

# ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ

ΥΠΟΛΟΓΙΣΤΩΝ

ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΔΙΚΤΥΩΝ



## ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΜΕΑΣ

ΓΡΑΦΙΚΑ ΥΠΟΛΟΓΙΣΤΩΝ

ΘΕΜΑ

ΑΝΙΧΝΕΥΣΗ ΣΕ ΤΡΙΣΔΙΑΣΤΑΤΟ ΕΙΚΟΝΙΚΟ ΠΕΡΙΒΑΛΛΟΝ  
ΜΟΥΣΕΙΟΥ

ΕΠΙΒΛΕΠΟΝΤΕΣ ΚΑΘΗΓΗΤΕΣ: ΧΟΥΣΤΗΣ ΗΛΙΑΣ  
ΜΠΟΖΑΝΗΣ ΠΑΝΑΓΙΩΤΗΣ

ΟΝΟΜΑΤΑ: ΜΑΝΩΛΑ ΖΩΗ  
ΜΑΡΝΑΡΗ ΑΛΕΞΙΑ

## Ευχαριστίες

Αρχικά, θα θέλαμε να ευχαριστήσουμε τον καθηγητή μας και βασικό επιβλέποντα της εργασίας μας, κύριο Ηλία Χούστη, για τη δυνατότητα ενασχόλησης με ένα θέμα άκρως ενδιαφέρον και πρωτοποριακό στον τομέα των γραφικών ηλεκτρονικών υπολογιστών και την καθοδήγησή του κατά τη διάρκεια εκπόνησης του.

Ευχαριστούμε θερμά την κυρία Ιωάννα Καφφέ για τις παρατηρήσεις και τις υποδείξεις της για την βελτίωση της εργασίας, καθώς και τον κύριο Αθανάσιο Γκαιτατζή, προϊστάμενο του Τμήματος Εικονικής Πραγματικότητας στο Ίδρυμα Μείζονος Ελληνισμού, για την προτροπή ενασχόλησης με το συγκεκριμένο θέμα και την διάθεση κάποιου απαραίτητου υλικού.

Ακόμα, ευχαριστούμε τον πρόεδρο του τμήματος μας κύριο Γεώργιο Σταμούλη, για την στήριξη του και τις συμβουλές του καθ' όλη τη διάρκεια των σπουδών μας.

Τέλος, ευχαριστούμε την οικογένεια μας και τους φίλους μας για την συνεχή κατανόηση και συμπαράσταση τους.

# Περιεχόμενα

Εισαγωγή .....	6
Αντικείμενο και οργάνωση της μελέτης .....	8
Κεφάλαιο 1 : Γραφικά ηλεκτρονικών υπολογιστών .....	10
1.1 Εισαγωγή .....	10
1.2 Βασικές έννοιες .....	11
1.2.1 Βασικοί ορισμοί .....	11
1.2.2 Κατηγορίες γραφικών .....	14
1.2.3 Μοντέλο χρώματος .....	15
1.3 Μετασχηματισμοί (Transformations) .....	17
1.4 Διαδικασία Rendering .....	18
1.4.1 Μέθοδοι Rendering .....	19
1.4.2 Προοπτική προβολή .....	21
1.4.3 Τρισδιάστατα αντικείμενα .....	23
1.4.4 Το φαινόμενο σκίασης (shading) .....	23
1.4.5 Εισαγωγή φωτός και υλικών για τα αντικείμενα .....	24
1.4.6 Ορισμός βάθους οθόνης .....	24
1.4.7 Culling τρισδιάστατων αντικειμένων .....	25
1.4.8 Αντικείμενα έξω από την οθόνη .....	25
Κεφάλαιο 2 : Στερεοσκοπία (Stereoscopy) .....	26
2.1 Εισαγωγή .....	26
2.2 Στερεοσκοπικές και αυτοστερεοσκοπικές τεχνικές .....	27
2.2.1 Αυτοστερεοσκοπικές τεχνικές .....	27
2.2.2 Στερεοσκοπικές τεχνικές .....	28
2.3 Οι βασικές έννοιες της στερεοσκοπίας .....	32
2.3.1 Τρισδιάστατη προβολή σε ένα επίπεδο .....	32
2.3.2 Βασική μονοσκοπία .....	34
2.3.3 Δύο μάτια, μία οθόνη, δύο εικόνες .....	35
2.3.4 Interaxial .....	36
2.3.5 Βάθος οθόνης (Screen depth) .....	37
2.3.6 Αριστερή/Δεξιά προβολή .....	37
2.3.7 Παράλλαξη .....	38
2.3.8 Παράγοντας βάθους (Depth factor) .....	39
2.3.9 Μέσα/Έξω από την οθόνη .....	40
Κεφάλαιο 3 : Συσκευές εισόδου .....	41
3.1 Εισαγωγή .....	41
3.2 Πληκτρολόγιο (Keyboard) .....	42
3.3 Ποντίκι (Mouse) .....	43
3.3.1 Περιγραφή και λειτουργία .....	44
3.4 Joystick .....	47
3.4.1 Εισαγωγή .....	47
3.4.2 Τεχνικά χαρακτηριστικά .....	48
3.5 Webcam .....	49
3.5.1 Ορισμός .....	49
	3

3.6	Σύνθετες συσκευές .....	51
3.6.1	Wii remote .....	51
Κεφάλαιο 4 : Ανίχνευση (tracking) .....		54
4.1	Εισαγωγή .....	54
4.1.1	Τρισδιάστατη ανίχνευση στο πεδίο Computer Vision .....	55
4.1.2	Μαθηματικά εργαλεία .....	56
4.1.3	Παραμετροποίηση της “στάσης” της κάμερας (Camera Pose Parameterization) .....	60
4.1.4	Εκτίμηση του πίνακα εξωτερικών παραμέτρων .....	62
4.1.5	Τεχνικές βελτιστοποίησης ελαχίστων τετραγώνων .....	66
4.2	Face tracking .....	68
4.2.1	Εισαγωγή .....	68
4.2.2	Βασικό μαθηματικό υπόβαθρο .....	70
4.3	Eye Tracking .....	71
4.3.1	Μέθοδοι .....	72
4.3.2	Τεχνολογίες .....	74
4.3.3	Ανιχνεύοντας κινήσεις ματιού .....	74
4.3.4	Μελλοντική αξία .....	78
4.4	Hand Tracking .....	79
4.4.1	Εισαγωγή .....	79
4.4.2	Η διαδικασία ανίχνευσης χεριού .....	80
Κεφάλαιο 5 : Το ελεύθερο λογισμικό Sweet Home 3D .....		85
5.1	Εισαγωγή .....	85
5.2	Προετοιμασία συστήματος .....	86
5.2.1	Ελάχιστες Προδιαγραφές .....	86
5.2.2	Εγκατάσταση προγράμματος .....	86
5.3	Περιγραφή προγράμματος .....	86
5.4	Σχεδίαση ενός νέου χώρου .....	89
5.5	Άλλα χαρακτηριστικά του προγράμματος .....	95
5.6	Αξιολόγηση προγράμματος .....	96
Κεφάλαιο 6 : Το ελεύθερο λογισμικό EaZD .....		97
6.1	Εισαγωγή .....	97
6.2	Προετοιμασία συστήματος .....	98
6.2.1	Ελάχιστες Προδιαγραφές .....	98
6.2.2	Εγκατάσταση προγράμματος .....	98
6.3	Η διεπαφή deferred renderer .....	98
6.3.1	Εισαγωγή του deferred renderer σε εφαρμογές .....	99
6.3.2	Γεωμετρικοί “κάδοι” (geometry bins) .....	100
6.3.3	Φώτα και σκιές .....	101
6.3.4	Το μοντέλο του ουρανού .....	102
6.3.5	Απόδοση .....	103
6.4	Η διεπαφή SceneGraph3D .....	104
6.4.1	Εισαγωγή του SceneGraph3D σε εφαρμογές .....	106
6.4.2	Αρχικοποίηση και φόρτωση σκηνής .....	106
6.4.3	Ο κύκλος ζωής του SceneGraph .....	107
6.4.4	Η γλώσσα περιγραφής μιας σκηνής του SceneGraph3D .....	108

Κεφάλαιο 7 : Η εφαρμογή .....	110
7.1 Η βιβλιοθήκη OpenCV .....	110
7.1.1 OpenCV και ανίχνευση αντικειμένων .....	111
7.1.2 Διαδικασία ανίχνευσης.....	113
7.1.3 Χρησιμοποιούμενες δομές OpenCV στην εφαρμογή μας.....	114
7.1.4 “Σύλληψη” εικόνας από μία webcamera .....	114
7.1.5 Κώδικας για τις δικές μας ανιχνεύσεις.....	116
7.2 Σχεδιασμός εικονικού μουσείου .....	121
7.3 Σύνδεση αρχείου κάμερας με το EaZD .....	127
7.3.1 Αρχικά βήματα .....	127
7.3.2 Δήλωση γεγονότων (event declaration) .....	129
7.3.3 Συγγραφή αρχείου επικεφαλίδας .....	129
7.3.4 Συγγραφή αρχείου πηγαίου κώδικα .....	129
7.3.5 Συγγραφή αρχείου σκηνης .....	130
 Κεφάλαιο 8 : Εφαρμογές και συσκευές που σχετίζονται με ανίχνευση.....	142
8.1 Charade .....	143
8.2 Wii Nintendo .....	144
8.3 EyeToy .....	146
8.3.1 Εισαγωγή .....	146
8.3.2 Ιστορία .....	147
8.3.3 Σχεδιασμός .....	147
8.3.4 Τεχνικοί Περιορισμοί .....	148
8.3.5 Επιπλέον δυνατότητες .....	148
8.4 CamSpace .....	149
8.5 Project Natal .....	150
8.5.1 Ιστορία .....	151
8.5.2 Τεχνολογία .....	151
8.6 EyePoint .....	152
 Κεφάλαιο 9 : Προτάσεις για μελλοντική βελτίωση της εφαρμογής.....	155
 Βιβλιογραφία .....	156
 Παράρτημα .....	159

## Εισαγωγή

Η ρεαλιστική αναπαράσταση εικόνων και σκηνών αποτελεί τον απόλυτο στόχο της τεχνολογίας απεικόνισης, που θα ωφελήσει ένα μεγάλο εύρος εφαρμογών πληροφορικής, εκπαίδευσης, ιατρικής και ψυχαγωγίας. Η αναπαράσταση σε δύο διαστάσεις έχει αναπτυχθεί οριακά και το επόμενο λογικό βήμα είναι η χρήση τριών διαστάσεων, προσφέροντας φυσική αίσθηση του βάθους που είναι αναπόσπαστο συστατικό της ανθρώπινης όρασης. Χάρη στην ταχεία αύξηση της υπολογιστικής ισχύος και την πρόοδο των οπτικών συστημάτων, το ενδιαφέρον της ερευνητικής κοινότητας στράφηκε τα τελευταία χρόνια στην ανάπτυξη εφαρμογών τρισδιάστατης εικόνας.

Τα γραφικά υπολογιστών (computer graphics) είναι ένας κλάδος της επιστήμης των υπολογιστών που ασχολείται με τη θεωρία και την τεχνολογία σύνθεσης εικόνων σε ηλεκτρονικό υπολογιστή. Τα τρισδιάστατα γραφικά υπολογιστών αποτελούν προσπάθειες απεικόνισης γραφικών τριών διαστάσεων στην - απεικόνισης δύο διαστάσεων - οθόνη μιας ψηφιακής συσκευής (π.χ. ενός υπολογιστή). Το γεγονός ότι η απεικόνιση χρησιμοποιεί τρεις διαστάσεις τα καθιστά ιδιαίτερα ρεαλιστικά. Τέτοιου είδους γραφικά χρησιμοποιούνται συνήθως από προγράμματα όπως παιχνίδια υπολογιστών και εικονικούς κόσμους. Τα τρισδιάστατα γραφικά βρίσκουν επίσης εφαρμογή στον κινηματογράφο, για τη δημιουργία σκηνών εικονικών κόσμων αλλά και ειδικών εφέ, που είναι αδύνατον να γυριστούν ως πραγματικές σκηνές.

Η εικονική πραγματικότητα ή Virtual Reality (VR) είναι μία τεχνολογία η οποία υποστηρίζει τη δημιουργία μίας διαφορετικής μορφής διεπαφής (interface) σε ένα υπολογιστικό σύστημα. Η μεγάλη σημασία αυτής της διεπαφής έγκειται στο ότι ο χρήστης καλείται να αλληλεπιδράσει με το σύστημα μέσω πράξεων, κινήσεων και εκτιμήσεων που μοιάζουν με τις καθημερινές του ενέργειες, στο πραγματικό του περιβάλλον και όχι μέσω της πληκτρολόγησης εντολών, ή με την χρήση ποντικιού και κάνοντας κλικ σε παραθυρικά οπτικά στοιχεία. Αν και δεν υπάρχει κάποιος συγκεκριμένος και αυστηρός ορισμός για τον όρο εικονική πραγματικότητα, αντιλαμβάνεται από τους περισσότερους ως η σύνθεση στον υπολογιστή μοντέλων που απεικονίζουν έννοιες, αντικείμενα ή περιβάλλοντα, ρεαλιστικά ή νοητά, σε τρεις ή περισσότερες διαστάσεις. Ταυτόχρονα, ο χρήστης έχει την δυνατότητα να μπει μέσα στο εικονικό περιβάλλον και να το εξερευνήσει σε πραγματικό χρόνο.

Η εικονική πραγματικότητα συνεπάγεται την χρήση προηγμένων τεχνολογιών, συμπεριλαμβανομένων των υπολογιστών και των διάφορων περιφερειακών μονάδων

πολυμέσων, για να παραγάγει ένα μιμούμενο περιβάλλον όπου οι χρήστες αντιλαμβάνονται τα αντικείμενα και τα γεγονότα σαν να είναι στον πραγματικό κόσμο. Με την χρήση των ειδικά σχεδιασμένων μετατροπέων και αισθητήρων, οι χρήστες αλληλεπιδρούν με τις επιδειχθείσες εικόνες, μεταχειρίζονται και μετακινούν τα εικονικά αντικείμενα και εκτελούν διάφορες ενέργειες με τρόπο που να προκαλεί αίσθημα πραγματικής παρουσίας στον εικονικό κόσμο, δηλαδή το αίσθημα της εμπύθισης (Patrice Weiss, Adam S.Jessel, 1998).

Η επιλογή της τεχνολογίας η οποία θα χρησιμοποιηθεί για την ανάπτυξη της εκάστοτε εφαρμογής θα γίνει βάσει του βαθμού κατά τον οποίο η τεχνολογία αυτή ικανοποιεί τις ανάγκες και τις προδιαγραφές οι οποίες περιγράφονται στο σενάριο της ζητούμενης εφαρμογής. Η τεχνολογία των αλληλεπιδραστικών πολυμέσων μπορεί να προσφέρει την δυνατότητα δημιουργίας ενός αλληλεπιδραστικού περιβάλλοντος αλλά το περιεχόμενο αυτό να περιορίζεται στις δύο διαστάσεις. Οποιαδήποτε απόπειρα εισαγωγής της τρίτης διάστασης στην απεικόνιση του χώρου θα έπρεπε να χρησιμοποιεί κίνηση δισδιάστατων κινούμενων στοιχείων πάνω σε ψευδο-τρισδιάστατες (προοπτικές ή αξονομετρικές) απεικονίσεις περιβαλλόντων. Μία τέτοια προσπάθεια δε θα οδηγούσε σε αληθοφανή και ρεαλιστική απεικόνιση χώρου. Η δημιουργία αληθοφανών χώρων θα ήταν δυνατή μόνο μέσω της χρησιμοποίησης τρισδιάστατων γραφικών και animation, αλλά στην περίπτωση αυτή θα περιοριζόμασταν σε επιλεκτική θέαση ψηφιοποιημένου βίντεο αυτού του animation, όπου ο χρήστης θα είχε ελάχιστα περιθώρια επιλογής ως προς το που κινείται.

Αντίθετα, κάνοντας χρήση τεχνολογίας εικονικής πραγματικότητας μας δίνεται η δυνατότητα δημιουργίας περιβαλλόντων καθαρά τρισδιάστατων μέσα στα οποία μπορεί το άτομο να κινηθεί ελεύθερα με τρόπο ρεαλιστικό. Σε αυτά τα περιβάλλοντα είναι δυνατή η εισαγωγή δισδιάστατου περιεχομένου (εικόνες, σκίτσα κλπ.) τόσο εντάσσοντάς το μέσα στο τρισδιάστατο περιβάλλον, όσο και στο συνολικό σύστημα διεπαφής της εφαρμογής. [4]



Εικόνα 1 : Τρισδιάστατη απεικόνιση της αρχαίας αγοράς της Αθήνας στο Google Earth (από το Ίδρυμα Μείζονος Πολιτισμού).

## Αντικείμενο και οργάνωση της μελέτης

Αντικείμενο αυτής της διπλωματικής εργασίας αποτελεί η μελέτη τρόπων ανίχνευσης μερών του ανθρώπινου σώματος σε ένα εικονικό περιβάλλον μουσείου με τη χρήση μιας απλής κάμερας. Συγκεκριμένα, χρησιμοποιήθηκε το πρόγραμμα SweetHome3D για τον σχεδιασμό του μουσείου, το οποίο στη συνέχεια εισήχθη στη μηχανή rendering EaZD, ώστε να προκύψει το τρισδιάστατο εικονικό περιβάλλον εργασίας το οποίο πέρα από την εξωτερική και εσωτερική διαμόρφωση των χώρων παρουσιάζει στους επισκέπτες διάφορες κατηγορίες στερεών τα οποία περιστρέφονται. Μια απλή εμπορική κάμερα μπορεί να ανιχνεύει μέρη του σώματος του χρήστη και συγκεκριμένα, το πρόσωπο, τα χέρια ή τα μάτια του, με αποτέλεσμα να σταματά η περιστροφή των στερεών εντός του μουσείου.

Το 1<sup>ο</sup> κεφάλαιο αποτελεί μια αναφορά στα γραφικά ηλεκτρονικών υπολογιστών, τις βασικές έννοιες και κατηγορίες των γραφικών, τους μετασχηματισμούς που υποστηρίζονται και τις τεχνικές rendering.

Το 2<sup>ο</sup> κεφάλαιο αναφέρεται στη τρισδιάστατη αντίληψη των αντικειμένων μέσω της στερεοσκοπίας. Αναλύονται οι βασικές έννοιες της στερεοσκοπίας καθώς και οι τεχνικές που χρησιμοποιούνται για την τρισδιάστατη παρατήρηση (στερεοσκοπικές και αυοστερεοσκοπικές).

Το 3<sup>ο</sup> κεφάλαιο δίνει μια συνοπτική εικόνα για τις διάφορες συσκευές που μπορούν να αποτελέσουν είσοδο σε έναν ηλεκτρονικό υπολογιστή, ξεκινώντας από το ποντίκι και



το πληκτρολόγιο που αποτέλεσαν τις πρώτες συσκευές εισόδου και καταλήγοντας σε πιο καινοτόμες συσκευές που παρέχουν απεριόριστες δυνατότητες σε πρωτοποριακές εφαρμογές.

