καθώς το πλήθος των συσκευών που θα μπορούσε να εφαρμοστεί είναι πολύ μεγάλο.

*Υβριδικό σύστημα με γάντια οπτικής ίνας για αναγνώριση κινήσεων δακτύλων*

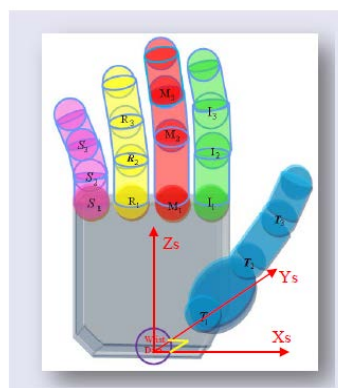
Τέλος, στη μέθοδο του (Gan, 2011) περιγράφεται ο σχεδιασμός και η ανάπτυξη ενός καινοτόμου συστήματος διάδρασης ανθρώπου-υπολογιστή που συνδυάζει με μοναδικό τρόπο ένα ζευγάρι γάντια συνδεδεμένα με αισθητήρες οπτικής ίνας για να αναγνωρίζουν κινήσεις των δακτύλων. Το σύστημα είναι υβριδικό, και βασίζεται στην αδράνεια και στους υπέρηχους για να συλλέξει τις θέσεις των χεριών και τον

προσανατολισμό τους. Επίσης συνδυάζει ένα σύστημα στερεοσκοπικής οθόνης για να παράσχει το οπτικό αποτέλεσμα της διάδρασης.

Πιο συγκεκριμένα, το προτεινόμενο σύστημα ελέγχεται από έναν κανονικό υπολογιστή ο οποίος συνδέεται με τις προαναφερόμενες συσκευές (βλ. παρακάτω εικόνα) για τη λήψη των δεδομένων της θέσης του χεριού, του προσανατολισμού του και της μορφής του. Για να γίνει αυτό, χρησιμοποιήθηκαν διαφορετικά σημεία εντοπισμού, μέσω ενός κινηματικού μοντέλου που υλοποιήθηκε, πάνω στο γάντι (όπως φαίνεται παρακάτω), από τα οποία το σύστημα μπορεί να εξάγει ακριβείς πληροφορίες για τα παραπάνω δεδομένα. Επίσης, χρησιμοποιήθηκε και ένας ανιχνευτής που τοποθετείται στο κεφάλι και χρησιμοποιήθηκε για τη λήψη πληροφοριών σχετικά με τη θέση και τον προσανατολισμό του κεφαλιού του χρήστη. Με βάση τη θέση του κεφαλιού, το σύστημα μπορεί και προσαρμόζει την προβολή των αντικειμένων πάνω στη στερεοσκοπική οθόνη, προβάλλοντας τα αντικείμενα από την κατάλληλη οπτική γωνία με βάση την κλίση του κεφαλιού.



Εικόνα 32. Προτεινόμενο σύστημα.

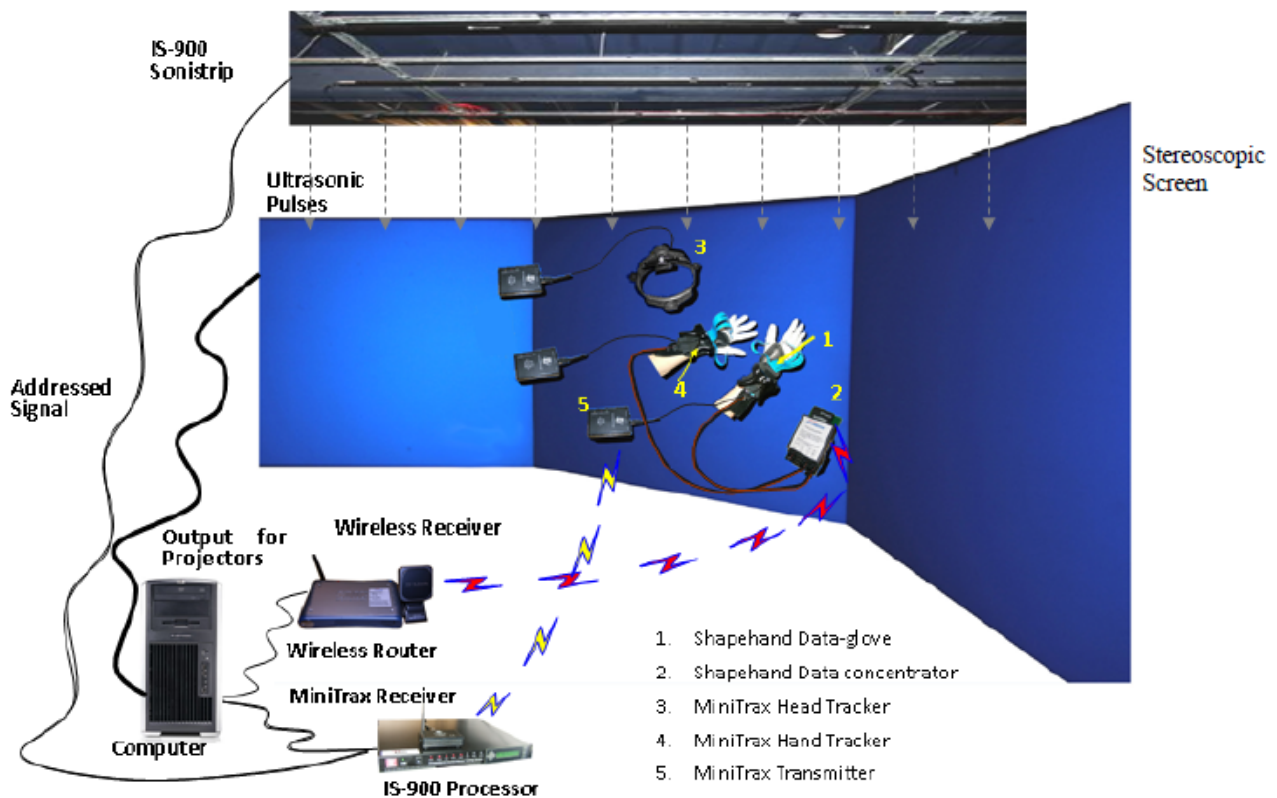


Εικόνα 33. Το κινηματικό μοντέλο του χεριού, όπως ερμηνεύεται από το σύστημα μέσω του γαντιού.