Μια εισαγωγή στην ανίχνευση και συγκεκριμένα στην τρισδιάστατη ανίχνευση και τα μαθηματικά εργαλεία που χρησιμοποιούνται, γίνεται στο κεφάλαιο 4. Εκεί, αναλύονται διεξοδικά όλες οι λεπτομέρειες για την ανίχνευση διαφόρων μερών του σώματος του ανθρώπου όπως του προσώπου, των ματιών του και του χεριού του.

Ακολουθεί μια πλήρης περιγραφή του λογισμικού SweetHome3D για την διαρρύθμιση και διακόσμηση εσωτερικών και εξωτερικών χώρων, στο κεφάλαιο 5, ενώ στο κεφάλαιο 6 περιγράφεται αναλυτικά η πλατφόρμα EaZD, η οποία αποτελεί μια ιδιαίτερα ενδιαφέρουσα μηχανή rendering.

Λεπτομερής ανάλυση της εφαρμογής γίνεται στο κεφάλαιο 7. Αρχικά περιγράφεται η βιβλιοθήκη OpenCV, η οποία διαθέτει μια σειρά από έτοιμες συναρτήσεις και δομές γραμμένες σε C++ για την επεξεργασία εικόνων και βίντεο. Γίνεται αναφορά στις δομές OpenCV που χρησιμοποιήθηκαν στην παρούσα εφαρμογή και εξηγείται ο κώδικας “σύλληψης” ενός πλαισίου από μία webcam, αλλά και πως γίνονται οι ανιχνεύσεις. Ακολουθεί ο τρόπος με τον οποίο συνδέθηκε η κάμερα με το EaZD και στη συνέχεια παρουσιάζεται ο σχεδιασμός του μουσείου με το πρόγραμμα Sweethome3D. Περιγράφεται έπειτα, η δημιουργία του εικονικού μουσείου με εισαγωγή στο EaZD του προκύπτοντος αρχείου από το Sweethome3D.

Στο κεφάλαιο 8 παρουσιάζονται κάποιες έτοιμες εφαρμογές, projects και συσκευές που επιτρέπουν ανίχνευση του ανθρώπινου σώματος, κυρίως στο πεδίο των ηλεκτρονικών παιχνιδιών, χωρίς να αποκλείεται ωστόσο η εισαγωγή τους σε πολλούς άλλους τομείς.

Προτάσεις για τη βελτίωση της παρούσας εργασίας δίνονται στο Κεφάλαιο 9. Και τέλος, ακολουθούν η βιβλιογραφία που χρησιμοποιήθηκε και παράρτημα με την ερμηνεία βασικών όρων της εργασίας.

# Κεφάλαιο 1

## Γραφικά ηλεκτρονικών υπολογιστών

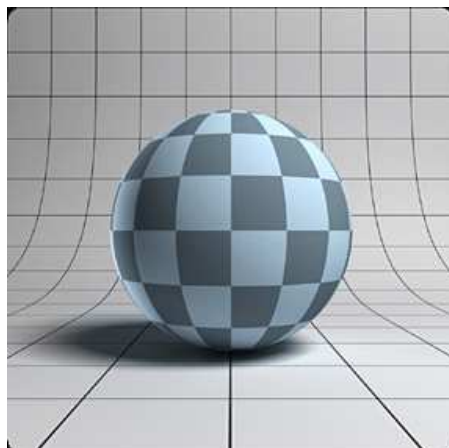
### 1.1 Εισαγωγή

Η λέξη γραφικά προέρχεται από την ελληνική λέξη γραφικός και δηλώνει οπτικές αναπαραστάσεις σε κάποια επιφάνεια, όπως ο τοίχος, το χαρτί ή η οθόνη ενός υπολογιστή για να στιγματίσουν, να ενημερώσουν, να απεικονίσουν ή να διασκεδάσουν. Τέτοια παραδείγματα είναι οι φωτογραφίες, οι ζωγραφιές, τα διάφορα διαγράμματα, οι χάρτες και άλλες εικόνες. Τα γραφικά συνήθως συνδυάζουν κείμενο και χρώμα. [ 26 (όρος graphics)]

Τα γραφικά ηλεκτρονικών υπολογιστών, είναι γραφικά που δημιουργούνται στον υπολογιστή. Η ανάπτυξη τους έχει βοηθήσει αρκετά στην επικοινωνία με τον ηλεκτρονικό υπολογιστή, καθώς την κατέστησαν πιο εύκολη και πιο απλή, ώστε να κατανοηθούν και να ερμηνευτούν καλύτερα πολλοί τύποι πληροφοριών.

Ο όρος γραφικά υπολογιστών χρησιμοποιείται ως γενική έννοια για να περιγράψει “όλα εκείνα στον υπολογιστή που δεν είναι κείμενο ή ήχος”. Τυπικά, όμως, αναφέρεται σε πολλά πράγματα, όπως : η επεξεργασία και αναπαραστάση πληροφοριών εικόνας από έναν υπολογιστή, οι διάφορες τεχνολογίες που χρησιμοποιούνται για την δημιουργία εικόνων και το πεδίο της επιστήμης υπολογιστών που μελετά μεθόδους για την δημιουργία και επεξεργασία ψηφιακού οπτικού περιεχομένου.

Έχουν αναπτυχθεί πολλά ισχυρά εργαλεία για την οπτικοποίηση πληροφοριών. Οι παραγόμενες εικόνες από έναν υπολογιστή μπορούν να κατηγοριοποιηθούν σε διάφορους τύπους : δισδιάστατα (2D), τρισδιάστατα (3D), πέντε διαστάσεων (5D) και κινούμενα (animated) γραφικά. Καθώς η τεχνολογία βελτιώνεται, τα τρισδιάστατα γραφικά υπολογιστών γίνονται όλο και πιο κοινά, αλλά εξακολουθούν να χρησιμοποιούνται ευρέως και τα δισδιάστατα γραφικά. [ 26 (όρος computer graphics)]



Εικόνα 2 : η σφαίρα είναι αδιαμφισβήτητα μία από τις βασικές και πρωταρχικές αρχές των γραφικών υπολογιστών

## 1.2 Βασικές έννοιες

### 1.2.1 Βασικοί ορισμοί

#### *Εικόνα – Image*

Η εικόνα είναι ένα τεχνούργημα που παρουσιάζει αντικείμενα ή/και ανθρώπους. Ο όρος περιλαμβάνει δύο διαστάσεων αντικείμενα, όπως οι φωτογραφίες και μερικές φορές αναπαραστάσεις τριών διαστάσεων. Οι εικόνες “συλλαμβάνονται” από οπτικές συσκευές, όπως οι κάμερες, οι καθρέφτες, τα τηλεσκόπια και άλλες τέτοιες συσκευές, αλλά και από φυσικά πρόσωπα ή φαινόμενα, όπως το ανθρώπινο μάτι και οι επιφάνειες νερού.

Μία ψηφιακή εικόνα (digital image) είναι η αναπαράσταση μιας δύο διαστάσεων εικόνας σε δυαδική μορφή, ως μια ακολουθία δηλαδή μηδενικών και ένα. Οι ψηφιακές εικόνες υπάρχουν τόσο στα vector graphics, όσο και στα bitmap images, αλλά στα τελευταία χρησιμοποιούνται περισσότερο. [ 26 (όρος computer graphics)]

#### *Pixels*

Το pixel (ή στοιχείο εικόνας) είναι ένα σημείο σε μια εικόνα της κατηγορίας γραφικών bitmap. Συνήθως, είναι τοποθετημένα σε ένα πλέγμα δύο διαστάσεων και αναπαριστώνται με την χρήση τελειών ή μικρών τετραγώνων. [ 26 (όρος computer

graphics]] Ο αριθμός των pixels σε μια εικόνα καθορίζει την ανάλυση της εικόνας. Η ανάλυση μιας εικόνας κυμαίνεται μεταξύ 320\*200 και 2000\*1500. [7]

Κάθε pixel αποτελεί ένα δείγμα της αρχικής εικόνας και όπως είναι φυσικό, όσο πιο πολλά δείγματα υπάρχουν, τόσο πιο ακριβής είναι η αρχική αναπαράσταση. Το κάθε pixel φέρει πληροφορίες για την εικόνα, όπως για παράδειγμα την ποσότητα του χρώματος που έχει. [ 26 (όρος computer graphics)]



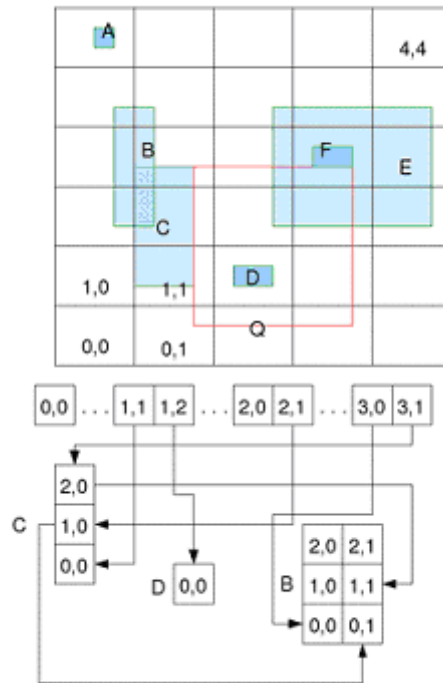
Εικόνα 3 : Μία εικόνα χαμηλής ανάλυσης

### Γεωμετρικό μοντέλο

Οι πληροφορίες της σκηνής αποθηκεύονται στην μνήμη του υπολογιστή. Για την επεξεργασία τους χρησιμοποιείται η υπολογιστική γεωμετρία ή απλά γεωμετρία. Υπάρχουν δυο βασικοί κλάδοι της : η combinatorial computational geometry ή algorithmic geometry, η οποία συμπεριφέρεται στα γεωμετρικά αντικείμενα ως ξεχωριστές οντότητες, και η numerical computational geometry ή machine geometry ή geometric modeling, η οποία ασχολείται κυρίως με την αναπαράσταση των αντικειμένων του πραγματικού κόσμου σε κατάλληλη μορφή για υπολογισμούς στον υπολογιστή (κυρίως σε προγράμματα CAD/CAM). Τα γραφικά χρησιμοποιούν κατά κύριο λόγο τον πρώτο κλάδο της γεωμετρίας αξιοποιώντας τους “κάδους” (geometry bin). Ένας “κάδος” είναι μια δομή δεδομένων που επιτρέπει την αναζήτηση περιοχής με αποδοτικό και εύκολο τρόπο. Αν, για παράδειγμα, στην σκηνή υπάρχουν ορθογώνια, μία αναζήτηση θα μπορούσε να είναι “δεδομένου ενός ορθογωνίου, επιστροφή όλων των ορθογωνίων που σχετίζονται με αυτό”.

Η δομή αυτή χωρίζει την διδιάστατη περιοχή σε “κάδους” ίδιου μεγέθους, συνήθως σχήματος ορθογωνίου. Ένας “κάδος” περιέχει όλα τα υποψήφια ορθογώνια που αναζητούνται. Όλοι οι υποψήφιοι οργανώνονται σε πίνακες δύο διαστάσεων. Το μέγεθος του πίνακα ενός υποψηφίου είναι ο αριθμός των “κάδων” τους οποίους τέμνει. Κάθε κάδος περιέχει την κεφαλή μιας απλά διασυνδεδεμένης λίστας. Εάν ένας

υποψηφίος τέμνει κάποιον “κάδο”, γίνεται στοιχείο της λίστας του “κάδου” αυτού. Κάθε στοιχείο στον πίνακα ενός υποψηφίου είναι ένας κόμβος σύνδεσης στην αντίστοιχη λίστα “κάδου”. [26 (όρος computational geometry)]



Εικόνα 4 : μία δομή “κάδου”. Η απάντηση στην αναζήτηση που αναφέρθηκε για το ορθογώνιο Q είναι τα ορθογώνια C, D, F και E. Επίσης, διακρίνουμε το μέγεθος του υποψηφίου B, έχει 6 στοιχεία, αφού τέμνει 6 “κάδους”, και οργανώνονται σε έναν πίνακα 3 σειρών και 2 στηλών.

*Τρόποι αναπαράστασης της γεωμετρίας της σκηνής στην μνήμη του υπολογιστή.*

### 1. Πολύγωνα

Η πιο κλασική μέθοδος για μοντελοποίηση τρισδιάστατης γεωμετρίας είναι η χρήση πολυγώνων. Ένα αντικείμενο προσεγγίζεται μέσω ενός πολυγωνικού πλέγματος, δηλαδή ένα σύνολο συνδεδεμένων πολυγώνων χρησιμοποιείται, για την δημιουργία του αντικειμένου. Τα πιο κοινά πολύγωνα αυτής της μεθόδου, είναι τα τρίγωνα λόγω απλότητας και γενικότητας. Ο προφανής, όμως, περιορισμός των τριγώνων είναι η παραγωγή ενός επίπεδου αποτελέσματος. Παρ’ όλα αυτά έχουν αναπτυχθεί τεχνικές (smoothing ή interpolation techniques) για την βελτίωση του προβλήματος αυτού. Κάθε πολύγωνο μπορεί να περιγραφεί από την λίστα των συντεταγμένων των κορυφών του στον τρισδιάστατο χώρο.



Εικόνα 5 : Μία αγελάδα μοντελοποιημένη σε ένα πλέγμα τριγώνων

## 2.Πρωτεύοντα στερεά

Οι πιο γνωστές κλασικές γεωμετρικές οντότητες χρησιμοποιούνται απευθείας. Κάποιες από αυτές τις οντότητες είναι ο κύβος, ο κύλινδρος, η σφαίρα και ο κώνος. Για παράδειγμα, μία σφαίρα μπορεί να περιγραφεί από τις συντεταγμένες του κέντρου και της ακτίνας του.

## 3.Συνεχείς καμπύλες (Smooth patches)

Πιο πολύπλοκες μαθηματικές οντότητες επιτρέπουν την αναπαράσταση πολύπλοκων λείων αντικειμένων. Οι πιο γνωστές είναι οι καμπύλες spline και τα NURBS. Είναι, ωστόσο, δύσκολο να τις χειριστούμε, καθώς δεν ελέγχουν απευθείας την επιφάνεια, αλλά χρησιμοποιούν κάποια σημεία ελέγχου τα οποία συνδέονται έμμεσα με το τελικό σχήμα. [7]

### 1.2.2 Κατηγορίες γραφικών

Υπάρχουν δύο είδη γραφικών : raster (παράγονται από pixels) και vector (παράγονται από μονοπάτια ή paths). Τα γραφικά raster είναι πιο γνωστά με την ονομασία bitmap images.

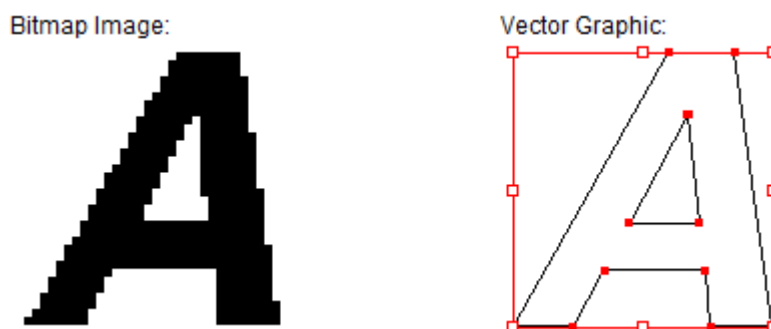
Τα γραφικά bitmap images χρησιμοποιούν ένα πλέγμα ατομικών pixels, όπου κάθε pixel μπορεί να είναι ένα διαφορετικό χρώμα ή μια σκιά.

Τα γραφικά vector χρησιμοποιούν μαθηματικές σχέσεις μεταξύ σημείων και των μονοπατιών που τα συνδέουν για να δημιουργήσουν μια εικόνα.

Το πρόβλημα που αντιμετωπίζουν τα bitmap images είναι η εμφάνιση τους καθώς το μέγεθος αυξάνει. Σε όλο και μεγαλύτερο μέγεθος οι άκρες τους εμφανίζονται όλο και πιο

οδοντωτές. Αυτό διορθώνεται εφαρμόζοντας τεχνικές αντιαύτισης (anti-aliasing techniques). Κατά την διάρκεια αυτών των τεχνικών επιχειρούνται ομαλές μεταβάσεις ανάμεσα στα pixels των γωνιών.

Τα bitmap images απαιτούν υψηλή ανάλυση και επεξεργασία anti-aliasing για μια πιο ομαλή εμφάνιση. Χρησιμοποιούνται με καλύτερα αποτελέσματα σε φωτογραφίες και εικόνες με χαμηλή σκίαση. Αντιθέτως, τα γραφικά vector περιγράφονται μαθηματικά και είναι αρκετά ομαλά σε οποιοδήποτε μέγεθος και με οποιαδήποτε ανάλυση. Ωστόσο, υπάρχει η διαδικασία rasterization κατά την οποία μία εικόνα σε μορφή vector μετατρέπεται σε μορφή raster (αποτελείται πλέον από pixels). [22]

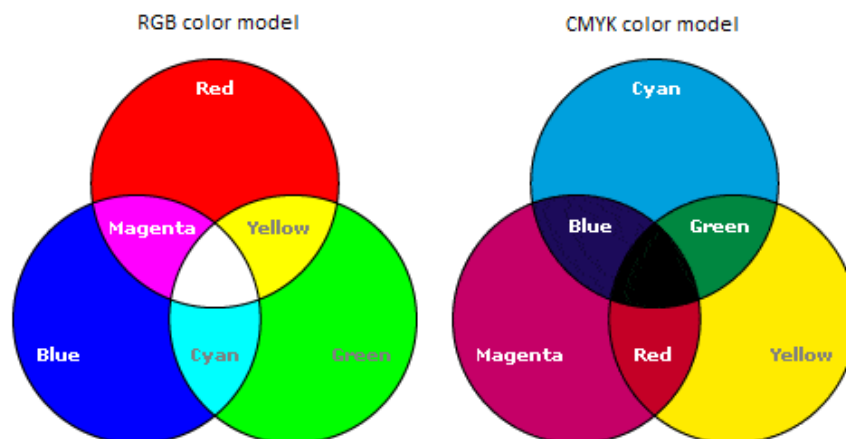


Εικόνα 6 : Αναπαράσταση ενός bitmap image και ενός vector graphic. Είναι 4 φορές μεγαλύτερα από το κανονικό μέγεθος και φαίνεται πως οι γωνίες ενός bitmap γίνονται οδοντωτές καθώς το μέγεθος αυξάνει

### 1.2.3 Μοντέλο χρώματος

Η χρήση χρωμάτων είναι σημαντική για τα γραφικά υπολογιστών. Μοντέλο χρώματος (color model) είναι ένα σύστημα δημιουργίας μεγάλου εύρους χρωμάτων από λίγα βασικά χρώματα. Τα μοντέλα διακρίνονται σε αυτά που είναι αφαιρετικά (subtractive) και σε αυτά που είναι προσθετικά (additive). Τα προσθετικά χρησιμοποιούν φως για να δείξουν κάποιο χρώμα και είναι αποτέλεσμα διαχεόμενου φωτός, ενώ τα αφαιρετικά δίνουν χρώμα που είναι αποτέλεσμα ανακλώμενου φωτός.