Το υλικό που χρησιμοποιήθηκε (και φαίνεται παρακάτω) αποτελείται από έναν υπολογιστή γραφείου και τρία υποσυστήματα, το σύστημα του γαντιού που χρησιμοποιείται για τη λήψη των δεδομένων της χειρονομίας, από το σύστημα εντοπισμού της θέσης του κεφαλιού και του χεριού, και τέλος μια στερεοσκοπική οθόνη. Χρησιμοποιήθηκαν επίσης μια συσκευή επεξεργασίας των δεδομένων που μεταδίδονται με ασύρματο τρόπο, ένα μόντεμ/ρούτερ και κάποιοι υπέρηχοι αισθητήρες. Το λογισμικό τρέχει σε περιβάλλον Windows και υλοποιήθηκε σε C++ χρησιμοποιώντας κλάσεις από το MFC (Microsoft



Foundation Classes) για τη δημιουργία του περιβάλλοντος αλληλεπίδρασης και ελέγχου. Επίσης χρησιμοποιήθηκαν βιβλιοθήκες από την OpenGL για τη γραφική απεικόνιση των αντικειμένων.



Εικόνα 34. Σχηματικό διάγραμμα του υλικού του συστήματος.

Στα πειράματα που έγιναν για το παραπάνω σύστημα φαίνεται η αποτελεσματικότητα της μεθόδου σε διάφορες περιπτώσεις, όπως γράψιμο, ομιλία με ταυτόχρονες κινήσεις χεριών, κτλ. Το προτεινόμενο σύστημα επιτρέπει το χρήστη να επιλέξει, να μεταφράσει, να περιστρέψει, να μεγενθύνει και να προβάλει εικονικά αντικείμενα στις τρεις διαστάσεις, χρησιμοποιώντας φυσικές κινήσεις των χεριών. Το σύστημα γαντιών που χρησιμοποιήθηκε ήταν το ShapeHand μαζί με τρεις InterSense MiniTrax συσκευές ανίχνευσης κινήσεων, μία για το κεφάλι, και δύο για τα χέρια, μαζί με μία μεγάλη στερεοσκοπική οθόνη. Δοκιμάστηκαν διάφορες κινήσεις των δακτύλων και των χεριών (πχ. λύγισμα δακτύλων, προέκταση, σφίξιμο, κτλ) και το σύστημα μπορούσε να αναγνωρίσει τις κινήσεις αυτές άμεσα. Για τις δοκιμές χρησιμοποιήθηκε επίσης και μια κάμερα υψηλής ταχύτητας, για να καταγράψει τις καθυστερήσεις στην απόκριση του συστήματος για το mode της δυναμικής αναγνώρισης των κινήσεων των χεριών (μεγ. 64 fps). Επίσης, η απόκριση του συστήματος ήταν πάρα πολύ ικανοποιητική, ενώ η εκπαίδευση που απαιτείται για να γίνει σωστή χρήση του συστήματος είναι πολύ μικρή.

## Γ. Συγκριτική Αξιολόγηση Τεχνικών και Συμπεράσματα

Στις προηγούμενες ενότητες έγινε εκτενής αναφορά σε σύγχρονες μεθόδους Αλληλεπίδρασης Ανθρώπου – Υπολογιστή και περιγράφηκε ο τρόπος υλοποίησής τους όσον αφορά τον αλγόριθμο ή την τεχνική που ακολουθήθηκε, καθώς και το υλικό ή λογισμικό που χρησιμοποιήθηκε στην εκάστοτε περίπτωση.

Όπως προκύπτει από τα παραπάνω, οι μέθοδοι που χρησιμοποιήθηκαν διαφοροποιούνται αρκετά όσον αφορά την μεθοδολογία που ακολούθησαν, αλλά και όσον αφορά την πειραματική διαδικασία και τα αποτελέσματά τους. Κάποιες μέθοδοι, όπως αυτές που εφαρμόζονται στα συστήματα της Microsoft (Kinect) και του Wii της Nintendo, έχουν ήδη υλοποιηθεί και τυποποιηθεί σε εμπορικά προϊόντα με μεγάλη απήχηση στο κοινό λόγω της καλής απόκρισης που έχουν και της ελευθερίας κινήσεων/χειρονομιών που προσφέρουν στο χρήστη. Και οι δύο συσκευές παρουσιάζουν καλή και γρήγορη απόκριση σε σχέση με τις χειρονομίες του χρήστη, αλλά το Kinect σαφώς επιτρέπει μεγαλύτερη ευελιξία στις κινήσεις του, καθώς εισάγει λιγότερους περιορισμούς όσον αφορά το δραστικό πεδίο αναγνώρισης των χειρονομιών από το σύστημα, λόγω της κάμερας που διαθέτει. Επίσης, η διεπαφή που χρησιμοποιείται από το κάθε σύστημα είναι διαφορετική (στο ένα ασύρματη μέσω κάμερας ενώ στο άλλο μέσω χειριστηρίων), κάτι που επηρεάζει την ελευθερία κινήσεων των χρηστών και το πεδίο εφαρμογών της κάθε συσκευής (για παράδειγμα, με το Wii ίσως να είναι πιο δύσκολο σε απαιτητικές εφαρμογές πολλών χρηστών να αλληλεπιδράσουν ταυτόχρονα 2 ή περισσότεροι χρήστες με το σύστημα). Όσον αφορά το σύστημα και τα τεχνικά χαρακτηριστικά τους, το Kinect ενσωματώνει μια κάμερα με αισθητήρα βάθους και μπορεί να ανιχνεύσει τις κινήσεις χωρίς τη χρήση φυσικού ελεγκτή (χειριστηρίου). Σε αντίθεση, το Wii έχει όλα τα κουμπιά ελέγχου στο χειριστήριο, ενώ με έναν αισθητήρα IR ανιχνεύεται η θέση του στο χώρο. Στο Kinect μπορείς να περιηγηθείς στα μενού και να παίζεις παιχνίδια χωρίς να ακουμπήσεις καμία συσκευή (πχ. χειριστήριο ή κονσόλα), ενώ υπάρχει και μικρόφωνο για την εισαγωγή φωνητικών εντολών. Στο Wii το Wiimote αρχικά δεν αποκρινόταν όπως έπρεπε και οι χρήστες παρατηρούσαν καθυστέρηση στις κινήσεις του στο παιχνίδι, όμως με την εισαγωγή του Wii MotionPlus αυτή η αίσθηση βελτιώθηκε. Μειονέκτημα του Kinect είναι ότι απαιτείται αρκετός χώρος για την αλληλεπίδραση του χρήστη με το σύστημα (συνιστάται ο χρήστης να βρίσκεται σε απόσταση μεγαλύτερη από 2-3 μέτρα μακριά από τη κάμερα), ενώ με κόστος στα 150\$ είναι αρκετά προσιτό. Επίσης, η δυνατότητα να αλληλεπιδράσουν με το σύστημα περισσότεροι παίκτες δίνοντας ταυτόχρονα φωνητικές εντολές στο σύστημα, το κάνει να ξεχωρίζει από το Wii, του οποίου η περιορισμένη δυνατότητα να υποστηρίξει γραφικά HD και το υψηλότερο κόστος του (199\$), το καθιστούν μη ελκυστικό πολλές φορές.