Υπάρχουν αρκετοί τύποι μοντέλων χρώματος που χρησιμοποιούνται στα γραφικά, τα βασικά όμως είναι δύο : το RGB (Red-Green-Blue, κόκκινο-πράσινο-μπλε) που χρησιμοποιείται για την απεικόνιση στην οθόνη του υπολογιστή και το CMYK (Cyan-Magenta-Yellow-black, κυανό-μωβ-κίτρινο-μαύρο) που χρησιμοποιείται για την εκτύπωση εικόνων.



Εικόνα 7 : το μοντέλο RGB είναι προσθετικό μοντέλο χρώματος, ενώ το CMYK είναι αφαιρετικό

Η γκάμα χρωμάτων του μοντέλου RGB προέρχεται από τα βασικά προσθετικά χρώματα κόκκινο, πράσινο και μπλε. Όταν συνδυάζονται και τα 3 αυτά χρώματα προκύπτει το άσπρο. Οι υπολογιστές χρησιμοποιούν για κάθε χρώμα του μοντέλου αυτού 24 bits. Έτσι υπάρχουν 256 παραλλαγές για κάθε ένα από τα προσθετικά αυτά χρώματα. Επομένως, τα χρώματα που είναι διαθέσιμα στο μοντέλο είναι 16.777.216 (256 για το κόκκινο X 256 για το πράσινο X 256 για το μπλε).

Τα χρώματα, δηλαδή, είναι αποτέλεσμα διαφορετικών πυκνοτήτων του κόκκινου, του πράσινου και του μπλε φωτός. Η πυκνότητα αντιπροσωπεύεται σε μια κλίμακα από το 0 μέχρι το 255, με το 0 να είναι η μηδενική ποσότητα του συγκεκριμένου φωτός (δεν εκπέμπεται καθόλου τέτοιου είδους φως) και το 255 να είναι η μέγιστη. Για παράδειγμα, για να παράγουμε το χρώμα μωβ, η πυκνότητα του φωτός που πρέπει να πέσει σε κάθε βασικό χρώμα είναι : για το κόκκινο  $R = 255$ , για το πράσινο  $G = 0$  και για το μπλε  $B = 255$ . Το μαύρο παράγεται από την πλήρη απουσία φωτός, δηλαδή  $R = 0$ ,  $G = 0$  και  $B = 0$ .

Το μοντέλο CMYK είναι επίσης γνωστό και ως “διαδικασία τεσσάρων χρωμάτων” (four-color process) ή απλά ως “διαδικασία” (process). Η γκάμα χρωμάτων του μοντέλου προέρχεται από τα βασικά αφαιρετικά χρώματα κυανό, μωβ και κίτρινο. Θεωρητικά, όταν τα μελάνια αυτών των χρωμάτων συνδυάζονται, δίνουν μαύρο. Παρ’ όλα αυτά, εξαιτίας των προσμειξεων στα μελάνια, όταν αυτά τα χρώματα συνδυαστούν δίνουν ένα μουντό καφέ αποτέλεσμα. Το μαύρο μελάνι έχει προστεθεί σε αυτό το σύστημα για να αντισταθμίσει τις προσμειξεις αυτές.

Τα χρώματα, δηλαδή, είναι αποτέλεσμα ποσοστών του κυανού, του μωβ, του κίτρινου και του μαύρου μελανιού. Για παράδειγμα το κόκκινο χρώμα παράγεται παίρνοντας 14% κυανού, 100% μωβ, 99% κίτρινου και 3% μαύρου. Το άσπρο παράγεται από την



πλήρη απουσία μελανιού πάνω σε άσπρο χαρτί, δηλαδή ποσοστό όλων των χρωμάτων 0%. [22]

### 1.3 Μετασχηματισμοί (Transformations)

Οι μετασχηματισμοί είναι αναπόσπαστο κομμάτι των γραφικών. Πρόκειται για γεωμετρικές λειτουργίες που εφαρμόζονται σε όλα τα σημεία ενός αντικειμένου και μπορεί να σχετίζονται με έναν ή με κάποιο συνδυασμό των τριών αξόνων ( $x, y, z$ ). Είναι αποτέλεσμα πολλαπλασιασμού πινάκων. Ένα αντικείμενο μπορεί να μετακινηθεί, να περιστραφεί ή να αλλάξει μέγεθος. Πολλοί μετασχηματισμοί μπορούν να συνδυαστούν μεταξύ τους και να αλλάξουν τελείως το αρχικό αντικείμενο.

#### *Μετατόπιση (Translation)*

Είναι η μετατόπιση ενός σημείου από την αρχική του θέση σε μια άλλη, με βάση μια συγκεκριμένη σταθερά.

#### *Αλλαγή κλίμακας (Scaling)*

Πρόκειται για διαδικασία αλλαγής μεγέθους του αντικειμένου. Είναι είτε μεγέθυνση, είτε σμίκρυνση των συντεταγμένων του.

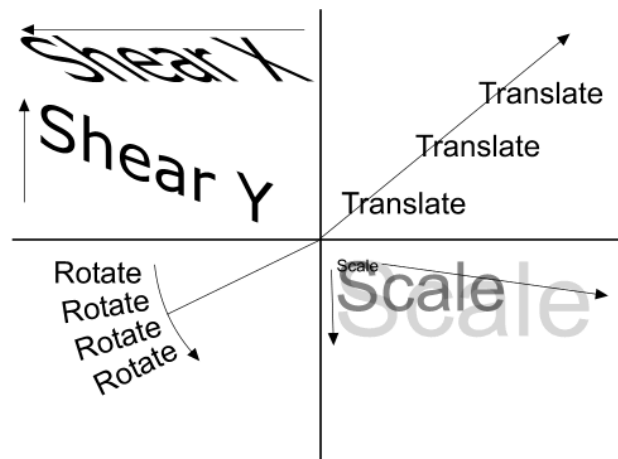
#### *Περιστροφή (Rotation)*

Το αντικείμενο στρέφεται γύρω από τον άξονα της επιλογής μας κατά την γωνία που θέλουμε.

#### *Στρέβλωση (Shearing)*

Παραμορφώνεται το αντικείμενο στον έναν άξονα αναφορικά με τους άλλους άξονες, μετατρέποντας για παράδειγμα ορθογώνια σε παραλληλόγραμμα.

Τα συστήματα διαδικασίας rasterization χρησιμοποιούν, γενικά, μία στοίβα μετασχηματισμών προκειμένου να εφαρμοστούν όλοι οι μετασχηματισμοί και με την κατάλληλη σειρά στο αντικείμενο. Η στοίβα αυτή είναι μια κλασσική στοίβα που αποθηκεύει πίνακες. [14]



Εικόνα 8 : Οι τέσσερις μετασχηματισμοί στους δύο άξονες

## 1.4 Διαδικασία Rendering

Ένα δισδιάστατο/τρισδιάστατο αρχείο σκηνής περιέχει αντικείμενα με βάση μια αυστηρά ορισμένη γλώσσα ή δομή. Για την περιγραφή της σκηνής χρησιμοποιούνται σε συνδυασμό πολλές πληροφορίες : γεωμετρία σκηνής, καθορισμός της οπτικής γωνίας, φωτισμός, σκίαση, υλικό αντικειμένων. Οι πληροφορίες αυτές περνούν, στην συνέχεια, σε ένα πρόγραμμα rendering προκειμένου να επεξεργαστούν και να δώσουν ένα αρχείο εικόνας (είτε σε μορφή vector είτε σε μορφή bitmap). Το πρόγραμμα rendering βασίζεται, συνήθως, σε κάποιο λογισμικό αν και κάποια είναι διαθέσιμα ως ξεχωριστά ανεξάρτητα προγράμματα. [26 (όρος computer graphics)]

Η διαδικασία rendering μετατρέπει ένα μοντέλο σε μια εικόνα. Λέγοντας μοντέλο εννοούμε την περιγραφή ενός τρισδιάστατου αντικειμένου σε κάποια δομή. Η δομή αυτή μπορεί να περιέχει πληροφορίες γεωμετρίας, υψής, φωτισμού, σκίασης και οπτικής αντίληψης του αντικειμένου. Τέλος, λέγοντας εικόνα εννοούμε μια ψηφιακή εικόνα, δηλαδή, την αναπαράσταση μιας δισδιάστατης εικόνας στο δυαδικό σύστημα.

3D rendering είναι η διαδικασία αυτόματης μετατροπής τρισδιάστατων μοντέλων πλέγματος (wire frame models) σε δισδιάστατες εικόνες με τρισδιάστατα ρεαλιστικά στοιχεία στην οθόνη του υπολογιστή. [26 (όρος rendering)]

### 1.4.1 Μέθοδοι Rendering

Πολλοί αλγόριθμοι rendering έχουν ερευνηθεί, ενώ το λογισμικό που χρησιμοποιείται για την διαδικασία αυτή μπορεί να χρησιμοποιήσει διάφορες τεχνικές ώστε να προκύψει η τελική εικόνα.

Η παρακολούθηση κάθε σωματιδίου φωτός σε μια σκηνή είναι σχεδόν πάντα μη εφαρμόσιμη και απαιτεί ένα μεγάλο χρονικό διάστημα. Ακόμα και αν πρόκειται για ένα σύνολο αρκετά μεγάλο ώστε να παραχθεί μια εικόνα, απαιτείται ένα υπερβολικά μεγάλο χρονικό διάστημα, εάν η δειγματοληψία δεν είναι σωστά επιλεγμένη.

Ως εκ τούτου, προέκυψαν τέσσερις οικογένειες πιο αποδοτικών τεχνικών μοντελοποίησης της μεταφοράς φωτός. Η τεχνική rasterization, συμπεριλαμβανομένου του scanline rendering, προβάλλει γεωμετρικά, αντικείμενα της σκηνής στο επίπεδο εικόνας, χωρίς προηγμένα οπτικά εφέ. Η προσέγγιση ray casting θεωρεί τη σκηνή, όπως παρατηρείται από ένα συγκεκριμένο σημείο θέασης, υπολογίζοντας την παρατηρούμενη εικόνα με βάση μόνο τη γεωμετρία και πολύ βασικούς νόμους της έντασης αντανάκλασης. Ακολουθεί η τεχνική ray tracing η οποία είναι παρόμοια με την ray casting, αλλά χρησιμοποιεί πιο προηγμένη οπτική προσομοίωση χρησιμοποιώντας συνήθως τις τεχνικές Monte Carlo για να αποκτήσει πιο ρεαλιστικό αποτέλεσμα. Η τέταρτη τεχνική μεταφοράς φωτός είναι η radiosity. Δεν εφαρμόζεται συνήθως ως μέθοδος rendering, αλλά, αντίθετα, υπολογίζει τη διέλευση του φωτός καθώς αφήνει την πηγή φωτός και φωτίζει τις επιφάνειες. Αυτές οι επιφάνειες προωθούνται συνήθως στην οθόνη, χρησιμοποιώντας μία από τις τρεις άλλες τεχνικές.

#### *Scanline rendering and rasterization*

Μια αναπαράσταση της εικόνας σε υψηλό επίπεδο, περιέχει στοιχεία που δεν θεωρούνται pixels, αλλά αρχέτυπα (primitives). Σε ένα σχηματικό διάγραμμα, για παράδειγμα, τμήματα γραμμών και καμπυλών πιθανόν να είναι primitives, όπως και τα παράθυρα ή τα κουμπιά σε μια γραφική διεπαφή χρήστη, ή τα τρίγωνα και τα πολύγωνα στο 3D rendering.

Αν μία pixel ανά pixel προσέγγιση για την διαδικασία rendering είναι ανέφικτη ή υπερβολικά αργή, μια προσέγγιση primitive ανά primitive μπορεί να αποδειχθεί χρήσιμη. Σε αυτή την περίπτωση εξετάζεται κάθε primitive και καθορίζεται ποια pixel της εικόνας επηρεάζει, ώστε να τροποποιηθούν ανάλογα τα pixels αυτά. Η διαδικασία λέγεται rasterization και είναι η μέθοδος rendering που χρησιμοποιείται από όλες τις σύγχρονες κάρτες γραφικών.

Η διαδικασία rasterization είναι συχνά πιο γρήγορη από την διαδικασία rendering pixel ανά pixel. Πρώτον, μεγάλες περιοχές της εικόνας μπορεί να μην περιέχουν primitives, οπότε οι περιοχές αυτές θα αγνοηθούν, ενώ η διαδικασία pixel ανά pixel πρέπει να περάσει μέσα από αυτές. Δεύτερον, η μέθοδος rasterization μπορεί να βελτιώσει τη συνοχή της κρυφής μνήμης και να μειώσει την περιττή εργασία με την αξιοποίηση του γεγονότος ότι τα pixel που αντιστοιχούν σε ένα primitive τείνουν να είναι συνεχόμενα στην εικόνα. Για τους λόγους αυτούς, η μέθοδος rasterization είναι συνήθως η προσέγγιση που επιλέγεται, όταν απαιτείται διαδραστική διαδικασία rendering. Ωστόσο, η προσέγγιση pixel ανά pixel μπορεί να παράγει εικόνες υψηλότερης ποιότητας και είναι πιο ευέλικτη, διότι δεν εισάγει τόσες υποθέσεις σχετικά με την εικόνα όπως η rasterization.

### *Ray casting*

Στη μέθοδο ray casting η γεωμετρία που μοντελοποιείται αναλύεται pixel προς pixel, γραμμή προς γραμμή. Όταν ένα αντικείμενο τέμνεται, η τιμή του χρώματος στο σημείο τομής μπορεί να υπολογιστεί χρησιμοποιώντας διάφορες μεθόδους. Στην απλούστερη, η τιμή χρώματος του αντικειμένου στο σημείο τομής γίνεται η τιμή του εν λόγω pixel. Μια πιο εξελιγμένη μέθοδος είναι η τροποποίηση της τιμής χρώματος με ένα συντελεστή φωτισμού, αλλά χωρίς τον υπολογισμό της σχέσης του με μια προσομοιωμένη πηγή φωτός.

Επιπρόσθετα μπορούν να χρησιμοποιηθούν προσομοιώσεις οπτικών ιδιοτήτων. Γίνεται ένας απλός υπολογισμός της ακτίνας από το αντικείμενο προς το σημείο θέασης και ένας υπολογισμός της γωνίας πρόσπτωσης των ακτινών από τη φωτεινή πηγή. Από τα αποτελέσματα αυτά, καθώς και από την καθορισμένη ένταση των φωτεινών πηγών, υπολογίζεται η τιμή των pixel.

Η μέθοδος ray casting χρησιμοποιείται κυρίως για προσομοιώσεις πραγματικού χρόνου, όπως αυτές που χρησιμοποιούνται σε τρισδιάστατα παιχνίδια στον υπολογιστή και κινούμενα σχέδια, όπου η λεπτομέρεια δεν είναι σημαντική, ή όταν σκοπός είναι να επιτευχθεί καλύτερη απόδοση σε υπολογιστικό στάδιο. Αυτό συνήθως συμβαίνει όταν ένας μεγάλος αριθμός καρέ πρέπει να μετατραπεί σε κινούμενη εικόνα. Οι προκύπτουσες επιφάνειες έχουν μια χαρακτηριστική επίπεδη εμφάνιση, όταν δεν χρησιμοποιούνται πρόσθετα τεχνάσματα.

### *Ray tracing*

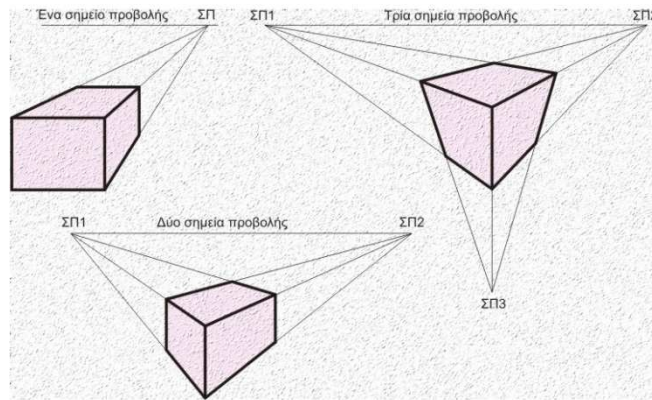
Η μέθοδος ray tracing έχει ως στόχο να προσομοιώσει την φυσική ροή του φωτός, η οποία θεωρείται ως σωματίδια. Χρησιμοποιούνται γνωστοί νόμοι της οπτικής όπως “η γωνία πρόσπτωσης είναι ίση με τη γωνία αντανάκλασης” και πιο εξελιγμένοι νόμοι που σχετίζονται με την διάθλαση και την ομαλότητα των επιφανειών.

Σκοπός είναι η δημιουργία μιας εικόνας με τον εντοπισμό της διαδρομής του φωτός μέσα από pixels σε ένα επίπεδο εικόνας και προσομοίωση των συνεπειών των συναντήσεων του με εικονικά αντικείμενα. Η τεχνική είναι σε θέση να παράγει έναν πολύ υψηλό βαθμό οπτικού ρεαλισμού, συνήθως υψηλότερο από εκείνο των τυπικών μεθόδων scanline rendering, αλλά με μεγαλύτερο υπολογιστικό κόστος. Το γεγονός αυτό καθιστά την ray tracing καταλληλότερη για εφαρμογές όπου η εικόνα εμφανίζεται αργά κατά τον χρόνο, όπως σε φωτογραφίες και φιλμ ή τηλεοπτικά ειδικά εφέ, ενώ είναι πιο ακατάλληλη για εφαρμογές σε πραγματικό χρόνο, όπως σε ηλεκτρονικά παιχνίδια, όπου η ταχύτητα είναι κρίσιμη. Η μέθοδος ray tracing είναι ικανή για προσομοίωση μια ευρείας ποικιλίας οπτικών εφέ, όπως ανάκλαση και διάθλαση, σκέδαση και χρωματική εκτροπή. [26(όρος rendering)]

### **1.4.2 Προοπτική προβολή**

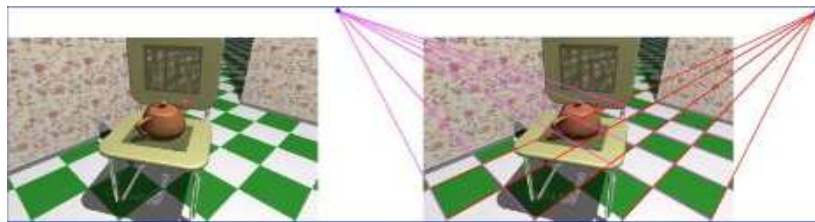
Οι συσκευές απεικόνισης είναι δισδιάστατες, το μαθηματικό μοντέλο, όμως, είναι τρισδιάστατο. Εμείς προσπαθούμε να προβάλλουμε τα τρισδιάστατα αντικείμενα σε ένα επίπεδο προβολής. Προσπαθούμε, δηλαδή, να βρούμε μετασχηματισμούς μετατροπής των τρισδιάστατων αντικειμένων από το παγκόσμιο σύστημα συντεταγμένων στο σύστημα συντεταγμένων παρατηρητή.

Η προοπτική προβολή (perspective projection) πλησιάζει αρκετά σε αυτό που πραγματικά βλέπει το ανθρώπινο μάτι. Αυτό συμβαίνει γιατί στην προβολή αυτή, ο παρατηρητής θεωρείται ότι βρίσκεται κοντά στο αντικείμενο. Στην προοπτική προβολή υπάρχουν τέσσερα βασικά χαρακτηριστικά: η θέση του παρατηρητή, το αντικείμενο το οποίο προβάλλεται, το επίπεδο προβολής και τέλος οι ακτίνες προβολής που σχεδιάζονται από τη θέση του παρατηρητή προς όλα τα σημεία του αντικειμένου.



Εικόνα 9 : παρουσιάζεται ένα αντικείμενο όπως αυτό έχει σχεδιαστεί προοπτικά χρησιμοποιώντας ένα, δύο και τρία σημεία προβολής.

Στην προοπτική προβολή τα αντικείμενα που είναι πιο κοντά στον παρατηρητή φαίνονται μεγαλύτερα. Επίσης, οι παράλληλες ευθείες φαίνεται να συγκλίνουν σε κάποιο σημείο. Όταν σχεδιάζουμε τρισδιάστατα αντικείμενα σκεφτόμαστε την οθόνη ως ένα δισδιάστατο παράθυρο στον τρισδιάστατο κόσμο. Η γεωμετρία που χρησιμοποιούμε είναι αυτή των όμοιων τριγώνων.



Εικόνα 10 : λόγω της προοπτικής προβολής οι παράλληλες ευθείες συγκλίνουν σε κάποιο σημείο

Η τρισδιάστατη σκηνή προβάλλεται από ένα σημείο, το λεγόμενο κέντρο προβολής που δεν είναι κάποιο άλλο από την θέση της κάμερας ή του παρατηρητή σε ένα επίπεδο προβολής.

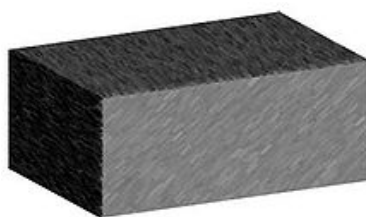
Σε εφαρμογές πραγματικού χρόνου, με την διαδικασία rendering επιθυμούμε να δείξουμε όσο πιο πολλές πληροφορίες μπορεί να επεξεργαστεί το μάτι σε ένα τριακοστό του δευτερολέπτου. Ο κύριος στόχος εδώ είναι η ταχύτητα. Στην περίπτωση αυτή, εκμεταλλευόμαστε τον τρόπο που το μάτι ερμηνεύει τον κόσμο, έτσι η τελική εικόνα δεν είναι κατά ανάγκη τέλεια απεικόνιση του πραγματικού κόσμου, αλλά αρκετά κοντά σε αυτό που αντιλαμβάνεται το ανθρώπινο μάτι. [1]

### 1.4.3 Τρισδιάστατα αντικείμενα

Όλα τα τρισδιάστατα αντικείμενα πρέπει να αποδοθούν χρησιμοποιώντας μία μοναδική προοπτική προβολή σε ένα δεδομένο πλαίσιο. Για παράδειγμα, ο χώρος του ουρανού (sky box) πρέπει να σχεδιαστεί δίνοντας ρεαλιστικό βάθος, δηλαδή, να βρίσκεται πολύ πιο μακριά από την κανονική σκηνή. Τέλος, όλα τα τρισδιάστατα αντικείμενα πρέπει να έχουν βάθος, με συνοχή, που να ταιριάζει με την σκηνή στην οποία βρίσκονται. [10]

### 1.4.4 Το φαινόμενο σκίασης (shading)

Το φαινόμενο σκίασης αναφέρεται στην τεχνική απόκτησης βάθους σε τρισδιάστατα μοντέλα, χρησιμοποιώντας διαφορετικά επίπεδα σκοταδιού. Πρόκειται για μια διαδικασία στην σχεδίαση κατά την οποία, το αντικείμενο όσο απομακρύνεται από το μάτι σχεδιάζεται πιο πυκνό και πιο επισκιασμένο, ενώ όσο πλησιάζει γίνεται λιγότερο πυκνό και επισκιασμένο. [26 (όρος computer graphics)]



Εικόνα 11 : Παράδειγμα ενός αντικειμένου σχεδιασμένο με βάση το φαινόμενο σκίασης

Η σκίαση αναφέρεται επίσης στην εφαρμογή shaders. Με τον όρο shader περιγράφουμε ένα σύνολο εντολών, σε κάποιο λογισμικό, που χρησιμοποιείται κυρίως για τον υπολογισμό επιδράσεων φαινομένων rendering με έναν υψηλό βαθμό ευελιξίας. Χρησιμοποιούνται, δηλαδή, για να προγραμματίσουν την μονάδα επεξεργασίας γραφικών (graphics processing unit, GPU), η οποία έχει αντικαταστήσει κατά μεγάλο βαθμό την διαδικασία που επέτρεπε μόνο κοινούς γεωμετρικούς μετασχηματισμούς. Οι γλώσσες προγραμματισμού που αναπτύσσονται στα πλαίσια αυτά είναι γνωστές ως shading languages. Λόγω της ευρείας ανάπτυξης των γραφικών δημιουργήθηκαν αρκετές τέτοιες γλώσσες, με πιο ευρέως διαδεδομένη την GLSL shading language. [26 (όρος shader)]

### 1.4.5 Εισαγωγή φωτός και υλικών για τα αντικείμενα

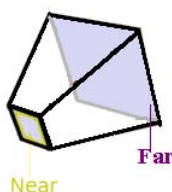
Προσθέτοντας στην σκηνή πηγές φωτός, η εικόνα αποκτά μεγαλύτερη ρεαλιστικότητα. Τα αντικείμενα μπορούν να γίνουν πιο φωτεινά ή όχι, ανάλογα με το πώς το φως πέφτει πάνω τους. Αυτό εξαρτάται από την θέση του αντικειμένου, την θέση της πηγής και την οπτική γωνία από την οποία κοιτάζει κανείς.

Πιο κοντά στην πραγματικότητα ερχόμαστε, δίνοντας και στα αντικείμενα την επιφάνεια που τους ταιριάζει. Η διαδικασία texture mapping χρησιμοποιεί διδιάστατες εικόνες (που ονομάζονται textures) και τις προσαρμόζει στα τρισδιάστατα αντικείμενα δίνοντας τους με αυτόν τον τρόπο καλύτερη εμφάνιση. [26 (όρος computer graphics)]

### 1.4.6 Ορισμός βάθους οθόνης

Το οπτικό πεδίο του υπολογιστή, συνήθως, στηρίζεται σε ένα είδος πυραμίδας, όπου η μύτη της κορυφής της έχει αφαιρεθεί. Έτσι έχουμε μια πυραμίδα με μια επίπεδη κορυφή. Επίπεδο αποκοπής, ή αλλιώς clipping plane, λέγεται η κορυφή ή η βάση αυτής της πυραμίδας. Η επίπεδη κορυφή της, που είναι παράλληλη στην οθόνη, λέγεται near clipping plane ή κοντινό επίπεδο αποκοπής, ενώ η βάση της πυραμίδας λέγεται far clipping plane ή μακρινό επίπεδο αποκοπής. Τα αντικείμενα που βρίσκονται μπροστά από το near clipping plane ή πίσω από το far clipping plane δεν εμφανίζονται στην οθόνη.

Το βάθος της σκηνής πρέπει να κυμαίνεται μεταξύ μιας φοράς του βάθους της οθόνης και 100 αυτού. [10]



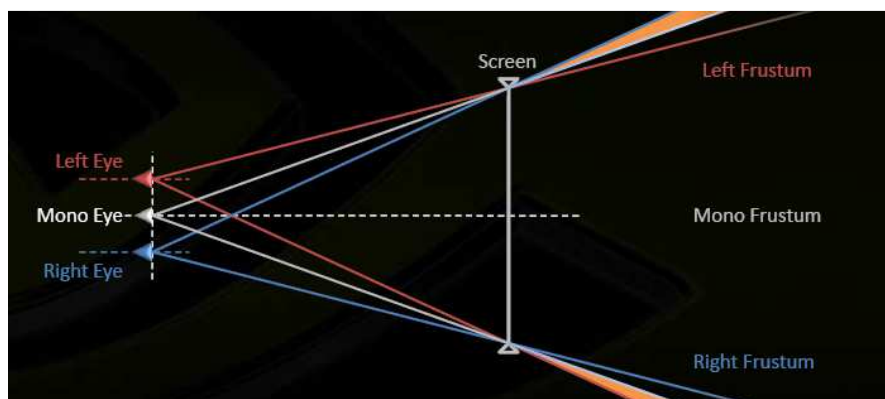
Εικόνα 12 : σχηματική απεικόνιση του κοντινού και μακρινού επιπέδου αποκοπής



### 1.4.7 Culling τρισδιάστατων αντικειμένων

Η διαδικασία culling καθορίζει τα ορατά τμήματα των τρισδιάστατων περιβαλλόντων. Βασίζεται στην απλή αρχή πως εάν κάτι δεν φαίνεται, δεν χρειάζεται να σχεδιαστεί. Αυτή η διαδικασία κάνει την εμφάνιση των αντικειμένων πιο γρήγορη και πιο αποδοτική, αφού μειώνεται ο αριθμός αυτών που πρέπει να σχεδιάσει το πρόγραμμα. Για παράδειγμα, σε μία σκηνή δρόμου δεν υπάρχει ανάγκη να σχεδιαστούν τα πολύγωνα των πλευρών των κτιρίων που κοιτούν μακριά από την κάμερα, καθώς καλύπτονται τελείως από τις πλευρές που κοιτούν την κάμερα.

Ίσως η πιο προφανής μορφή της διαδικασίας culling είναι το frustum culling, που βασίζεται στο γεγονός ότι μόνο τα αντικείμενα στο τρέχον οπτικό πεδίο πρέπει να σχεδιαστούν. Εάν ένα πολύγωνο βρίσκεται ολόκληρο έξω από την πυραμίδα (το οπτικό πεδίο του υπολογιστή), δεν μπορεί να είναι ορατό και απορρίπτεται. Εάν είναι τμηματικά μέσα, κόβεται στις ακμές της πυραμίδας έτσι ώστε τα εξωτερικά του τμήματα να αφαιρούνται.



Εικόνα 13 : Η διαδικασία culling πρέπει να επεκτείνεται οριζόντια ώστε να φαίνεται επιπλέον χώρος του δεξιού και αριστερού frustum μετά το βάθος της σκηνής.

### 1.4.8 Αντικείμενα έξω από την οθόνη

Το ανθρώπινο μυαλό δεν μπορεί να καταλάβει την ύπαρξη αντικειμένων που αιωρούνται έξω από την οθόνη. Πρέπει να γίνει επιπλέον προσπάθεια ώστε να δώσουμε στους χρήστες ένα αποτέλεσμα που να τους πείσει. Είναι αναγκαίο να βεβαιωθούμε ότι τα αντικείμενα δεν αποκόπτονται από τις άκρες της οθόνης. Η μετακίνηση των τελευταίων από το εσωτερικό της οθόνης προς την εξωτερική περιοχή πρέπει να γίνεται αργά ώστε να δοθεί χρόνος στο μάτι να προσαρμοστεί. Είναι απαραίτητο να γίνονται ομαλές οπτικές μεταβάσεις και να μην τρεμοπαίζουν οι εικόνες. [10]

## Κεφάλαιο 2

### Στερεοσκοπία (Stereoscopy)

#### 2.1 Εισαγωγή

Η λέξη στερεοσκοπία υποδηλώνει τρισδιάστατη αντίληψη των αντικειμένων. Οι στερεοσκοπικές τεχνικές εκμεταλλεύονται την διοπτρική όραση του ανθρώπου για να δημιουργήσουν την εντύπωση βάθους. Αυτό γίνεται παρουσιάζοντας δύο όψεις της ίδιας σκηνής, οι οποίες έχουν ληφθεί από ελαφρά διαφορετικές γωνίες και βασίζονται στον ανθρώπινο εγκέφαλο για να τις ενώσει σε μια τρισδιάστατη αναπαράσταση. Η στερεοσκοπία σαν ιδέα υπάρχει για περισσότερο από έναν αιώνα, αλλά μόνο τις τελευταίες δύο δεκαετίες έχει αναθερμανθεί το ενδιαφέρον για την αξιοποίησή της. Αυτό έγινε δυνατό με την ανάπτυξη τεχνικών πολυπλεξίας/ αποπολυπλεξίας εικόνας, οι οποίες χρειάζονται για να διαχωρίζονται οι όψεις που προορίζονται για κάθε μάτι, αλλά και στην ολοένα μεγαλύτερη αύξηση της υπολογιστικής ισχύος των σύγχρονων υπολογιστών.

Εφευρέθηκε αρχικά από τον Sir Charles Wheatstone το 1840. Η λέξη "stereo" προέρχεται από την ελληνική λέξη "στερεός". Με την στερεοσκοπία, επομένως, βλέπουμε ένα αντικείμενο με τρεις χωρικές διαστάσεις, πλάτος, ύψος και βάθος (ή αλλιώς  $x$ ,  $y$  και  $z$ ). Η πρόσθεση της τρίτης διάστασης, του βάθους δηλαδή, είναι που κάνει ιδιαίτερα ξεχωριστή την ικανότητα να βλέπεις στερεοσκοπικά. [26, (όρος stereoscopy)]

Υπάρχουν πολλά πλεονεκτήματα στο να βλέπουμε τον κόσμο με στερεοσκοπική αντίληψη. Με αυτόν τον τρόπο αντιλαμβανόμαστε που βρίσκονται τα αντικείμενα σε σχέση με το σώμα μας με ακόμη μεγαλύτερη ακρίβεια, ιδιαίτερα όταν αυτά κινούνται, είτε προς εμάς, είτε προς αντίθετη κατεύθυνση. Μπορούμε, επίσης, να βλέπουμε στερεά αντικείμενα από διαφορετικές πλευρές, χωρίς να κινούμε το κεφάλι μας, αλλά και να εκτιμούμε και να μετράμε πόσο ελεύθερο χώρο έχουμε.

## 2.2 Στερεοσκοπικές και αυτοστερεοσκοπικές τεχνικές

Οι τρέχουσες τεχνικές που χρησιμοποιούνται για τρισδιάστατη παρατήρηση μπορούν να διαιρεθούν σε δύο κατηγορίες, οι οποίες καλύπτουν τις στερεοσκοπικές και τις αυτοστερεοσκοπικές μεθόδους.

### 2.2.1 Αυτοστερεοσκοπικές τεχνικές

Με τις αυτοστερεοσκοπικές τεχνικές απεικονίζονται τρισδιάστατες εικόνες που μπορεί κανείς να δει χωρίς τη χρήση ειδικού εξοπλισμού ή γυαλιών από την πλευρά του χρήστη. Αυτές οι μέθοδοι δίνουν αντίληψη βάθους στο θεατή ακόμα και αν η εικόνα παράγεται από μια επίπεδη συσκευή.

Περιλαμβάνουν ειδικές οπτικές διατάξεις, οι οποίες ενσωματώνονται στη διάταξη απεικόνισης και μ' αυτό τον τρόπο εφοδιάζουν τον παρατηρητή με στερεοσκοπικά είδωλα τα οποία του προσφέρουν από μόνα τους αντίληψη βάθους. Οι τεχνικές αυτές βασίζονται στις κυλινδροειδείς συστοιχίες φακών (lenticular lenses) ή σε ειδικές συσκευές που ονομάζονται parallax barriers. Αν ο θεατής τοποθετήσει το κεφάλι του σε συγκεκριμένες θέσεις θέασης, θα αντιληφθεί μια διαφορετική εικόνα από κάθε μάτι, που δίνει την στερεοσκοπική εικόνα. Η κοινή πρακτική που ακολουθείται προς το παρόν για την εκμετάλλευση αυτών των τεχνολογιών είναι η ανάπτυξη εξειδικευμένου λογισμικού και υλικού από κάθε κατασκευαστή η οποία οδηγεί σε αύξηση του κόστους, μείωση των τρισδιάστατων θεμάτων προς απεικόνιση και έλλειψη υποστήριξης από τις υπάρχουσες εφαρμογές.

Το κυριότερο πλεονέκτημα των αυτοστερεοσκοπικών τεχνικών είναι το ότι παρέχουν στον εγκέφαλο αντίληψη πραγματικά στερεοσκοπικών ειδώλων χωρίς τους περιορισμούς που δημιουργούνται από τα γυαλιά ή από άλλα σύνεργα, πράγμα που είναι καίριας σημασίας για ορισμένες εφαρμογές. Επίσης, αυτές οι απεικονίσεις μπορεί να έχουν πολλαπλές ζώνες θέασης που επιτρέπουν σε πολλούς χρήστες να βλέπουν μία εικόνα ταυτόχρονα. Ωστόσο, η πολύωρη θέαση αυτοστερεοσκοπικών απεικονίσεων με τις παραπάνω τεχνικές προκαλεί συχνά πίεση στα μάτια και πονοκεφάλους. [4]

## 2.2.2. Στερεοσκοπικές τεχνικές

Οι στερεοσκοπικές τεχνικές περιλαμβάνουν ορισμένες διατάξεις όπως τα γυαλιά συγχρονισμένων κλειστρών, ή τα πολωτικά γυαλιά ή – πιο πρόσφατα – διατάξεις απεικόνισης με LCD σε κράνος που φοριέται στο κεφάλι (Head Mounted Display, HMD). Υπάρχει μεγάλη ποικιλία μεθόδων για την στερεοσκοπική θέαση, οι βασικότερες όμως είναι οι ενεργητικές ή active stereo και οι παθητικές ή passive stereo. [4]

### *Ενεργητικές μέθοδοι*

Χρησιμοποιούν ηλεκτρονικά γυαλιά για να δημιουργήσουν στο χρήστη την αίσθηση του τρισδιάστατου. Λειτουργούν παρουσιάζοντας τις εικόνες στον παρατηρητή σε γρήγορη διαδοχή από το δεξί στο αριστερό μάτι (συνήθως περισσότερα από 96 καρέ (frames) το δευτερόλεπτο). Ωστόσο, ο χρήστης μπορεί να βλέπει μόνο μια εικόνα κάθε φορά καθώς οι δύο φακοί εναλλάσσονται σε διαφανείς και αδιαφανείς κατά την παρουσίαση των εικόνων. Αυτό γίνεται με τη χρήση εκπομπών υπερύθρων. Κατά συνέπεια, όταν η εικόνα του αριστερού ματιού παρουσιάζεται από τον προβολέα, ο αριστερός φακός είναι διαφανής, ενώ ο δεξιός αδιαφανής.