Άλλα συστήματα που έχουν τυποποιηθεί και μετατραπεί σε εμπορικά προϊόντα είναι τα συστήματα αλληλεπίδρασης που έχουν αναπτύξει οι μεγάλες εταιρείες του χώρου της τηλεόρασης. Συστήματα όπως το Smart Interaction και το Smart TV της Samsung, της Toshiba που παρουσιάζεται στο (Miller, 2008), της Lenovo το Smart TV, της Hitachi στο (Saenz, 2009), το Magic Motion της LG ή το Dragon TV της Nuance (που το έχει δώσει στην Panasonic για να ενσωματωθεί στις τηλεοράσεις της) είναι μόνο μερικά από τα συστήματα που αναπτύσσονται ή έχουν ήδη αναπτυχθεί για να αποτελέσουν την επόμενη γενιά τηλεοράσεων. Παρόλο που τα περισσότερα από αυτά περιέχουν τις ίδιες λειτουργίες (πχ. αναγνώριση φωνής ή αναγνώριση χειρονομιών), εντούτοις το καθένα από αυτά παρουσιάζει διαφορετική απόδοση και απόκριση στις χειρονομίες των χρηστών. Κάποια από αυτά (όπως αυτό της Samsung) παρουσιάζει περιορισμένη δυνατότητα χειρισμών και μεγάλη καθυστέρηση στην αναγνώριση φωνής, κάτι που το κάνει να υστερεί έναντι κάποιων άλλων, όπως αυτό της Nuance, το οποίο θεωρείται αρκετά καινοτόμο όσον αφορά τα επίπεδα και τις δυνατότητες αναγνώρισης της φωνής. Επίσης, η τιμή του κάθε συστήματος παίζει σαφώς ρόλο, αλλά για κάποια προϊόντα δεν έχει ανακοινωθεί ακόμα τιμή (όπως αυτό της Lenovo), επομένως η σύγκριση σε αυτό το επίπεδο δεν είναι δυνατή. Σημαντικό είναι επίσης το γεγονός ότι κάποια συστήματα, όπως της Samsung, της LG, της Hitachi και της Toshiba, έχουν ήδη μια μικρή ωριμότητα στο χώρο, αφού κυκλοφορούν ήδη εμπορικά προϊόντα τους με δυνατότητες αλληλεπίδρασης με τους ανθρώπους εδώ και λίγα χρόνια, σε αντίθεση με τα άλλα συστήματα, όπως αυτό της Panasonic ή της Lenovo, που μόλις πρόσφατα ξεκίνησαν την παραγωγή και διάθεση αυτών των προϊόντων στην αγορά. Μια άλλη διαφορά είναι το γεγονός ότι το σύστημα της Toshiba παρουσιάζει έναν διαφορετικό τρόπο αλληλεπίδρασης από τα άλλα

συστήματα, χρησιμοποιώντας και κατευθύνοντας έναν κέρσορα πάνω σε κάποιο μενού και όχι μέσω της αναγνώρισης κάποιων προκαθορισμένων μοτίβων χειρονομιών/κινήσεων από το οπτικό σύστημα αναγνώρισης της τηλεόρασης για την εκτέλεση της αντίστοιχης εντολής. Επίσης, το ίδιο σύστημα παρουσιάζει μια ακόμα καινοτομία έναντι των υπολοίπων δίνοντας τη δυνατότητα αναγνώρισης εικόνων και την εκτέλεση προκαθορισμένων εντολών/λειτουργιών ανάλογα με την εικόνα που αναγνωρίστηκε. Τέλος, ένας άλλος σημαντικός παράγοντας είναι η ευχρηστία και η απόκριση/ακρίβεια αναγνώρισης των χειρονομιών του συστήματος αλληλεπίδρασης που χρησιμοποιούν οι τηλεοράσεις, με τη Samsung και τη Toshiba να υπερέχουν σε αυτόν τον τομέα έναντι των υπολοίπων προϊόντων, καθώς το σύστημά τους είναι ικανό να αναγνωρίζει τις χειρονομίες με παρόμοια ακρίβεια με αυτή του Wii, του Xbox ή ακόμα και με αυτήν του ποντικιού.

Όσον αφορά τις άλλες μεθόδους που βρίσκονται ακόμα κυρίως σε ερευνητικό επίπεδο, υπάρχει μεγάλη διαφοροποίηση των τεχνικών, του υλικού, του λογισμικού, αλλά και των μεθόδων αξιολόγησης τους, με αποτέλεσμα μια ευθεία σύγκριση μεταξύ τους να μην είναι δυνατή και δίκαιη με βάση τα στοιχεία που παρουσιάζονται σε κάθε μία από αυτές. Παρ' όλ' αυτά, μπορούμε να συγκρίνουμε κάποια γενικά χαρακτηριστικά τους και να δώσουμε κάποιες κατευθύνσεις ως προς τη καταλληλότερη μέθοδο που θα μπορούσε να εφαρμοστεί ανά περίπτωση.

Το σύστημα που παρουσιάζεται στο (Lenman, Bretzner and Thuresson, 2002) θα μπορούσε να αποτελεί ένα μελλοντικό βήμα των συστημάτων που έχουν υλοποιηθεί ως τώρα στις τηλεοράσεις, καθώς μπορεί ο χρήστης να ελέγχει πλήθος οικιακών ηλεκτρικών συσκευών μόνο μέσω ενός συστήματος αναγνώρισης χειρονομιών, ανάδρασης και μέσω κάποιων μενού (που υπάρχουν ούτως ή άλλως στις τηλεοράσεις). Το μειονέκτημα της μεθόδου είναι ότι απαιτείται αρχική εκπαίδευση για το σύστημα, ενώ ο χρήστης θα πρέπει να ασχοληθεί και να μάθει τις λειτουργίες του μενού, κάτι που το κάνει δύσχρηστο σε σχέση με τα άλλα συστήματα που έχουν υλοποιηθεί μέχρι τώρα στις τηλεοράσεις. Επίσης, το περιορισμένο πλήθος επιλογών του μενού είναι ένας ακόμα αποτρεπτικός παράγοντας, ενώ όπως ισχυρίζονται και οι ίδιοι συγγραφείς, το σύστημα δεν είναι ακόμα έτοιμο για εμπορική χρήση καθώς απαιτούνται ακόμα αρκετές βελτιώσεις.