Τα γυαλιά συγχρονισμένων κλειστρών (LCD shutter glasses) χρησιμοποιούνται σε συνδυασμό με μία οθόνη για την δημιουργία της ψευδαίσθησης μιας τρισδιάστατης εικόνας. Κάθε φακός περιέχει υγρούς κρυστάλλους και ένα φίλτρο πόλωσης που του δίνει τη δυνατότητα να γίνεται αδιαφανής όταν εφαρμόζεται τάση, διαφορετικά να γίνεται διαφανής. Η διαδικασία αυτή ελέγχεται από έναν εκπομπό υπερύθρων. Οι φακοί αλλάζουν διαδοχικά τη διαφάνεια τους σε συγχρονισμό με το ρυθμό ανανέωσης της οθόνης, με αποτέλεσμα να προκύπτει διαδοχικά διαφορετική αντίληψη για το κάθε μάτι. [26, (όρος LCD shutter glasses)]



*Εικόνα 14 : γυαλιά συγχρονισμένων κλειστρών*

Με τα γυαλιά συγχρονισμένων κλειστρών μειώνονται οι παραμορφώσεις. Ωστόσο, μπορεί να παρατηρηθεί τρεμοπαίξιμο (flicker), καθώς το κάθε μάτι αντιλαμβάνεται μόνο το μισό από τον πραγματικό ρυθμό ανανέωσης της οθόνης, εκτός αν έχουμε πολύ υψηλό ρυθμό ανανέωσης. Επίσης, επειδή κάθε φορά ο ένας φακός είναι διαφανής,

λιγότερο φως φτάνει στα μάτια του παρατηρητή με αποτέλεσμα μια αρκετά σκοτεινή εικόνα. Το μεγαλύτερο πρόβλημα, όμως, είναι η χρήση των γυαλιών, τα οποία είναι ακριβά και εύθραυστα. Πρέπει, επίσης, όλοι οι χρήστες να βρίσκονται στην εμβέλεια του εκπομπού υπερύθρων, διαφορετικά τα γυαλιά δε γνωρίζουν πότε πρέπει να αλλάζουν τους φακούς τους σε αδιαφανείς και πότε σε διαφανείς.

Η ενεργή στερεοσκοπική προβολή έχει το πλεονέκτημα του να μπορεί να εξετάζει σκηνές που δεν περιέχουν αρκετή πληροφορία για την διαδικασία stereo matching, όπως είναι οι ακμές και οι γωνίες. Παράγει έναν πολύ πυκνό χάρτη βάθους αυξάνοντας την τρισδιάστατη αντίληψη και είναι εφαρμόσιμη σε μεγάλο εύρος αντικειμένων και εργασιών.

Δυστυχώς, όμως, οι στερεοσκοπικές τεχνικές παρουσιάζουν αρκετά μειονεκτήματα. Οι μόνοι προβολείς που έχουν την δυνατότητα να δείχνουν στερεοσκοπικά είναι οι CRT projectors (αν και μπορούν και μερικοί καινούριοι LCD και DLP projectors). Το πρόβλημα με αυτούς είναι ότι δεν είναι φορητοί και δεν παρέχουν πολύ φωτεινές εικόνες. Παρατηρείται, ακόμη, το φαινόμενο του “ghosting”, το οποίο προκαλείται από το χρόνο που χρειάζεται η δέσμη ηλεκτρονίων να σαρώσει όλη την περιοχή στο κάτω μέρος της εικόνας. Επίσης, οι κάρτες γραφικών που είναι απαραίτητες για τα στερεοσκοπικά συστήματα είναι ακριβές και δεν έχουν την ίδια ποιότητα με τις μη στερεοσκοπικές κάρτες. Τέλος, όπως ήδη αναφέρθηκε, τα γυαλιά αποτελούν ακριβό και δύσχρηστο εξοπλισμό. [42]

### *Παθητικές μέθοδοι*

Για να επιτευχθεί τρισδιάστατη απεικόνιση χρησιμοποιούνται δύο προβολείς για μια απλή οθόνη, ένας για κάθε μάτι. Ένα φίλτρο τοποθετείται μπροστά από το φακό του κάθε προβολέα. Οι προβολείς παρουσιάζουν δύο εικόνες σε μια ειδική σκηνή, η οποία αντανακλά σωστά το πολωμένο φως. Έπειτα, οι χρήστες φοράνε γυαλιά πόλωσης (polarizing glasses), που επιτρέπουν να έρχεται το φως από τον έναν προβολέα, και μπορούν να βλέπουν τρισδιάστατες εικόνες. Η χρήση πολωμένου φωτός απαιτεί τη χρήση μιας ειδικής οθόνης προβολής (silver screen) ώστε να διατηρηθεί η στερεοσκοπική αντίληψη.

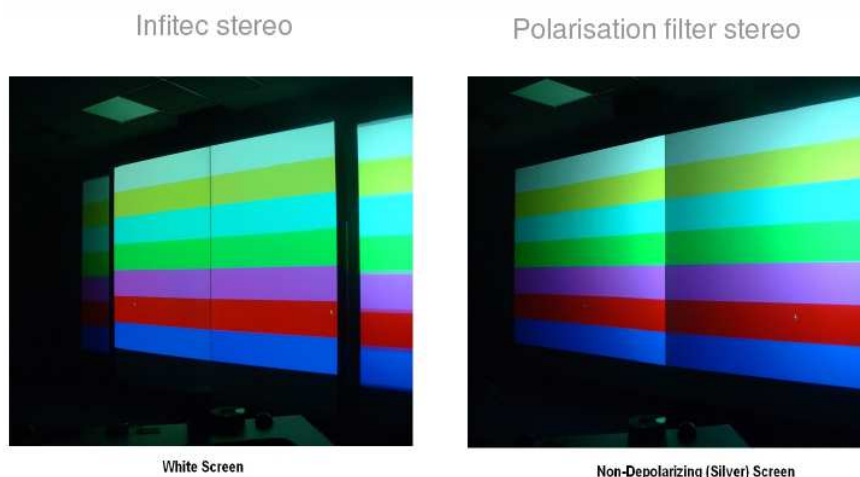
Μπορούμε να χρησιμοποιήσουμε δύο ειδών φίλτρα για την πόλωση του φωτός των προβολέων, το γραμμικό και το κυκλικό. Τα γραμμικά φίλτρα είναι πιο συνηθισμένα, αν και όταν χρησιμοποιούνται οι χρήστες δε πρέπει να κινούν το κεφάλι τους, διαφορετικά θα μειωθεί η στερεοσκοπική αντίληψη, κάτι που δε συμβαίνει με τα κυκλικά φίλτρα.

Το κυριότερο πλεονέκτημα των παθητικών μεθόδων είναι το χαμηλό τους κόστος. Μπορούμε να χρησιμοποιήσουμε LCD ή DLP προβολείς που προσφέρουν μεγαλύτερη φωτεινότητα. Αντιθέτως, στις ενεργητικές στερεοσκοπικές τεχνικές απαιτούνται υψηλές συχνότητες, συνήθως μεγαλύτερες από 96 Hz, ώστε να μειωθεί το φαινόμενο του flickering, περιορίζοντας μας στην χρήση προβολέων CRT ή των νέων και πολύ ακριβών προβολέων DLP. Ένα άλλο πλεονέκτημα είναι ότι τα γυαλιά που χρησιμοποιούνται στα παθητικά στερεοσκοπικά συστήματα είναι πολύ πιο φτηνά και λιγότερο εύθραυστα, άρα τα συστήματα αυτά είναι κατάλληλα για χρήση σε αίθουσα. Από την άλλη πλευρά, αν και φτηνά, τα πλαστικά γραμμικά φίλτρα τελικά στραβώνουν από τη θερμότητα των προβολέων. Λύση σε αυτό το πρόβλημα είναι να τοποθετήσουμε το φίλτρο σε γυαλί μέσα στον προβολέα κάτι που συμβάλλει και στο να διατηρείται καθαρό. Είναι όμως ακριβό. Ακόμη και αν επιλέξουμε τα κυκλικά φίλτρα, υπάρχει κάποιο γραμμικό συστατικό, και η διαρρύθμιση τους είναι αρκετά δύσκολη. Από την άλλη αν χρησιμοποιήσουμε γραμμικά φίλτρα, οι χρήστες θα πρέπει να διατηρούν το κεφάλι τους σταθερό ώστε να έχουν τρισδιάστατο αποτέλεσμα. Αυτό βέβαια δεν είναι και τόσο κακό σε μία αίθουσα, καθώς οι μαθητές θα αναγκαστούν να μείνουν προσεχτικοί. Τέλος, η οθόνη που χρησιμοποιείται για την προβολή παθητικών στερεοσκοπικών εικόνων, πρέπει να είναι ειδικά πολωμένη (υπάρχουν διάφορες επιλογές και τιμές). Εναλλακτικά, θα μπορούσαμε να αγοράσουμε μια Z-screen η οποία αποτελεί ένα φίλτρο που μπαίνει σε προβολέα CRT και μετατρέπει την ενεργητική στερεοσκοπική προβολή σε παθητική. [42][43]

### *Συμπληρωματικές μέθοδοι*

#### *Infitec τεχνολογία*

Επιπρόσθετα των μεθόδων πόλωσης, η τεχνολογία Infitec χρησιμοποιεί μια διαφορετική μέθοδο διαχωρισμού των αντιλήψεων των ματιών, με τη βοήθεια φασματικών δεσμών χρωμάτων, που είναι διαφορετικές για κάθε μάτι, μέσω των φίλτρων των προβολέων και των γυαλιών. Επιτυγχάνεται τέλειος διαχωρισμός, απουσία flickering, δεν απαιτείται συγχρονισμός και πρόκειται για φθινό εξοπλισμό ανεξάρτητο silver οθόνης. Ωστόσο, μπορεί να προκύψουν λάθη χρωμάτων σε περίπτωση που οι προβολείς δεν έχουν ρυθμιστεί σωστά, οπότε πρέπει να δώσουμε ιδιαίτερη βαρύτητα στην διαρρύθμιση και στην ευθυγράμμιση των προβολέων. [12]



Εικόνα 15 : διαφορά στην απεικόνιση με τεχνολογία Infitec και τεχνολογία φίλτρων πόλωσης

### Μέθοδος Anaglyph

Μια μικρή παραλλαγή της παθητικής στερεοσκοπικής μεθόδου είναι η μέθοδος Anaglyph η οποία βασίζεται σε οπτικές διαδικασίες του μυαλού για την ένωση των δύο διασταυρωμένων εικόνων σε μία στερεοσκοπική. Παρουσιάζει ομοιότητες με την πόλωση, αλλά αντί οι εικόνες να περνούν μέσα από φίλτρα πόλωσης, το πραγματικό χρώμα των εικόνων ρυθμίζεται ελαφρώς και έπειτα ο θεατής φορά γυαλιά με φίλτρα ταιριάσματος χρώματος, που είναι συνήθως τα κόκκινα-μπλε γυαλιά. Η θέαση των ανάγλυφων μέσω των κατάλληλα χρωματισμένων γυαλιών έχει σαν αποτέλεσμα το κάθε μάτι να βλέπει μια ελαφρώς διαφορετική εικόνα. Για παράδειγμα, σε ένα ανάγλυφο ρυθμισμένο με τη μέθοδο κόκκινο-μπλε, το μάτι που καλύπτεται από το κόκκινο φίλτρο βλέπει τα κόκκινα μέρη της εικόνας σαν "λευκά" και τα μπλε μέρη της εικόνας σαν "μαύρα" (με τον εγκέφαλο να παρέχει κάποια προσαρμογή για το χρώμα). Αντίστοιχα, το μάτι που καλύπτεται από το μπλε φίλτρο βλέπει τα μπλε μέρη της εικόνας σαν "λευκά" και τα κόκκινα σαν "μαύρα". Οι πραγματικά άσπρες ή μαύρες περιοχές γίνονται αντιληπτές το ίδιο και από τα δύο μάτια. Ο εγκέφαλος συνδυάζει τις εικόνες που λαμβάνει από κάθε μάτι, και ερμηνεύει τις διαφορές τους σαν το αποτέλεσμα των διαφορετικών αποστάσεων. [16]



Εικόνα 16 : γυαλιά με φίλτρα ταιριάσματος χρώματος κόκκινο-μπλε

Οι μέθοδοι αυτές δίνουν αποτελέσματα με χαμηλότερη ποιότητα από αυτά των μεθόδων Infitec ή polarised, αλλά υπάρχει μόνο μία εικόνα προβολής και πολύ φθηνά γυαλιά. Υπάρχουν αρκετές Anaglyph μέθοδοι, όπως η Colorcode που στοχεύει στην επίλυση των προβλημάτων ποιότητας.

Η μέθοδος Colorcode 3-D (κίτρινο- μπλε) είναι μια νέα ευρεσιτεχνία που έχει σαν στόχο την βελτίωση του χρώματος και του βάθους της απεικόνισης σε σχέση με τις πιο παλιές μεθόδους Anaglyph κόκκινο/ μπλε (red/ cyan) και κόκκινο/ πράσινο(red/ green). Χρησιμοποιεί τα χρώματα κίτρινο και βαθύ μπλε για την οθόνη και τα χρώματα του κεχριμπαριού και του βαθύ μπλε για τους φακούς των γυαλιών. Ο φακός με το χρώμα του κεχριμπαριού επιτρέπει να περάσει στο μάτι λίγο από το κάθε χρώμα του φάσματος προκαλώντας μεγάλη σκίαση όταν έχουμε μικρό εύρος όρασης. Χωρίς την χρήση φίλτρων οι εικόνες εμφανίζουν μπλε και κίτρινα περιθώρια αλλά έχουν καλύτερη ποιότητα χρωμάτων σε σχέση με τα αποτελέσματα που δίνουν οι μέθοδοι Anaglyph κόκκινο/ πράσινο χωρίς φίλτρα. Παρ' όλα αυτά, η μέθοδος αναπτύσσει διαδικασίες για την μείωση εμφάνισης περιθωρίων σε μεγάλα βάθη. [26, (όρος stereoscopy)]



Εικόνα 17 : γυαλιά με φίλτρα ταιριάσματος χρώματος κεχριμπάρι-μπλε

## 2.3 Οι βασικές έννοιες της στερεοσκοπίας

### 2.3.1 Τρισδιάστατη προβολή σε ένα επίπεδο

Ένα βασικό στοιχείο της οπτικής αντίληψης του χρήστη, έχει να κάνει με την οπτική αντίληψη του βάθους. Η αντίληψη του βάθους είναι η φυσική ικανότητα του ανθρώπου να αντιλαμβάνεται τον κόσμο στις τρεις διαστάσεις και τις σχέσεις μεταξύ των αντικειμένων στο χώρο. Το ανθρώπινο οπτικό σύστημα αντιλαμβάνεται το οπτικό βάθος μέσω φυσιολογικών και ψυχολογικών στοιχείων. Όσον αφορά τα φυσιολογικά στοιχεία, σε ορισμένες περιπτώσεις απαιτείται η χρήση και των δυο ματιών



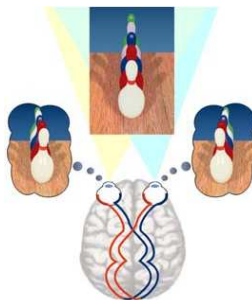
(διοπτρικός) και σε ορισμένες η χρήση του ενός ματιού (μονοσκοπικός). Στα ψυχολογικά στοιχεία, όλες οι περιπτώσεις είναι μονοφθαλμικού χαρακτήρα.

Στον πραγματικό κόσμο το ανθρώπινο μάτι χρησιμοποιεί όλα τα διαθέσιμα στοιχεία οπτικού βάθους για να προσδιορίσει τις αποστάσεις και τις σχέσεις μεταξύ των αντικειμένων. Για να έχουμε διαθέσιμα όλα αυτά τα στοιχεία και στον εικονικό κόσμο, απαιτείται η παροχή κάποιου είδους στερεοσκοπικής εικόνας.

Δεδομένου ότι το βασικό χαρακτηριστικό της τρισδιάστατης προβολής είναι η αίσθηση του βάθους που οφείλει να προκαλεί στον θεατή, υπάρχουν δύο βασικές τεχνικές προσομοίωσης, που χρησιμοποιούνται για την απόδοση των τριών διαστάσεων. Στα μη στερεοσκοπικά συστήματα, όπου η προσομοίωση των τριών διαστάσεων γίνεται αξιοποιώντας ιδιότητες του αντικειμένου της όρασης (χρωματισμοί, υφές, σκιάσεις, σχετικές θέσεις κ.α.), ο χρήστης βλέπει μια δισδιάστατη εικόνα και επιφορτίζεται με το νοητικό έργο της “μετάφρασής” της σε τρεις διαστάσεις. Στα στερεοσκοπικά συστήματα, μέσω ειδικών γυαλιών, προβάλλεται σε κάθε οφθαλμό διαφορετική οπτική γωνία της εικόνας και αξιοποιείται η διοφθαλμική διαφορά (binocular disparity) που αποτελεί το βασικό στοιχείο από το οποίο το νευροφυσιολογικό σύστημα όρασης του ανθρώπου παράγει την αντίληψη του βάθους (3<sup>η</sup> διάσταση).

Η έρευνα γύρω από την στερεοσκοπική όραση ή την “αίσθηση του βάθους” σε νευρωνικό επίπεδο έχει καταδείξει την ύπαρξη νευρώνων εξειδικευμένων να αποκρίνονται επιλεκτικά στις διαφορές λόγω της διοφθαλμικής όρασης. Οι άνθρωποι έχουν 2 μάτια, το ένα δίπλα στο άλλο, μπροστά στο κεφάλι τους. Χάρη στην κοντινή τους αυτή θέση, κάθε μάτι βλέπει την ίδια περιοχή από μια ελαφρά διαφορετική γωνία. Οι δύο αυτές όψεις έχουν αρκετά κοινά σημεία, αλλά κάθε μάτι προσφέρει κάποιες οπτικές πληροφορίες που δεν προσφέρει το άλλο.

Κάθε μάτι έχει την δική του οπτική αντίληψη για το αντικείμενο που βλέπει, και οι δύο διαφορετικές αντιλήψεις στέλνονται στον εγκέφαλο για επεξεργασία. Όταν αυτές φτάνουν ταυτόχρονα στο πίσω μέρος του, ενώνονται σε μια εικόνα. Ο εγκέφαλος συνδυάζει τις δύο αντιλήψεις σε μία ανακαλύπτοντας τις ομοιότητες και προσθέτοντας τις μικρές διαφορές. Το συνδυασμένο αποτέλεσμα είναι μία στερεοσκοπική εικόνα τριών διαστάσεων.