Όσον αφορά εφαρμογές με περιορισμένο πεδίο αναγνώρισης από το σύστημα ή με περιορισμένες δυνατότητες κίνησης από το χρήστη, θα μπορούσε να εφαρμοστεί πιο αποδοτικά η καινοτόμος μέθοδος που παρουσιάζεται στη δουλειά των Sato, Rouryev and Harrison (2012). Τα ικανοποιητικά αποτελέσματα της μεθόδου μετά τα πειράματα και η πολλά υποσχόμενη τεχνική που εφαρμόζει για την αναγνώριση των κινήσεων και την αλληλεπίδραση με το σύστημα, υπερτερεί σε σχέση με τις προηγούμενες μεθόδους σε αυτόν τον τομέα των εφαρμογών. Τέτοια παραδείγματα θα μπορούσαν να είναι εφαρμογές που ο χρήστης θα πρέπει να αλληλεπιδρά με το σύστημα υπό συνθήκες περιορισμένου φωτισμού, αψής, ή ακόμα και σε περιπτώσεις που ο χρήστης έχει κινητικά προβλήματα και μπορεί να εκτελεί μόνον περιορισμένες και μικρές κινήσεις των χεριών.

Αντίστοιχα με τις προηγούμενες τεχνικές, στις τεχνικές που χρησιμοποιούν σαν διαπαφή το γάντι, η μέθοδος που προτείνεται από το MIT στο (Wang and Popovic, 2009) είναι πιο εύχρηστη, φθηνή και αρκετά ακριβής σε σχέση με αυτήν που παρουσιάζεται στο (Garg, Aggarwal and Sofat, 2009) και στο (Gan, 2011), καθώς η υλοποίηση ενός αποδοτικού αλγόριθμου για την αναγνώριση ενός έγχρωμου γαντιού προσφέρει σαφώς μεγαλύτερη ευελιξία, ελευθερία κινήσεων και καλύτερη αλληλεπίδραση με το σύστημα, απ' ότι ένα γάντι δεδομένων με καλώδια που περιορίζει τις δυνατότητες του συστήματος και τις κινήσεις του χρήστη. Με βάση και τα πειράματα που έγιναν, αποδείχθηκε ότι οι χρήστες προτιμούν περισσότερο την οπτική αναγνώριση των κινήσεων τους από το σύστημα αλληλεπίδρασης, απ' ότι όταν το σύστημα χρησιμοποιεί γάντια δεδομένων για τη μετάδοση των πληροφοριών. Επίσης, στη δεύτερη περίπτωση, το κόστος του συστήματος είναι αρκετά μεγαλύτερο από αυτό της οπτικής αναγνώρισης των κινήσεων, και δεν απαιτείται ειδικό υλικό/εξοπλισμό, εξειδικευμένη παραμετροποίηση του γαντιού, κτλ. για να λειτουργήσει το σύστημα. Βέβαια, η ακρίβεια των κινήσεων δεν είναι η ίδια στα παραπάνω συστήματα και η επιλογή του καθενός εξαρτάται και από την εφαρμογή. Για παράδειγμα, το τελευταίο σύστημα που παρουσιάζεται στο (Gan, 2011) μπορεί να ανιχνεύει με ακρίβεια μικρές κινήσεις των δακτύλων, ακόμα και γραφή, με ελάχιστη εκπαίδευση του χρήστη, κάτι που το κάνει ιδανικό για τέτοιου είδους αλληλεπίδραση έναντι των υπολοίπων συστημάτων. Ακόμα και η στερεοσκοπική κάμερα στο κεφάλι βοηθάει στην καλύτερη αντίληψη του χώρου από το χρήστη, όσον αφορά την αλληλεπίδραση του με το σύστημα, κάτι που δεν υπάρχει στα άλλα συστήματα. Τέλος, η αποθήκευση μοτίβων χειρονομιών σε μια βάση δεδομένων πριν τη χρήση του συστήματος θα πρέπει να θεωρείτε δεδομένη σχεδόν σε κάθε τέτοιο σύστημα, ενώ δεν μπορεί να γίνει

απευθείας σύγκριση στην ταχύτητα και την ακρίβεια ανάκτησης αυτών των δεδομένων από τη βάση δεδομένων, λόγω των διαφορετικών μεθόδων που εφαρμόστηκαν και των διαφορετικών δεδομένων που αποθηκεύτηκαν.

Στην περίπτωση του (Keir et. al., 2006) το σύστημα είναι σαφώς πιο απλό και ευέλικτο από τα παραπάνω, καθώς δε γίνεται χρήση γαντιού, αλλά η διεπαφή του χρήστη με το σύστημα μπορεί να καθοριστεί/προγραμματιστεί απευθείας από το χρήστη (πχ. μπορεί να είναι το ποντίκι του H/Y, ένα κινητό τηλέφωνο, κτλ). Παραπλήσια μέθοδος είναι και αυτή που παρουσιάζεται στο (Leber, 2012), όπου και πάλι ο χρήστης μπορεί να προσαρμόσει τη συγκεκριμένη μέθοδο και σε άλλες μικρές και κινητές συσκευές (πχ. κινητό τηλέφωνο), με τη διαφορά ότι σε αυτήν την περίπτωση θα πρέπει να υπάρχει ενσωματωμένη κάμερα στη συσκευή για την αναγνώριση της χειρονομίας. Στην πρώτη περίπτωση, το κόστος του συστήματος είναι αρκετά πιο φθινό από τα προηγούμενα, ενώ η εκπαίδευση που απαιτείται από το χρήστη είναι μικρή, σε σχέση με τα πλεονεκτήματα που προσφέρονται μέσω της αξιοπιστίας και της ακρίβειας απόκρισης του συστήματος. Βέβαια οι τροποποιήσεις που θα πρέπει να γίνουν στο σύστημα για την κάθε περίπτωση αποτελούν μια πρόσθετη δυσκολία που δεν υπάρχει στα άλλα συστήματα, ενώ στο σύστημα του Leber (2012) υπάρχουν κάποιοι περιορισμοί όσον αφορά τις κινήσεις που μπορεί να κάνει το χέρι πάνω στην επιφάνεια αλληλεπίδρασης, το εύρος της αναγνωρίσιμης περιοχής και τον τρόπο αναγνώρισης των χειρονομιών. Το υλικό στο (Keir et. al., 2006) αλλά και στο (Leber, 2012) δεν είναι τόσο εξειδικευμένο όπως στις προηγούμενες περιπτώσεις (πχ. όπως στο (Wang and Popovic, 2009) και στο (Gan, 2011)), ενώ το λογισμικό βασίζεται σε ένα SDK και στη C++, όπως και στις περισσότερες μεθόδους, με χρήση και πάλι μια βάσης δεδομένων για αποθήκευση των μοτίβων των αναγνωρίσιμων χειρονομιών, αλλά με τη διαφορά ότι εδώ δεν αποθηκεύονται εικόνες (όπως στις προηγούμενες περιπτώσεις), αλλά δεδομένα σε μορφή XML, που σημαίνει ότι η επεξεργασία και ανάκτηση τους είναι πολύ πιο γρήγορη και απλή έναντι των υπολοίπων (πχ. δεν χρειάζονται εξειδικευμένοι και υπολογιστικά δαπανηροί αλγόριθμοι για την εύρεση της αντίστοιχης χειρονομίας μέσα από τη βάση δεδομένων).

Μια παραλλαγή αυτών των μεθόδων, που παρουσιάζεται στο (Kim and Fellner, 2004), δεν χρησιμοποιεί ακριβώς γάντια, αλλά κάποιους markers για τον εντοπισμό των χεριών και των χειρονομιών. Αυτή η μέθοδος είναι ακόμα πιο απλή στην υλοποίησή της, αλλά τα αποτελέσματα που εξάγει έχουν κάποιους περιορισμούς ως προς τα άτομα/χέρια που μπορούν να ανιχνευτούν με ακρίβεια, καθώς και στις υπάρχουσες περιβαλλοντικές συνθήκες που θα πρέπει να επικρατούν κατά τη διάρκεια της αλληλεπίδρασης του χρήστη με το σύστημα (πχ. θα πρέπει να υπάρχουν συγκεκριμένες συνθήκες φωτισμού και ο χρήστης δεν θα πρέπει να φοράει συγκεκριμένου χρώματος ρούχα). Κάποιοι είδους markers χρησιμοποιεί και η μέθοδος στο (Teixeira & et. al., 2006), όπου δοκιμάζονται κυρίως μη συμβατικοί μέθοδοι διάδρασης με χειρονομίες, οι οποίες δεν θα πρέπει να βασίζονται σε τυποποιημένες κινήσεις. Βέβαια και σε αυτό το σύστημα υπάρχουν περιορισμοί, με σοβαρότερο ότι το σύστημα στις περισσότερες των περιπτώσεων είχε αργό χρόνο απόκρισης, εκτός των προσδοκιών, και προκαλούσε μειωμένη εμπειρία διαδραστικότητας με το χρήστη. Παρ' όλ' αυτά, και τα δύο συστήματα είναι αρκετά απλά στη χρήση τους, το λογισμικό που χρησιμοποιήθηκε ήταν βασισμένο κυρίως στη C++, ενώ από υλικό δεν είχαν μεγάλες απαιτήσεις. Επίσης, με βάση τα πειράματα που έγιναν και στις δύο μεθόδους, το (Kim and Fellner, 2004) θα ήταν πιο χρήσιμο σε συγκεκριμένες εφαρμογές που άλλα παρόμοια συστήματα θα υστερούσαν (όπως αυτό στο (Teixeira & et. al., 2006)), όπως για παράδειγμα σε εφαρμογές που απαιτείται αλληλεπίδραση του χρήστη με το σύστημα σε σκοτεινές συνθήκες φωτισμού και χωρίς άμεση επαφή με το σύστημα.

Εναλλακτικές μέθοδοι αντί της χρήσης γαντιού είναι οι μέθοδοι που χρησιμοποιούν αλγόριθμους για την ανίχνευση του χρώματος του χεριού. Οι μέθοδοι που παρουσιάζονται στο (Manresa, Varona, Mas and Perales, 2005) και στο (Jaimés and Jianyi, 2005) είναι κάποιες από αυτές. Και στις δύο περιπτώσεις χρησιμοποιήθηκε απλό υλικό (πχ. κάμερες του εμπορίου και ένας απλός υπολογιστής) και στόχος ήταν το χαμηλό κόστος υλοποίησης των συστημάτων. Ενώ στο (Jaimés and Jianyi, 2005) το πεδίο αλληλεπίδρασης του χρήστη είναι μια επιφάνεια εργασίας με κάποιες προκαθορισμένες περιοχές ενδιαφέροντος, στο (Manresa, Varona, Mas and Perales, 2005) ο χρήστης έχει μεγαλύτερη ελευθερία κινήσεων σε όλο το χώρο. Όμως, ένας περιορισμός της μεθόδου αυτής είναι ότι μπορεί να εκτελέσει μόνον τέσσερις κινήσεις σε τέσσερις διαφορετικές κατευθύνσεις στο χώρο. Άρα η ελευθερία κινήσεων του χρήστη και η ευελιξία του συστήματος είναι αρκετά περιορισμένες σε σχέση με το δεύτερο σύστημα, παρόλο που οι κινήσεις σε αυτό εκτελούνται μόνον σε ένα επίπεδο επιφάνειας. Με βάση τα πειράματα, η ακρίβεια των αποτελεσμάτων της πρώτης μεθόδου ήταν περίπου 99% και ο χρόνος απόκρισης πολύ ικανοποιητικός, κάτι που κάνει το σύστημα αρκετά αξιόπιστο και αποδοτικό, ενώ στη δεύτερη μέθοδο, μετά από το στάδιο εκπαίδευσης που

είναι και το πιο χρονοβόρο, κάτω υπο συγκεκριμένες συνθήκες η ακρίβεια του μπορεί να φτάσει ακόμα και το 100%.