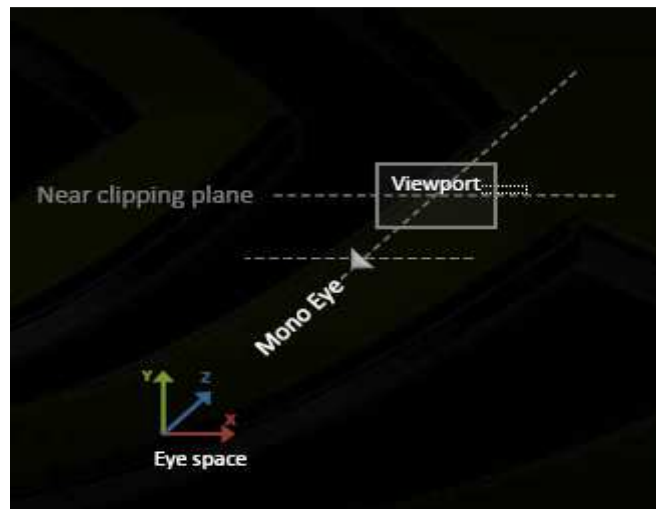


*Εικόνα 18* : οι αντιλήψεις του αντικειμένου από κάθε μάτι στέλνονται στον εγκέφαλο και προκύπτει η τελική εικόνα

Η στερεοσκοπία, εκμεταλλεύομενη την ιδιότητα αυτή του ανθρώπινου εγκεφάλου, χρησιμοποιεί δύο διαφορετικές οπτικές γωνίες για να αποτυπώσει μία εικόνα και να παράγει δύο σκηνές, μία για το δεξί και μία για το αριστερό μάτι. Έπειτα, η σκηνή για το αριστερό μάτι προβάλλεται στα ζυγά πλαίσια (0, 2, 4, κ.τ.λ.) και η σκηνή για το δεξί προβάλλεται στα μονά πλαίσια (1, 3, 5, κ.τ.λ.) με γρήγορο ρυθμό. Το αποτέλεσμα που προκύπτει για τον τελικό θεατή, ο οποίος φορά ειδικά γυαλιά, τα οποία συσκοτίζουν το μάτι που δεν πρέπει να βλέπει την εικόνα για κάθε πλαίσιο που προβάλλεται, είναι μια συνδυασμένη εικόνα που φαίνεται να έχει βάθος. [25]

### 2.3.2 Βασική μονοσκοπία

Στην μονοσκοπική προβολή γίνεται χρήση ενός μόνο καναλιού απεικόνισης. Τα στοιχεία βάθους των μονοσκοπικών εικόνων (monoscopic image depth cues) είναι στοιχεία, τα οποία αναγνωρίζονται σε στατικές οπτικές απόψεις μιας σκηνής, όπως είναι για παράδειγμα οι φωτογραφίες και οι πίνακες ζωγραφικής. Τα στοιχεία αυτά συνήθως προέρχονται από την κατάσταση της σκηνής όπως είναι: η παρεμβολή, η σκίαση, το μέγεθος, η γραμμική προοπτική, οι βαθμικές (gradient) υφές επιφανειών, το ύψος του οπτικού πεδίου, οι ατμοσφαιρικές επιδράσεις και η φωτεινότητα. [10]

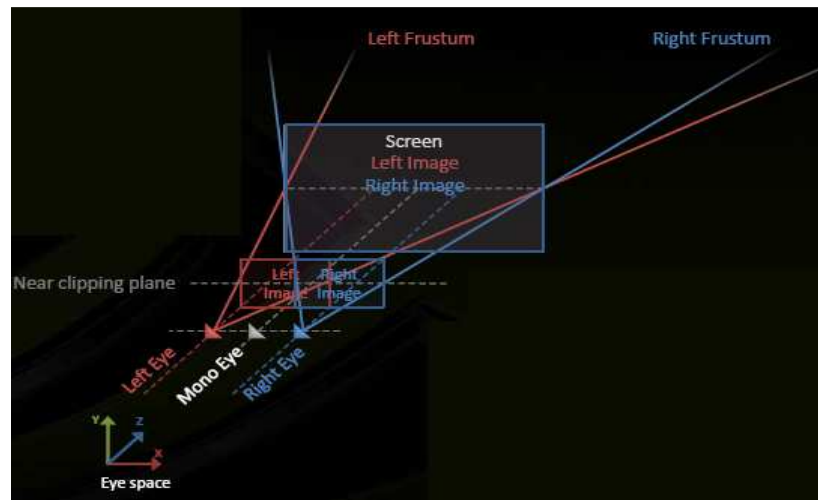


Εικόνα 19 : Βλέπουμε την σκηνή από ένα μάτι (mono eye), αναλύοντας την με μονοσκοπικό τρόπο (σημείο mono), και αυτή προβάλλεται στο κοντινό επίπεδο αποκοπής (Near Clipping plane).

### 2.3.3 Δύο μάτια, μία οθόνη, δύο εικόνες

Επειδή, όπως αναφέρθηκε και προηγουμένως, ο άνθρωπος έχει δύο μάτια, θα πρέπει να μπορούμε να βρούμε τη θέση των δύο ματιών σε σχέση με το ένα μάτι (mono eye) της μονοσκοπικής προβολής. Η θέση του αριστερού ματιού (left eye), λοιπόν, βρίσκεται εάν κάνουμε ολίσθηση της θέσης του ενός ματιού (mono eye), στον άξονα των x, προς τα αριστερά και η θέση του δεξιού ματιού (right eye) βρίσκεται εάν κάνουμε ολίσθηση της θέσης του στον άξονα των x, προς τα δεξιά.

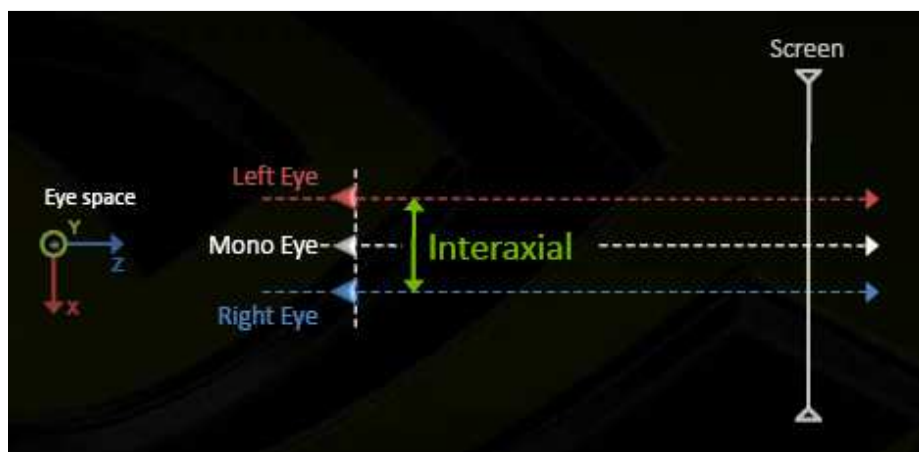
Έτσι, δημιουργούνται δύο οπτικά πεδία (frustums), ένα από το αριστερό μάτι και ένα από το δεξί, τα οποία συγκλίνουν και δημιουργούν την εικονική οθόνη (screen). Παράγονται με αυτόν τον τρόπο δύο πρότυπα στο κοντινό επίπεδο αποκοπής, σύμφωνα με το κάθε μάτι. Στη συνέχεια, παρουσιάζονται στον χρήστη συνδυασμένα και ανεξάρτητα από το κάθε μάτι στην πραγματική οθόνη, δίνοντας έτσι την αίσθηση του βάθους. [10]



Εικόνα 20 : τα δύο οπτικά πεδία από το αριστερό και το δεξί μάτι συγκλίνουν και δημιουργούν την εικονική οθόνη

### 2.3.4 Interaxial

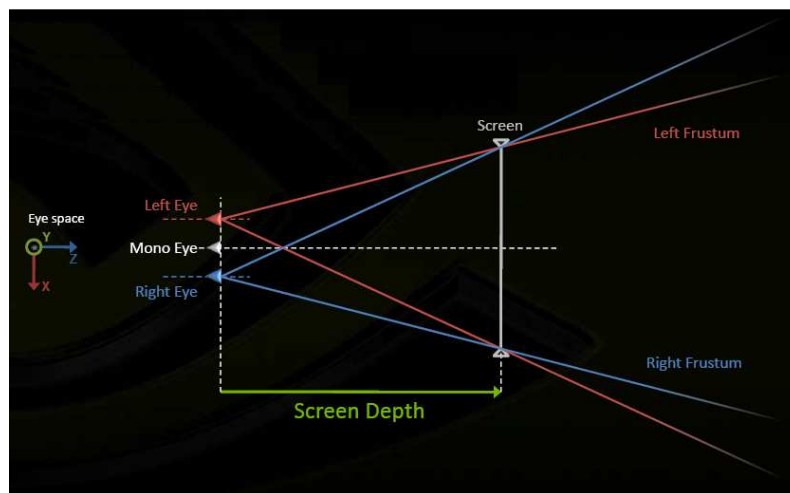
Interaxial ονομάζεται η απόσταση μεταξύ των δύο ματιών (αριστερό και δεξί) στον χώρο του ματιού (eye space). Έτσι, για να αναπαράγουμε μία στερεοσκοπική εικόνα, θα πρέπει η απόσταση από το κέντρο της αριστερής κάμερας ως το κέντρο της δεξιάς κάμερας να είναι ίση με την απόσταση interaxial, και περίπου ίση με 6.35 εκατοστά, τη μέση απόσταση μεταξύ των ανθρώπινων ματιών. Οι κατευθύνσεις του αριστερού, του δεξιού και του mono ματιού είναι παράλληλες, όπως δείχνει και η εικόνα. [10]



Εικόνα 21 : απεικόνιση του interaxial

### 2.3.5 Βάθος οθόνης (Screen depth)

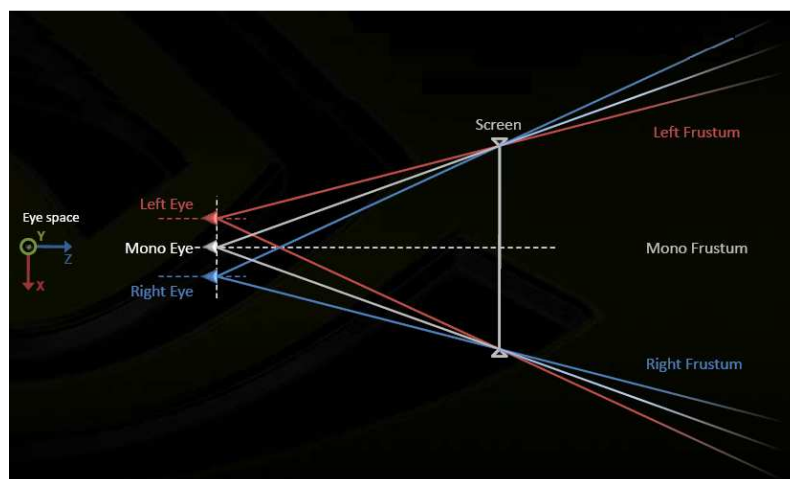
Για να μπορούμε να έχουμε τρισδιάστατη αντίληψη χρειάζεται να δουλεύουμε με τρεις διαστάσεις. Η οθόνη είναι επίπεδη και έχει πλάτος και ύψος. Μιλώντας για βάθος, αναφερόμαστε στην τρίτη διάσταση, μέσα και έξω από την οθόνη. Λέγεται και σύγκλιση. Ουσιαστικά, είναι το εικονικό βάθος της οθόνης στο χώρο του ματιού. Θεωρείται επίσης το επίπεδο στο οποίο τέμνεται το αριστερό και δεξιό frustum. [10]



Εικόνα 22: απεικόνιση του screen depth

### 2.3.6 Αριστερή / Δεξιά προβολή

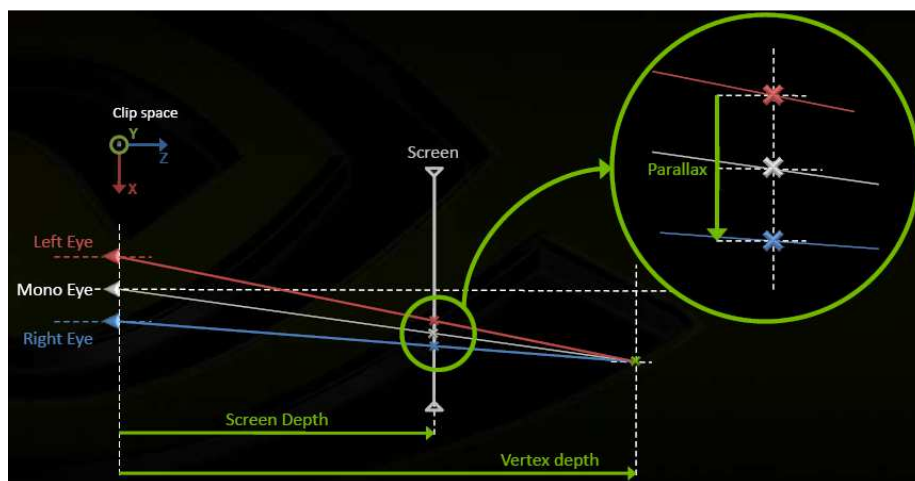
Ο πίνακας προβολής για κάθε μάτι είναι μια οριζοντίως τροποποιημένη εκδοχή του μονοσκοπικού πίνακα προβολής. Προκύπτει από την ολίσθηση της x συντεταγμένης δεξιά ή αριστερά. [10]



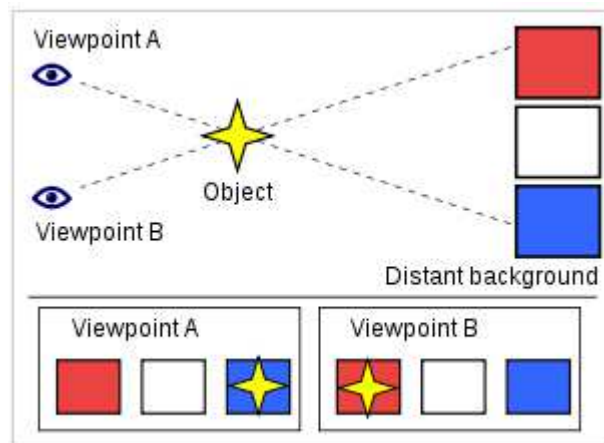
Εικόνα 23 : αριστερή και δεξιά προβολή

### 2.3.7 Παράλλαξη

Ο όρος προέρχεται από την ελληνική λέξη παράλλαξις που σημαίνει αλλαγή. Ονομάζεται η φαινομενική μετατόπιση της θέσης ενός αντικειμένου, ως προς ένα σύστημα αναφοράς, η οποία προκαλείται από μετατόπιση της θέσης παρατήρησής του. Αναλύεται σε δύο συνιστώσες, την x-παράλλαξη,  $P_x$ , και την y-παράλλαξη,  $P_y$ . Η x-παράλλαξη δίνεται από τη διαφορά των συντεταγμένων του ίδιου σημείου ως προς τον άξονα των x. Αντίστοιχα η y-παράλλαξη δίνεται από τη διαφορά των συντεταγμένων του ίδιου σημείου ως προς τον άξονα των y. Τα κοντινά αντικείμενα έχουν μεγαλύτερο parallax από τα πιο απομακρυσμένα αντικείμενα όταν παρατηρούνται από διαφορετικές θέσεις, οπότε το parallax μπορεί να χρησιμοποιηθεί για να καθορίσει τις αποστάσεις. Η παράλλαξη θεωρείται ως η διαδικασία για την δημιουργία βάθους στον χώρο του ματιού. [10][15]



Εικόνα 24 : παράλλαξη : προσημασμένη απόσταση στην οθόνη ανάμεσα στις προβολές ενός διανύσματος με βάση την αριστερή και δεξιά εικόνα



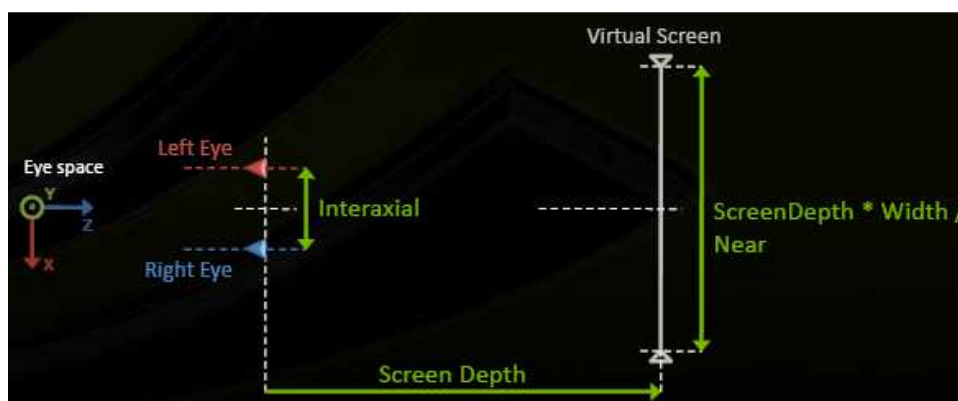
Εικόνα 25 : Όταν παρατηρούμε το αντικείμενο (αστέρι) από το σημείο A, "Viewpoint A", το μάτι μας βρίσκεται μπροστά από το κόκκινο τετράγωνο αλλά το αντικείμενο φαίνεται να είναι μπροστά από το μπλε τετράγωνο. Ενώ όταν αλλάζει το σημείο παρατήρησης, "Viewpoint B", το μάτι μας είναι μπροστά από το μπλε τετράγωνο και το αντικείμενο φαίνεται να είναι μπροστά από το κόκκινο τετράγωνο.