Στο (Dardas and Alhaj, 2011), παρόλο που παρουσιάζεται ένα αρκετά αποδοτικό και ακριβές σύστημα, το οποίο λειτουργεί με οπτική αναγνώριση και είναι αρκετά χαμηλού κόστους, το πεδίο δοκιμών του είναι αρκετά περιορισμένο (γίνεται αξιολόγηση μόνο για τέσσερις τυποποιημένες και συγκεκριμένες χειρονομίες), χωρίς να προσφέρει ελευθερία και ευελιξία στις κινήσεις του χρήστη ως προς τις χειρονομίες που θα αναγνωρίζονται από το σύστημα. Αυτό αποτελεί από μόνο του ένα σοβαρό μειονέκτημα έναντι των υπόλοιπων μεθόδων, ενώ και το εκτενές στάδιο εκπαίδευσης και προετοιμασίας αποτελεί έναν ακόμη περιορισμό του συστήματος. Σε αντίθεση, στο σύστημα των (Rauterberg, Bichsel, Meier and Fjeld, 1997) ο χρήστης έχει πλήρη ελευθερία κινήσεων, ενώ παροτρύνεται να χρησιμοποιήσει συνδυασμό των κινήσεων του σώματος του με φωνητικές εντολές για μεγαλύτερη ακρίβεια. Το σύστημα αυτό βέβαια είναι πιο πολύπλοκο από το προηγούμενο και συνδυάζει πλήθως αισθητηρίων, ενώ η εφαρμογή του απαιτεί την κατάλληλη διαμόρφωση ενός δωματίου για να μπορεί να γίνει σωστά η αλληλεπίδραση του χρήστη με το σύστημα.

## **Δ. Επίλογος και Μελλοντική Εργασία**

Η πρόσφατη εντυπωσιακή ανάπτυξη εφαρμογών κιναισθητικού ελέγχου έχει γίνει δυνατή κυρίως της διάθεσης διαφόρων μορφών αισθητήρων σε μικρά μεγέθη και την ολοκλήρωσή τους μέσω ασύρματων τεχνολογιών. Σύμφωνα με το «νόμο του Moore» (1958): «Ο αριθμός των transistors που μπορούν να τοποθετηθούν σε ένα ηλεκτρονικό κύκλωμα διπλασιάζεται κάθε 2 χρόνια»... ο νόμος επαληθεύεται από τότε ως σήμερα και η πρόβλεψη είναι ότι θα επαληθεύεται για μια ακόμα 10ετία. Το ίδιο συμβαίνει και με τις τεχνολογίες και επιστήμες που απαιτούνται για την διάδραση με χειρονομίες, οι οποίες τα τελευταία χρόνια έχουν παρουσιάσει σημαντική πρόοδο.

Παρότι έχουν αναπτυχθεί εντυπωσιακές εφαρμογές κιναισθητικού ελέγχου για την Αλληλεπίδραση Ανθρώπου-Υπολογιστή, οι περιορισμοί των παραπάνω τεχνολογιών είναι ακόμα πάρα πολλοί. Οι πιο σημαντικοί είναι οι εξής. Πρώτ' απ' όλα, τα πρότυπα (standards) που υπάρχουν ευρέως δεν είναι διαδεδομένα τεχνολογικά πρότυπα και προδιαγραφές στις οποίες μπορούν να βασιστούν οι σχεδιαστές, το οποίο αν και αφήνει χώρο για δημιουργικές δουλειές, απαιτείται πολύς κόπος για να χτιστούν οι βασικές υποδομές για ένα σωστό και εμπορικό σύστημα.

Επίσης, η ευρήτητα και η πρόοδος αυτών των τεχνολογιών είναι ένας ακόμα παράγοντας, που σε αντίθεση με άλλες περιπτώσεις, όπως αυτήν του Διαδικτύου όπου από νωρίς έγινε η προσπάθεια να είναι ένα ανοικτό μέσο στο ευρύ κοινό και θεσπίστηκε το W3C (World Wide Web Consortium) με στόχο την ανακάλυψη και προώθηση ανοικτών τεχνολογιών, στο πεδίο της κιναισθητικής αλληλεπίδρασης μεταξύ ανθρώπου-υπολογιστή δεν υπάρχουν μη κερδοσκοπικοί φορείς προώθησης των τεχνολογιών. Το αποτέλεσμα είναι ότι η σχετική έρευνα συχνά να κατευθύνεται στη βιομηχανία και να μην ανοίγει προς άλλους ερευνητές. Επίσης, πολλές εφαρμογές υλοποιούνται μόνον για ερευνητικούς σκοπούς, χωρίς εμπορική κατεύθυνση, που αποτελεί ακόμα ένα λόγω απόκλισης του κοινού από τη χρήση αυτών των τεχνολογιών.

Τέλος, οι πολλές και διαφορετικές τεχνολογίες, χωρίς ολοκλήρωση μεταξύ τους απαιτούν προγραμματιστική δουλειά, συχνά επίπονη και με λίγη υπάρχουσα τεκμηρίωση για να στηθούν εφαρμογές κιναισθητικού ελέγχου, ενώ η αξιοπιστία τους δεν είναι ιδιαίτερα μεγάλη. Για παράδειγμα, οι αλγόριθμοι υπολογιστικής όρασης επηρεάζονται αρκετά από τη ταχύτητα των κινήσεων του χρήστη, τον περιβάλλοντα φωτισμό, τα χρώματα του υποβάθρου σε σχέση με τα ρούχα του χρήστη, κ.α. Όλοι αυτοί οι παράγοντες δημιουργούν ένα ομιχλώδες τοπίο στην ανάπτυξη μεθόδων Αλληλεπίδρασης Ανθρώπου-Υπολογιστή και εισάγουν εν γέννη αρκετούς περιορισμούς.

Από την άλλη, παρόλο που οι χρήστες δείχνουν μια μεγάλη ανταπόκριση σε αυτό του είδους τη τεχνολογία και είναι αρκετά ανοιχτοί στο να δεχτούν αυτόν τον νέο τρόπο αλληλεπίδρασης με τους υπολογιστές στη ζωή τους. Απόδειξη είναι η μεγάλη απήχηση που είχαν οι συσκευές Wii και Kinect στο ευρύ κοινό, καθώς και η ραγδαία ανάπτυξη τηλεοράσεων με συστήματα αλληλεπίδρασης μεταξύ του χρήστη και της συσκευής. Συνεπώς, υπάρχουν πολλές προοπτικές ανάπτυξης σε αυτόν τον τομέα, με κύρια

έμφαση στις τεχνικές που βασίζονται στην οπτική αναγνώριση των χειρονομιών/κινήσεων, ενώ πολλές έρευνες τείνουν να επικεντρωθούν στην επίλυση και βελτίωση των ήδη υπαρχόντων τεχνολογιών, με έμφαση κυρίως στη δημιουργία εμπορικών και ανταγωνιστικών προϊόντων (όπως στην περίπτωση της έρευνας του MIT).

Εν κατακλείδι, με το ρυθμό που εξελίσσονται τα συστήματα στον τομέα αλληλεπίδρασης ανθρώπου-υπολογιστή, μπορούμε να περιμένουμε πολλά υποσχόμενες εξελίξεις στο μέλλον. Με τη μεγάλη απείχηση που έχουν αποκτήσει στην καθημερινότητα μας οι 3D εφαρμογές, τα συστήματα αλληλεπίδρασης με χειρονομίες θα αποκτήσουν άλλον νόημα. Είναι πολύ πιθανόν στο μέλλον ο χρήστης να μπορεί να επιλέξει και να χειριστεί διάφορα ψηφιακά αντικείμενα και κείμενα με μεγάλη ακρίβεια, να μπορεί να ταξιδεύει ψηφιακά από μια τοποθεσία σε μια άλλη και να αλληλεπιδρά με το περιβάλλον της τοποθεσίας αυτής, ενώ ακόμα θα μπορούσε να ελέγχει απομακρυσμένα κάποια συστήματα χρησιμοποιώντας χειρονομίες και χειριστήρια κατάλληλα για διεπαφές 3D χωρίς τη χρήση πληκτρολογίων ή άλλων μέσων. Μια άλλη θετική επίδραση των τεχνολογιών αυτών θα ήταν να δώσουν τη δυνατότητα στο χρήστη να μπορεί να κάνει εισαγωγές συμβολικών δεδομένων σε ένα σύστημα, κάτι που προς το παρόν είναι πολύ δύσκολο και παρουσιάζει ιδιαίτερα προβλήματα στις διεπαφές 3D (ιδιαίτερα όταν δε μπορεί να χρησιμοποιηθεί το πληκτρολόγιο). Τέλος, ίσως δούμε και τη δημιουργία και την ενσάρκωση των ανθρώπων σε ένα εικονικό περιβάλλον, δημιουργώντας εικονικούς κόσμους όπου ο καθένας μας θα μπορεί να αλληλεπιδρά με τα αντικείμενα ή τους άλλους εικονικούς συνανθρώπους του.

## Ε. Βιβλιογραφικές Πηγές

- Αβούρης Ν. (2000). “Εισαγωγή στην Επικοινωνία Ανθρώπου-Υπολογιστή” Εκδόσεις Διάυλος Αθήνα.
- Βολουδάκης. Κ. και Μαρμαροκόπος . Κ (2005). ”Μοντελοποίηση και Αναγνώριση ανθρωπίνων Χειρονομιών στην Επικοινωνία – Ανθρώπου Υπολογιστή”. Πτυχιακή Εργασία. Εθνικό Μετσόβιο Πολυτεχνείο.
- Kortum, P. (2008). HCI Beyond the GUI, Elsevier. <http://www.hcibeyondthegui.com>
- Smith, C.M. (2004). Human Factors in Haptic Interfaces, ACM Crossroads, <http://www.acm.org/crossroads/xrds3-3/haptic.html>
- Rubine, D. (1991). Specifying gestures by example, Proceedings of the 18th annual conference on Computer graphics and interactive techniques, ACM SIGGRAPH, pp. 329 – 337, 1991
- Krueger, M.W (1983). Artificial Reality. Reading, Mass.: Addison-Wesley.
- Buxton, B. (2007). Multi-Touch Systems that I Have Known and Loved, Available from <http://www.billbuxton.com/multitouchOverview.html>
- Myers, B.A. (1998). A Brief History of Human Computer Interaction Technology. Interactions 5, vol. 2, 44-54. <http://www.cs.cmu.edu/~amulet/papers/uihistory.tr.html>
- [http://jnd.org/dn.mss/natural\\_user\\_interfaces\\_are\\_not\\_natural.html](http://jnd.org/dn.mss/natural_user_interfaces_are_not_natural.html)
- Shaffer, D. (2009) Designing Gestural Interfaces, O’Reilly.
- Κουτσαμπάσης, Π. (2011) [Αλληλεπίδρασης Ανθρώπου – Υπολογιστή Ανθρώπου-Υπολογιστή: Μέθοδοι, Αρχές και Παραδείγματα](#), Κλειδάριθμος.
- Quam, D.L. (1990). Gesture Recognition With a DataGlove, Proc. IEEE National Aerospace and Electronics Conf., vol. 2.
- Yamato, J., Ohya, O.,Ishii, K. (1992). Recognizing Human Action in Time Sequential Images Using Hidden Markov Models, Proc. Computer Vision and Pattern Recognition, pp. 379-385. MIT’s Media Lab.
- Starner, T.E., Pentland, A. (1995). Visual Recognition of American Sign Language using Hidden Markov Models, Proc. Int’l Workshop on Automatic Face and Gesture Recognition, pp. 189-194.
- Starner, T.E., Pentland, A. (1995). Real-Time American Sign Language Recognition from Video using Hidden Markov Models, Technical Report TR-375.
- Nam, Y., Wohn, K.W. (1995). A state-based technique for the summarization and recognition of gesture, Proc. Fifth International Conference on Computer Vision, pp. 382-388.
- Wilson, A.D., Bobick, A.F. (1997). A State-Based Approach to the Representation and Recognition of Gesture, IEEE Transactions on Pattern Analysis And Machine Intelligence, 1, No 12 pp. 1325-1337.
- Wilson, A.D., Bobick, A.F. (1995). Using Configuration States for The Representation and Recognition of Gesture, Technical Report, MIT Media Lab Perceptual Computing Group, No 308.
- Darell, T.J., Pentland, A.P. (1992). Recognition of Space -Time Gestures using a Distributed Representation, MIT Media Laboratory Perceptual Computing Group TR 197.
- Darell, T.J., Pentland, A.P. (1993). Space-Time Gestures, Proc. Computer Vision and Pattern Recognition, pp. 335-340.
- Hyeon-Kyu, L., Jin, H.K. (1999). An HMM Based Threshold Model Approach for Gesture Recognition, IEEE Transactions on Pattern Analysis And Machine Intelligence, 21, No 10.
- Karpouzis, K., Raouzaïou, A., Drosopoulos, A., Ioannou, S., Balomenos, T., Tsapatsoulis, N., Koliass, S. (2004). Facial Expression and Gesture Analysis for Emotionally-Rich Man-Machine Interaction, Emotionally Rich Man-Machine Interaction Chapter 5, pp. 185-193.