### 2.3.8 Παράγοντας βάθους (Depth factor)

Στον εικονικό χώρο ο στερεοσκοπικός διαχωρισμός βασίζεται στο ποσοστό  $Interaxial / VirtualScreenWidth$  (πλάτος εικονικής σκηνής). Το πλάτος εικονικής σκηνής υπολογίζεται από τον τύπο

$$VirtualScreenWidth = ScreenDepth \text{ (Βάθος σκηνής)} * Width \text{ (πλάτος)} / Near.$$

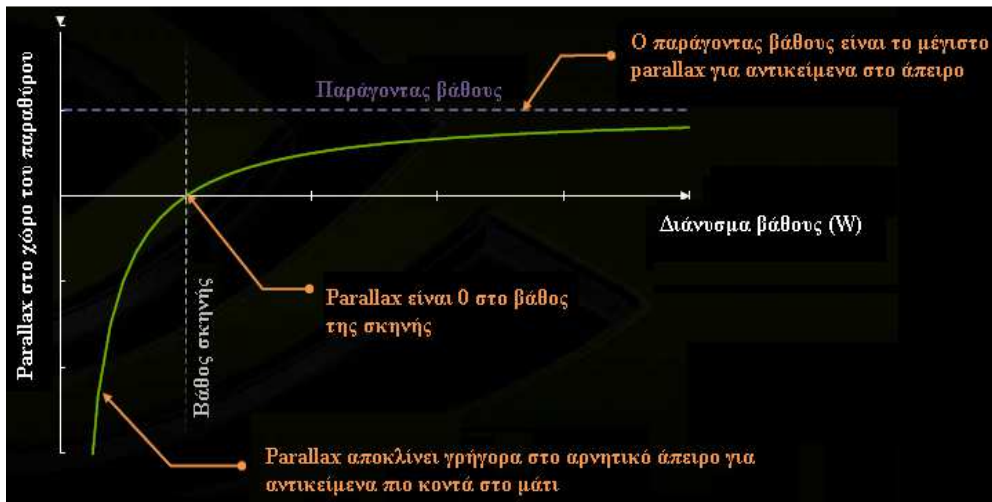
Μπορεί να γίνει προσαρμογή της αντίληψης του βάθους (depth perception) με την χρήση ενός παράγοντα κλίμακας (depth scale). [10]



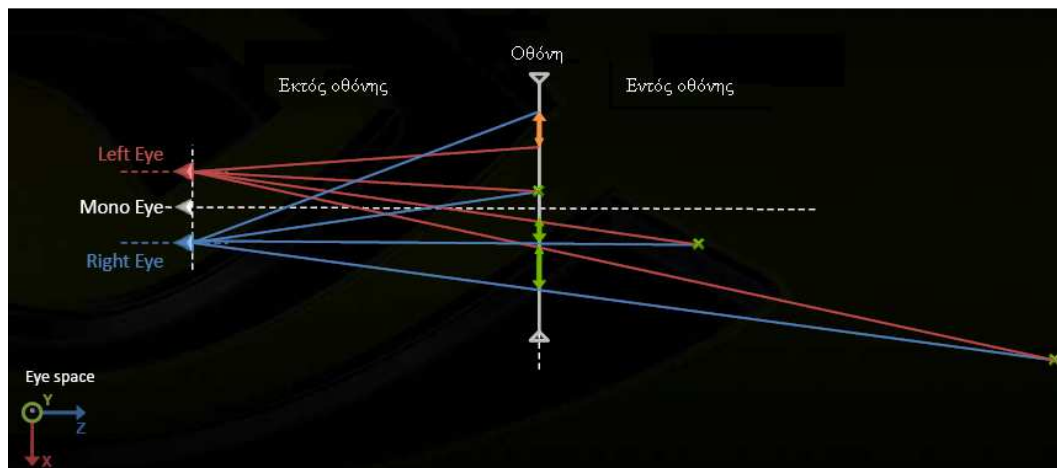
Εικόνα 26 :  $DepthFactor = DepthScale * Interaxial * Near / (Width * ScreenDepth)$

### 2.3.9 Μέσα/ Έξω από την οθόνη

Σε κανονικοποιημένο χώρο παραθύρου και μετά τον διαχωρισμό της προοπτικής προβολής (perspective division) με το διάνυσμα βάθους (w), το parallax δίνεται από :

$$\text{parallax} = \text{παράγοντας βάθους} * (1 - \text{βάθος σκηνής} / W). [10]$$


Εικόνα 27 : απεικόνιση παράλαξης σε σχέση με το διάνυσμα βάθους



Εικόνα 28 : Η αντίληψη βάθους δημιουργείται από το parallax. Όταν είναι αρνητικό βρισκόμαστε εκτός οθόνης



## Κεφάλαιο 3

### Συσκευές εισόδου

#### 3.1 Εισαγωγή

Η είσοδος των δεδομένων σε ένα διαδραστικό σύστημα είναι πολύ σημαντική και έχει να κάνει τόσο με την ευχρηστία, όσο με την λειτουργικότητα και την απόδοση του συστήματος. Είναι πολύ σημαντικό ο υπολογιστής να δέχεται εύκολα, γρήγορα και σωστά τα δεδομένα.

Μία συσκευή εισόδου είναι κάθε περιφερειακό (τμήμα εξοπλισμού του υλικού ενός υπολογιστή), που χρησιμοποιείται για να παρέχει δεδομένα και σήματα ελέγχου σε έναν υπολογιστή, ή σε ένα σύστημα επεξεργασίας πληροφορίας γενικότερα. [26 (όρος input device)] Μια συσκευή εισόδου είναι απλά κάτι που χρησιμοποιείται για να γίνει αλληλεπίδραση με τον υπολογιστή, αντίθετα με μια συσκευή εξόδου που εμφανίζει τα δεδομένα του χρήστη. Καθένας που έχει ασχοληθεί, έστω και μια φορά, με έναν υπολογιστή έχει χρησιμοποιήσει επίσης μια συσκευή εισόδου, όπως το πληκτρολόγιο ή το ποντίκι. Χωρίς συσκευές εισόδου, δεν θα υπήρχε κανένας τρόπος αλληλεπίδρασης με τους υπολογιστές. Οι πιο συνηθισμένες συσκευές εισόδου είναι το ποντίκι και το πληκτρολόγιο. Ωστόσο, joysticks, θόνες αφής, webcams και μικρόφωνα μπορούν επίσης να αποτελέσουν συσκευές εισόδου. Κάποιες πιο “εξωτικές” συσκευές εισόδου που οι περισσότεροι άνθρωποι δεν χρησιμοποιούν σε καθημερινή βάση είναι οι συσκευές τρισδιάστατης πλοήγησης. Παραδείγματα αυτού του τύπου συσκευών εισόδου, είναι τα ποντίκια που μπορούν να δείχνουν στον τρισδιάστατο χώρο, ή αναγνώστες διατρητών καρτών που χρησιμοποιούνταν τα πρώτα χρόνια. [28]

Οι συσκευές εισόδου μπορούν να ταξινομηθούν ανάλογα με τον τρόπο εισόδου (μηχανική κίνηση, ήχος, εικόνα κλπ.), το αν η είσοδος είναι διακριτή (π.χ. πάτημα πλήκτρων) ή συνεχής (π.χ. η θέση του ποντικιού, το οποίο αν και είναι ψηφιοποιημένο σε μια διακριτή ποσότητα, είναι αρκετά γρήγορο για να θεωρείται συνεχές) και τον αριθμό των βαθμών ελευθερίας που εμπλέκονται (π.χ. παραδοσιακά ποντίκια δύο διαστάσεων, ή τρισδιάστατοι πλοηγοί σχεδιασμένοι για εφαρμογές CAD).

Υπάρχουν και συσκευές εισόδου που χρησιμοποιούνται για τον προσδιορισμό της θέσης στο χώρο και ονομάζονται pointing devices (όπως είναι το ποντίκι και το joystick). Μπορούν να ταξινομηθούν ανάλογα με το αν η είσοδος είναι άμεση ή έμμεση. Με την άμεση είσοδο, ο χώρος εισόδου συμπίπτει με το χώρο απεικόνισης. Οι θόνες

αφής αποτελούν άμεση είσοδο, ενώ παράδειγμα έμμεσης εισόδου είναι το ποντίκι. Άλλη μια κατηγορία ταξινόμησης προκύπτει από το κατά πόσο η πληροφορία θέσης είναι απόλυτη (π.χ. σε μια οθόνη αφής) ή σχετική (π.χ. με ένα ποντίκι που μπορεί να αλλάζει θέσεις όποτε ο χρήστης το επιθυμεί). [26 (όρος input device)]

### 3.2 Πληκτρολόγιο (Keyboard)

Το πληκτρολόγιο είναι η πιο δημοφιλής συσκευή εισόδου σε έναν ηλεκτρονικό υπολογιστή. Μεταφράζει αριθμούς, γράμματα, σύμβολα και κλειδιά ελέγχου σε σήματα που μπορούν να αναγνωριστούν από τον υπολογιστή. Αναπαρίσταται σαν μια σειρά από πλήκτρα. Κάθε πλήκτρο μπορεί να χρησιμοποιηθεί είτε για να εισάγει κάποιο αριθμό ή γράμμα του αλφαβήτου, είτε για να εκτελέσει μια συγκεκριμένη λειτουργία του υπολογιστή. Τα περισσότερα αγγλικά πληκτρολόγια βασίζονται στον σχεδιασμό QWERTY. Ο σχεδιασμός αυτός πήρε το όνομα του από τα έξι πρώτα γράμματα που εμφανίζονται στην πρώτη σειρά γραμμάτων του πληκτρολογίου. Ένας εναλλακτικός σχεδιασμός για το πληκτρολόγιο είναι ο DVORAK, ο οποίος αναπτύχθηκε για να προσφέρει μία πιο αποτελεσματική μέθοδο πληκτρολόγησης, καθώς τα φωνήεντα είναι συγκεντρωμένα στο ίδιο μέρος. Παρόλα αυτά δεν έγινε ευρέως γνωστός. [27]

Η βασική χρήση του πληκτρολογίου είναι η εγγραφή κειμένου και αριθμών σε έναν επεξεργαστή κειμένου ή κάποιο άλλο σχετικό πρόγραμμα. Στους σύγχρονους υπολογιστές, η αναγνώριση των πλήκτρων που πατιούνται είναι ευθύνη του λογισμικού. Χρησιμοποιείται, επίσης, σε διάφορα ηλεκτρονικά παιχνίδια είτε ως ένα συνηθισμένο πληκτρολόγιο, είτε σαν ένα πληκτρολόγιο με ειδικά χαρακτηριστικά για ενέργειες που ενεργοποιούνται με συνδυασμούς πλήκτρων. Τέλος, μπορεί να δώσει εντολές στο λειτουργικό σύστημα, όπως ο συνδυασμός Control - Alt - Delete που εμφανίζει ένα παράθυρο διαχείρισης εργασιών.

Παρά την ανάπτυξη εναλλακτικών συσκευών εισόδου, όπως το ποντίκι, οθόνες αφής, αναγνώριση χαρακτήρων και φωνής το πληκτρολόγιο παραμένει η πιο ευρέως χρησιμοποιούμενη και ευπροσάρμοστη συσκευή άμεσης εισόδου. [26 (όρος keyboard)]



Εικόνα 29 : Ένα τυπικό πληκτρολόγιο

### 3.3 Ποντίκι (Mouse)

Εκτός από το πληκτρολόγιο, το ποντίκι είναι ένας εξίσου συνηθισμένος τύπος συσκευής εισόδου. Σχεδόν κάθε υπολογιστής διαθέτει ένα ποντίκι το οποίο κάνει την πλοήγηση στην οθόνη πολύ πιο εύκολη από την χρήση μόνο του πληκτρολογίου. [28] Όταν μετακινείται το ποντίκι, κινείται ανάλογα στην οθόνη ένα σύμβολο που έχει συνήθως τη μορφή βέλους και λέγεται δείκτης ή δρομέας (cursor). Ανήκει στις συσκευές εισόδου pointing και λειτουργεί ανιχνεύοντας την δισδιάστατη κίνηση στην οθόνη. Αποτελείται από ένα αντικείμενο (στο σχήμα του ποντικιού) που ο χρήστης χειρίζεται με το χέρι του, με ένα ή περισσότερα κουμπιά. Μπορεί να διαθέτει και άλλα στοιχεία, όπως “τροχούς”, που επιτρέπουν στον χρήστη να εκτελέσει διάφορες λειτουργίες ή επιπλέον κουμπιά ή χαρακτηριστικά που παρέχουν την δυνατότητα περισσότερου ελέγχου. Η κίνηση του ποντικιού μεταφράζεται σε κίνηση του κέρσορα στην οθόνη, το οποίο επιτρέπει τον ακριβή έλεγχο μια γραφικής διεπαφής χρήστη. [26 (όρος mouse)] Συνήθως χρησιμοποιούν μια “μπίλια”, φως, ή laser για να ανιχνεύσουν την κίνηση. Για να εκτελεστεί κάποια λειτουργία με το ποντίκι, θα πρέπει να πατηθεί κάποιο από τα πλήκτρα ή το κεντρικό κουμπί. Στις περισσότερες εφαρμογές χρησιμοποιείται κυρίως το αριστερό πλήκτρο του ποντικιού, το οποίο πατιέται με το δείκτη του δεξιού χεριού. Κατά την εγκατάσταση του ποντικιού στον υπολογιστή μπορεί να καθοριστεί η χρήση του ποντικιού από αριστερόχειρα. Σ' αυτή την περίπτωση κύριο πλήκτρο είναι το δεξί, το οποίο πατιέται με το δείκτη του αριστερού χεριού. [28]



*Εικόνα 30* : Ένα τυπικό ποντίκι με τα τρία κοινά χαρακτηριστικά του : δύο κουμπιά και μια ροδέλα κύλισης

### 3.3.1 Περιγραφή και λειτουργία

Το κλασικό ποντίκι είναι χειροκίνητη συσκευή κατάδειξης, τα κύρια μέρη του οποίου είναι το καλώδιο επικοινωνίας με τον Η/Υ και το κυρίως σώμα της συσκευής. Το κυρίως σώμα αποτελείται συνήθως από ένα σταθερό στέλεχος από πλαστικό, πάνω στο οποίο είναι προσαρμοσμένα δύο ή περισσότερα κουμπιά και ίσως μια ροδέλα κύλισης.

#### *Μηχανικό ποντίκι*

Στο κάτω μέρος του στελέχους υπάρχει μια μικρή σφαίρα από εύκαμπτο υλικό, η οποία μπορεί να κυλάει ελεύθερα. Κινώντας το ποντίκι πάνω σε μια σχετικά λεία επιφάνεια, η μπίλια μεταφέρει την κίνηση αυτή σε δύο κάθετες μεταξύ τους πλαστικές ράβδους στο εσωτερικό του κυρίου σώματος, οι οποίες έχουν προσαρμοσμένα στα άκρα τους δύο γρανάζια ή διάτρητους πλαστικούς δίσκους. Η κίνηση των δίσκων διακόπτει μια δέσμη φωτός από ένα LED προς μια φωτοδίοδο. Ένα ολοκληρωμένο κύκλωμα μέσω της συχνότητας των διακοπών αυτών παράγει ένα ψηφιακό σήμα, που αντιστοιχεί στην διανυσματική ταχύτητα (κατεύθυνση και τιμή) του ποντικιού. Το ψηφιακό σήμα μεταφέρεται στον Η/Υ μέσω του καλωδίου επικοινωνίας. Ο Η/Υ τότε μετακινεί τον κέρσορα στην οθόνη, βασιζόμενος στην προηγούμενη του θέση και στο ψηφιακό σήμα. [26 (όρος mouse)]



Εικόνα 31 : Λειτουργία ενός μηχανικού ποντικιού

1. κουνώντας το ποντίκι γυρίζει η μπίλια
2. τα X και Y γρανάζια πιάνουν την μπίλια και μεταδίδουν την κίνηση
3. οι οπτικοί δίσκοι κωδικοποίησης περιλαμβάνουν μικρές τρύπες
4. υπέρυθρα LEDs ακτινοβολούν μέσω των δίσκων
5. αισθητήρες συλλέγουν παλμούς φωτός για να μετατρέψουν σε X και Y διανύσματα

#### Οπτικό ποντίκι

Η τεχνολογία του οπτικού ποντικιού αναπτύχθηκε το 1999 από τις τεχνολογίες Agilent και χρησιμοποιεί μία μικροσκοπική κάμερα που τραβά χιλιάδες φωτογραφίες ανά δευτερόλεπτο, προκειμένου να καθορίσει την θέση και την ταχύτητα. Το οπτικό ποντίκι χρησιμοποιεί μία πολύ μικρή δίοδο εκπομπής που αναφέρεται ως LED και είναι κόκκινου χρώματος. Το LED αυτό στέλνει το φως που αντανακλάται στο mouse pad ή σε κάποια επιφάνεια σε ένα μικροεπεξεργαστή CMOS.

Πιο συγκεκριμένα, αρχικά, το LED παράγει ένα κόκκινο φως που στέλνεται στην επιφάνεια στην οποία βρίσκεται το ποντίκι. Το φως αντανακλάται από την επιφάνεια πίσω στον αισθητήρα του CMOS. Ο αισθητήρας, στη συνέχεια, στέλνει κάθε εικόνα που αντανακλάται πίσω στον μικροεπεξεργαστή DSP για ανάλυση. Χρησιμοποιώντας τις χιλιάδες εικόνων που το CMOS στέλνει στο DSP για ανάλυση, το DSP μπορεί να ανιχνεύσει πρότυπα και εικόνες και να καθορίσει εάν το ποντίκι έχει κουνηθεί, πόση απόσταση έχει διανύσει και με τι ταχύτητα. Μπορεί, επίσης, να υπολογίσει τις συντεταγμένες της θέσης του ποντικιού και να τις στείλει στον υπολογιστή στον οποίο είναι συνδεδεμένο το ποντίκι. Αφού ο υπολογιστής λάβει τις συντεταγμένες, δείχνει την κίνηση του ποντικιού, συνήθως κουνώντας τον κέρσορα στην οθόνη. Ο υπολογιστής ανανεώνεται διαρκώς με συντεταγμένες (περισσότερες από χιλιάδες ανά δευτερόλεπτο), καθώς η κίνηση είναι πολύ γρήγορη. Ωστόσο, η εμφάνιση του κέρσορα στην οθόνη, καθώς το ποντίκι αλλάζει θέσεις είναι αρκετά ομαλή.

Υπάρχουν αρκετά πλεονεκτήματα χρήσης οπτικού ποντικιού έναντι του μηχανικού με την μπίλια. Αρχικά, δεν υπάρχουν κινούμενα μέρη. Αφού το οπτικό ποντίκι

χρησιμοποιεί LED αντί για μια μικρή λαστιχένια μπίλια, δεν υπάρχει περίπτωση το LED να λερωθεί ή να κολλήσει. Αυτό οδηγεί σε υψηλή αξιοπιστία και σε μακροχρόνια χρήση χωρίς να χρειάζεται ιδιαίτερη συντήρηση. Επειδή η τεχνολογία είναι πιο εξελιγμένη, αφού χρησιμοποιούνται LEDs και μικροεπεξεργαστές CMOS, η απόδοση και τα αποτελέσματα κίνησης είναι σαφώς καλύτερα. Ένας ακόμη λόγος που τα οπτικά ποντίκια χρησιμοποιούνται περισσότερο από τα μηχανικά, είναι το γεγονός ότι μπορούν να λειτουργήσουν σε οποιαδήποτε επιφάνεια. Σε αντίθεση με τα μηχανικά που χρειάζονται mouse pads για να λειτουργήσουν σωστά και χωρίς προβλήματα, τα οπτικά δουλεύουν εξίσου καλά σε οποιαδήποτε επιφάνεια ή υλικό. [29]



Εικόνα 32 : Ένα οπτικό ποντίκι

#### Ποντίκι Laser

Το ποντίκι laser χρησιμοποιεί μια υπέρυθη δίοδο λέιζερ, αντί για ένα LED για να φωτίζει την επιφάνεια κάτω από τους αισθητήρες. Ήδη από το 1998, η Sun Microsystems παρείχε ένα ποντίκι laser με τους διακομιστές Sun SPARCstation και τους σταθμούς εργασίας. Ωστόσο, τα ποντίκια λέιζερ δεν μπόρεσαν να γίνουν ο κύριος κορμός της αγοράς μέχρι το 2004, όταν η Logitech, σε συνεργασία με την Agilent Technologies, παρουσίασαν το MX 1000 laser mouse. Αυτό το ποντίκι χρησιμοποιεί μια μικρή υπέρυθη ακτίνα λέιζερ αντί της ATA και αύξησε σημαντικά την ανάλυση της εικόνας που λαμβάνεται από το ποντίκι. [37] Το πλεονέκτημα είναι ότι λόγω της δίοδου λέιζερ, το ποντίκι δίνει καλύτερους χρόνους απόκρισης, ανίχνευσης και την δυνατότητα να χρησιμοποιείται σε ακόμα περισσότερες επιφάνειες. [29]



Εικόνα 33 : Ένα ποντίκι laser

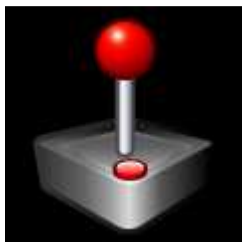
## 3.4 Joystick

### 3.4.1 Εισαγωγή

Το joystick είναι συσκευή εισόδου που αποτελείται από μία ράβδο τοποθετημένη σε μία βάση, έναν μικρό, δηλαδή, μοχλό που μπορεί να περιστραφεί γύρω από το ένα άκρο του. Η περιστροφή του μοχλού αντιστοιχίζεται σε κίνηση και έτσι το joystick αναφέρει την γωνία και την κατεύθυνση του στην συσκευή που ελέγχει. Χρησιμοποιείται για τον έλεγχο βιντεοπαιχνιδιών και συνήθως έχει ένα ή περισσότερα κουμπιά των οποίων η κατάσταση μπορεί να διαβαστεί από τον υπολογιστή. [26 (όρος joystick)]

Στην ιστορία των ηλεκτρονικών παιχνιδιών χρησιμοποιήθηκαν πολλοί διαφορετικοί τύποι ελεγκτών. Ο πιο γνωστός ελεγκτής είναι το joystick (ένα απλό κουτί με κουμπιά και μια ράβδο για τον έλεγχο της κίνησης στο παιχνίδι). Υπάρχουν, βέβαια, και πολλές παραλλαγές και διαφορετικοί τύποι joystick.