- Vogler, C. Metaxas, D. (1999). Parallel Hidden Markov Models for American Sign Language, Proc. Int'l, Conf. Computer Vision.
- Sato, M., Poupyrev, I., and Harrison, C. (2012). Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 483-492.
- Wang, R.Y. and Popovic, I. (2009). Real-Time Hand-Tracking with a Color Glove, *ACM Transaction on Graphics (SIGGRAPH 2009)*, 28(3).
- Keir, P., Payne, J., Elgoyhen, J., Horner, M., Naef, M. and Anderson, P. (2006). Gesture-recognition with Non-referenced Tracking. In *Proceedings of the IEEE conference on Virtual Reality (VR '06)*. pp. 151 – 158, IEEE Computer Society, Washington, DC, USA.
- Manresa, C., Varona, J., Mas, R. and Perales, F. J. (2005) Hand Tracking and Gesture Recognition for Human-Computer Interaction, In *Electronic Letters on Computer Vision and Image Analysis*, Vol. 5, No. 3, pp. 96-104.
- Kimura, D. (1993). *Neuromotor mechanisms in human communication*. Oxford, UK: Oxford University Press.
- Kendon, A. (1988). How gestures can become like words. In Potyatos, F. (ed), *Crosscultural perspectives in nonverbal communication*, p 131-141. Toronto, Canada: Hogrefe.
- McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago, USA: University of chicago press.
- Sturman, D.J., Zeltzer, D. (1993). A design method for "whole-hand human computer interaction". *ACM transactions on information systems*, 11(3), p219-238.
- Pook, P.K.; D.H. Ballard (1995). Teleassistance: A gestural sign language for teleoperation. Position paper for the workshop gesture at the user interface, CHI 95, Denver, CO, USA, May 1995.
- Kim, H. and Fellner, D.W. (2004). Interaction with Hand Gesture for a Back-Projection Wall. In *Proceedings of the Computer Graphics International (CGI '04)*. IEEE Computer Society, Washington, DC, USA, 395-402.
- Dardas, N. H. and Alhaj, M. (2011). Hand Gesture Interaction with a 3D Virtual Environment. *International Journal of The Research Bulletin of Jordan*. ACM, pp. 86-94.
- Leber, J. (2012). A New Chip to Bring 3-D Gesture Control to Smartphones.  
<http://www.technologyreview.com/news/507161/a-new-chip-to-bring-3-d-gesture-control-to-smartphones/>
- Saenz, A. (2009). New Hitachi TV Controlled By Gestures, <http://singularityhub.com/2009/10/29/new-hitachi-tv-controlled-by-gestures-video/>
- Miller, P. (2008). Toshiba's Cambridge Research Lab shows off gesture-controlled TVs, image recognition, <http://www.engadget.com/2008/08/29/toshibas-cambridge-research-lab-shows-off-gesture-controlled-tv/>
- Saenz, A. (2010). Hand Gesture Controls Trying For Mainstream in 2010, <http://singularityhub.com/2010/02/04/hand-gesture-controls-trying-for-mainstream-in-2010-video/>
- Rauterberg, M., Bichsel, M., Meier, M., Fjeld, M., (1997). A gesture based interaction technique for a planning tool for construction and design. *IEEE International Workshop on Robot and Human Communication*.
- Jaimes, Ro and Jianyi, Liu. (2005). Hotspot components for gesture-based interaction. In *Proceedings of the 2005 IFIP TC13 international conference on Human-Computer Interaction*. 3585, 1062-1066.
- Garg, P. & Aggarwal, N, & Sofat, S. (2009). Vision Based Hand Gesture Recognition. *World Academy of Science, Engineering and Technology* 49.



- Oskoei, M. A. & Huosheng Hu. (2009). Application of Feature Tracking in a Vision Based Human Machine Interface for XBOX. Proceedings of the 2009 IEEE International Conference on Robotics and Biomimetics December 19 -23, 2009, Guilin, Chin.
- Teixeira, J. M. & Farias, T. & Moura, G. & Lima, J. & Pessoa, S. & Teichrieb, V. (2006). Gefighters: an experiment for gesture-based interaction analysis in a fighting game. In SBGames (Brazil).
- Van Beurden, M. & IJsselsteijn, W. (2010). Range and variability in gesture-based interactions with medical images: Do non-stereo versus stereo visualizations elicit different types of gestures? In IEEE Virtual Reality 2010 Workshop on Medical Virtual Environments.
- Hassanpour, R., and Shahbahrami, A., (2010). Human Computer Interaction Using Vision-Based Hand Gesture Recognition. Journal of Advances in Computer Research, Feb., 21—30.
- Lenman, S., L., Bretzner, Thuresson, B., (2002). Computer Vision Based Hand Gesture Interfaces for Human-Computer Interaction. Technical Report TRITA-NA-D0209, CID-172, Royal Institute of Technology, Sweden.
- Gan, L., (2011). Real-time Immersive human-computer interaction based on tracking and recognition of dynamic hand gestures. Doctoral thesis, University of Central Lancashire.
- Mavridis, N., Giakoumidis N. and Lopes Machado, E., “A novel evaluation framework for tele-operation and a case study on natural human-arm-imitation through motion capture,” International Journal of Social Robotics, vol. 4, no. 1s, pp. 5-18, November 2012. DOI: 10.1007/s12369-011-0117-8
- Xbox 360 Handbücher und Spezifikationen, (2010). <http://support.xbox.com/de-DE/xbox-360/manuals-specs/manual-specs>
- Samsung, 2012. <http://www.samsung.com>
- [Katzmaier](http://reviews.cnet.com/8301-33199_7-57411497-221/samsung-smart-interaction-hands-on-with-voice-and-gesture-control/), D., 2012. Samsung Smart Interaction: Hands-on with voice and gesture control [http://reviews.cnet.com/8301-33199\\_7-57411497-221/samsung-smart-interaction-hands-on-with-voice-and-gesture-control/](http://reviews.cnet.com/8301-33199_7-57411497-221/samsung-smart-interaction-hands-on-with-voice-and-gesture-control/)