Μια δημοφιλής παραλλαγή του, που χρησιμοποιείται στα σύγχρονα βιντεοπαιχνίδια, είναι το αναλογικό joystick. Ένα joystick, δηλαδή με συνεχείς καταστάσεις που επιστρέφει την γωνία κίνησης σε οποιαδήποτε κατεύθυνση στο χώρο ή στο επίπεδο (χρησιμοποιώντας συνήθως ποτενσιόμετρο). Βρίσκει εφαρμογή στο πιλοτήριο των αεροσκαφών, στον έλεγχο μηχανών όπως γερανοί, φορτηγά, μη επανδρωμένα υποβρύχια οχήματα κλπ. Ένα επίσης γνωστό joystick είναι το ψηφιακό joystick, που δίνει μόνο σήματα on/off για τις τέσσερις διαφορετικές κατευθύνσεις και τους πιθανά μηχανικούς συνδυασμούς (όπως πάνω-δεξιά, κάτω-αριστερά, κλπ.). Τα ψηφιακά joysticks χρησιμοποιούνταν συχνά ως ελεγκτές παιχνιδιών στις κονσόλες βιντεοπαιχνιδιών και στους προσωπικούς υπολογιστές του 1980. Άλλα joysticks έχουν δυνατότητα ανάδρασης. Πρόκειται για ενεργές συσκευές και όχι απλά συσκευές εισόδου. Ο υπολογιστής μπορεί να επιστρέψει ένα σήμα στο joystick, έτσι ώστε αυτό να αντισταθεί στην κίνηση ή να δονηθεί. Μικροσκοπικά joysticks καθοδηγούμενα από τα δάχτυλα χρησιμοποιούνται σαν συσκευές εισόδου για μικρότερο ηλεκτρονικό εξοπλισμό, όπως τα κινητά τηλέφωνα. [45]



Εικόνα 34 : Το σκίτσο ενός απλού joystick

### 3.4.2 Τεχνικά χαρακτηριστικά

Τα joysticks διακρίνονται από ένα πολύ έξυπνο και προσεγμένο τέχνασμα. Δέχονται κάτι τελείως φυσικό – την κίνηση του χεριού μας – και το “μεταφράζουν” σε κάτι εντελώς μαθηματικό – μια συμβολοσειρά από 0 και 1 (την γλώσσα των υπολογιστών). Με ένα καλό joystick, η “μετάφραση” είναι τόσο άψογη (χωρίς κανένα ελάττωμα) και έτσι την ξεχνάμε εντελώς. Όταν κάποιος είναι απορροφημένος σε ένα παιχνίδι, πρέπει να αισθάνεται ότι αλληλεπιδρά άμεσα και απευθείας με τον εικονικό κόσμο. [44]



Εικόνα 35 : τμήματα joystick

1. Ράβδος
2. Βάση
3. Σκανδάλη
4. Επιπλέον κουμπιά
5. Διακόπτης αυτόματου πυροβολισμού
6. Ρυθμιστική βαλβίδα
7. Διακόπτης
8. Βεντούζες

Τα περισσότερα joysticks είναι δύο διαστάσεων, έχουν δύο άξονες κίνησης (όπως το ποντίκι), αλλά υπάρχουν και joysticks τριών διαστάσεων. Γενικά, μετακινώντας τη ράβδο αριστερά ή δεξιά, καθορίζεται η κίνηση στον x άξονα, ενώ μετακινώντας την μπροστά (πάνω) ή πίσω (κάτω), καθορίζεται η κίνηση στον y άξονα. Στα joysticks τριών διαστάσεων, στρίβοντας τη ράβδο αριστερά (αντίθετα με τη φορά των δεικτών ρολογιού) ή δεξιά (σύμφωνα με τη φορά των δεικτών του ρολογιού), καθορίζεται η κίνηση στον z άξονα. Επιπλέον, τα joysticks διαθέτουν συχνά ένα ή περισσότερα κουμπιά ενεργοποίησης (fire buttons), προκειμένου να προκαλέσουν κάποιου είδους δράση. Αυτοί είναι απλοί διακόπτες on/off.





Εικόνα 36 : Πιθανές κινήσεις της ράβδου ενός joystick (στις τρεις διαστάσεις)

Το joystick συνδέεται απευθείας στην κατάλληλη θύρα της κάρτας ήχου (παράλληλη θύρα). Υπάρχουν βέβαια και joysticks που διαθέτουν δική τους κάρτα, όμως είναι σχετικά ακριβά. Υπάρχει μια μεγάλη ποικιλία, από τα πιο φθηνά με χαμηλή ανταπόκριση μέχρι τα πιο ακριβά που διαθέτουν μεγάλη ευαισθησία στις κινήσεις. [26 (όρος joystick)]

## 3.5 Webcamera

Η webcamera ανήκει στην κατηγορία των συσκευών εισόδου video. Παρέχουν οπτικά δεδομένα στον υπολογιστή σε μορφή εικόνων ή βίντεο. Ανάλογη κατηγορία συσκευών εισόδου είναι και οι συσκευές audio οι οποίες παρέχουν στον υπολογιστή δεδομένα με την μορφή ήχου (όπως για παράδειγμα τα μικρόφωνα). [28]

### 3.5.1 Ορισμός

Η webcamera ή webcam είναι μια συσκευή καταγραφής βίντεο που συνδέεται με υπολογιστή (μέσω θύρας USB) ή σε δίκτυο υπολογιστών (χρησιμοποιώντας Ethernet ή Wi-Fi). Μία webcamera είναι ουσιαστικά το ίδιο με μια κάμερα που είναι συνδεδεμένη με έναν υπολογιστή, είτε άμεσα, είτε ασύρματα και συγκεντρώνει μια σειρά από εικόνες για την απομακρυσμένη επίδειξη τους οπουδήποτε. Η τεχνολογία της webcam χρησιμοποιείται ευρέως, από όλα τα είδη των ανθρώπων για πολλούς διαφορετικούς λόγους. Η πιο δημοφιλής χρήση της είναι σαν “βίντεο - τηλέφωνο”, που επιτρέπει σε έναν υπολογιστή να λειτουργεί ως τηλέφωνο με βίντεο. Μπορεί να χρησιμοποιηθεί με αυτό τον τρόπο σε πολλά προγράμματα όπως το Windows Live Messenger, το Skype

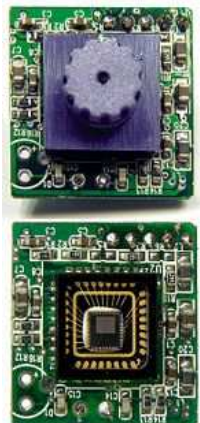
κλπ. Οι webcams είναι γνωστές για το χαμηλό κόστος παραγωγής τους και την ευελιξία τους, που τις καθιστά ως τον τύπο βίντεο-τηλεφωνίας με το χαμηλότερο κόστος. Το 1991, η πρώτη webcam χρησιμοποιήθηκε στο πανεπιστήμιο του Cambridge στο τμήμα Επιστήμης υπολογιστών και από τότε έχουν εισχωρήσει σε σπίτια, επιχειρήσεις, υπηρεσίες με απεριόριστες χρήσεις.



Εικόνα 37 : Μια τυπική web camera

Οι webcams περιλαμβάνουν συνήθως ένα φακό, έναν αισθητήρα εικόνας και μερικά βοηθητικά ηλεκτρονικά μέρη. Διάφοροι φακοί είναι διαθέσιμοι, αλλά ο πιο συνηθισμένος είναι ένας πλαστικός φακός που τοποθετείται με τέτοιο τρόπο ώστε να ρυθμίζεται η εστίαση της κάμερας. Φακοί σταθερής εστίασης, οι οποίοι δεν κάνουν καμία πρόβλεψη για ρύθμιση, είναι επίσης διαθέσιμοι. Οι αισθητήρες εικόνας μπορεί να είναι CMOS ή CCD. Η πρώτη κατηγορία χρησιμοποιείται σε χαμηλού κόστους κάμερες. Οι περισσότερες κάμερες είναι σε θέση να προσφέρουν ανάλυση βίντεο VGA, με ρυθμό των 30 καρέ ανά δευτερόλεπτο. Πολλές νεότερες συσκευές μπορούν να παράγουν βίντεο σε αναλύσεις πολλών megapixel και μερικές μπορούν να τρέχουν σε υψηλούς ρυθμούς καρέ, πχ. στα 120 καρέ ανά δευτερόλεπτο. Βοηθητικά ηλεκτρονικά μέρη συμβάλλουν στην ανάγνωση εικόνας από τον αισθητήρα και την διαβιβάζουν στον κεντρικό υπολογιστή. Οι περισσότερες κάμερες διαθέτουν ενσωματωμένα μικρόφωνα για να διευκολύνουν κλήσεις βίντεο και τηλεδιάσκεψης.

Ειδικό λογισμικό μπορεί να χρησιμοποιήσει τη ροή βίντεο από μια κάμερα για να βοηθήσει ή να βελτιώσει τον έλεγχο ενός χρήστη σε εφαρμογές και παιχνίδια. Χαρακτηριστικά βίντεο, τα οποία συμπεριλαμβάνουν πρόσωπα, σχήματα, μοντέλα και χρώματα μπορούν να παρατηρηθούν και να ανιχνευτούν για την παραγωγή αντίστοιχης μορφής ελέγχου. Για παράδειγμα, η θέση μιας πηγής φωτός μπορεί να ανιχνευθεί και να χρησιμοποιηθεί για να μιμηθεί το δείκτη του ποντικιού, ένα φως τοποθετημένο στο κεφάλι, επιτρέπει υπολογισμούς απουσία χεριών και θα μπορούσε να βελτιώσει την προσβασιμότητα του υπολογιστή. Τέτοιες δυνατότητες μπορούν να εφαρμοστούν σε παιχνίδια, παρέχοντας επιπλέον έλεγχο και αυξημένη διαδραστικότητα. [26 (όρος webcam)]



Εικόνα 38 : Οι webcams περιλαμβάνουν ένα φακό(πάνω σχήμα) , έναν αισθητήρα εικόνας(κάτω σχήμα) και βοηθητικό κύκλωμα.

### 3.6 Σύνθετες συσκευές

Διάφορα κουμπιά και το joystick μπορούν να συνδυαστούν σε μια ενιαία φυσική συσκευή που θεωρείται μια σύνθετη συσκευή εισόδου. Πολλές συσκευές ηλεκτρονικών παιχνιδιών διαθέτουν τέτοιου είδους ελεγκτές. Το ποντίκι μπορεί να θεωρηθεί και αυτό σαν σύνθετη συσκευή εισόδου, καθώς ανιχνεύει κίνηση και ταυτόχρονα παρέχει κουμπιά για εκτέλεση λειτουργιών, αλλά οι σύνθετες συσκευές εισόδου θεωρείται ότι έχουν περισσότερους από δύο διαφορετικούς τύπους εισόδου. Μία από τις πιο σημαντικές σύνθετες συσκευές εισόδου είναι ο ελεγκτής του παιχνιδιού Wii της Nintendo. [26 (όρος input device)]

#### 3.6.1 Wii remote

Το Wii remote, γνωστό και ως Wiimote είναι ο βασικός ελεγκτής για την κονσόλα Wii της Nintendo. Πρόκειται για μία εξαιρετικά ευέλικτη συσκευή που ενσωματώνει επιταχυνσιόμετρο (accelerometer) υψηλής ακρίβειας, το οποίο μπορεί να συλλάβει κίνηση με 3 βαθμούς ελευθερίας. Αντιλαμβάνεται, δηλαδή, κίνηση στους 3 άξονες του χώρου, ταχύτητα, θέση και κλίση, κάτι που επιτρέπει στο χρήστη να αλληλεπιδράσει με και να χειριστεί αντικείμενα στην οθόνη. Παράλληλα αναγνωρίζει που ακριβώς στοχεύει στην οθόνη, μπορεί να δονείται για μεγαλύτερο ρεαλισμό, ενώ με το ενσωματωμένο ηχείο του μεταφέρει τους ήχους που εκπέμπονται κοντά στο χρήστη. Διαθέτει επίσης, μια κάμερα υπέρυθρων που μπορεί να εντοπίσει τις πηγές υπέρυθρου

φωτός. Το Wiimote επικοινωνεί με ένα κεντρικό ελεγκτή (host controller), όπως η κονσόλα Wii ή ένας υπολογιστής, μέσω ασύρματης τεχνολογίας Bluetooth. Χρησιμοποιώντας αυτή την ασύρματη σύνδεση, μπορεί να μεταδώσει τα δεδομένα του επιταχυνσιόμετρου και τις συντεταγμένες των pixels της κάμερας σε συχνότητες των 100 Hz.

Τα στοιχεία του Wiimote του επιτρέπουν να χρησιμοποιηθεί με ποικίλους τρόπους, αλλά χρησιμοποιείται συνήθως ως διεπαφή υλικού για τα παιχνίδια που παίζονται στην κονσόλα Wii. Ο χρήστης κρατά το Wiimote, το μετακινεί και τα δεδομένα κίνησης αποστέλλονται στο κεντρικό ελεγκτή.

Συνδεδεμένη με την κονσόλα είναι μια μπάρα αισθητήρων, η οποία διαθέτει ένα σύνολο από λαμπάκια υπερύθρων (IR leds). Η κονσόλα υπολογίζει τη σχετική τρισδιάστατη θέση του Wiimote μέσω τριγωνοποίησης, με βάση τη σχετική θέση των πηγών υπερύθρων που εντοπίζονται από την κάμερα υπερύθρων.

Το Wiimote δεν ανιχνεύει μόνο το υπέρυθρο φως που εκπέμπεται από την μπάρα αισθητήρων. Μπορεί να χρησιμοποιηθεί σαν συσκευή ανίχνευσης για κάθε πηγή υπέρυθρου φωτός, συμπεριλαμβανομένων των κεριών, στυλό υπερύθρων, ακόμα και φυσικό φως της ημέρας. Χρησιμοποιώντας αυτή την ιδέα, το Wiimote μπορεί να τοποθετηθεί σε κάποιο σταθερό αντικείμενο και να ανιχνεύσει με ακρίβεια κάθε πηγή υπέρυθρου φωτός στο οπτικό πεδίο της κάμερας. Σε αυτή την περίπτωση, το επιταχυνσιόμετρο δεν θα χρησιμοποιηθεί αφού το Wiimote παραμένει ακίνητο.

Ένα άλλο χαρακτηριστικό είναι η επεκτασιμότητα του μέσω της χρήσης πρόσθετων εξαρτημάτων όπως το Nunchuk, το οποίο παρέχεται στα πλαίσια της κονσόλας Wii και συμπληρώνει το Wii remote παρέχοντας λειτουργίες ανάλογες με αυτές των ελεγκτών gamepad.

Ο ελεγκτής παρουσιάστηκε στο Tokyo Game Show στις 16 Σεπτεμβρίου του 2005 και η ονομασία Wii remote ανακοινώθηκε στις 27 Απριλίου του 2006. Από τότε έχει λάβει μεγάλη αναγνώριση χάρη στα μοναδικά χαρακτηριστικά του και τις διαφορές του σε σχέση με τους παραδοσιακούς ελεγκτές παιχνιδιών. Ταυτόχρονα, χάκερ χρησιμοποιούν το Wii remote για να ελέγξουν συσκευές ανεξαρτήτως του Wii, εξυπηρετώντας τους δικούς τους σκοπούς και αναπτύσσοντας τις δικές τους εφαρμογές. [26 (όρος Wii Remote)]



Εικόνα 39 : Το χειριστήριο Wii remote

## Κεφάλαιο 4

### Ανίχνευση (tracking)

#### 4.1 Εισαγωγή

Μία σημαντική πρόκληση για τους ερευνητές στο πεδίο computer vision είναι η ανάπτυξη συστημάτων που ανιχνεύουν την κίνηση οποιουδήποτε αντικειμένου, του προσώπου ή του σώματος του χρήστη. Σε πολλές περιπτώσεις, σχεδιάζονται τέτοια συστήματα έχοντας ως συγκεκριμένο σκοπό να προσφέρουν σε ανθρώπους με περιορισμένες ικανότητες την δυνατότητα να χρησιμοποιούν συστήματα υπολογιστών με αρκετά απλό τρόπο. Ένας άλλος σκοπός της ανάπτυξης τέτοιων συστημάτων είναι η μοντελοποίηση της αλληλεπίδρασης εικονικών τρισδιάστατων χαρακτήρων με χρήστες. Γενικά βρίσκουν εφαρμογή σε πολλές δυναμικές εφαρμογές Augmented Reality, όπως η ιατρική απεικόνιση, η συντήρηση και επισκευή, η ψυχαγωγία, η πλοήγηση των αεροσκαφών και η στόχευση, σε συστήματα ελέγχου της θέσης συσκευών όπως ενός ρομποτικού βραχίονα, σε διεπαφές ανθρώπου-μηχανής πιο φυσικές και διαισθητικές σε σχέση με τις παραδοσιακές. Στοχεύουν, δηλαδή γενικότερα, στην δημιουργία πρωτοποριακών διεπαφών με όλο και μεγαλύτερες και πιο ενδιαφέρουσες δυνατότητες.

Η δισδιάστατη ανίχνευση έχει ως στόχο, κατά κύριο λόγο την παρακολούθηση της προβολής εικόνας των αντικειμένων, ή μερών των αντικειμένων, των οποίων η τρισδιάστατη μετατόπιση καταλήγει σε μια κίνηση που μπορεί να μοντελοποιηθεί σαν δισδιάστατος μετασχηματισμός. Στη συνέχεια απαιτείται ένα προσαρμοζόμενο μοντέλο, για να χειριστεί αλλαγές εμφάνισης εξαιτίας των αποτελεσμάτων προοπτικής ή παραμόρφωσης. Μπορεί να προβάλει τη θέση της εικόνας του αντικειμένου από πλευράς του κέντρου βάρους και κλίμακας ή ενός συσχετισμένου μετασχηματισμού (affine transformation). Εναλλακτικά, πιο εξελιγμένα μοντέλα όπως splines, παραμορφωμένα πρότυπα ή δισδιάστατα αρθρωτά μοντέλα μπορούν να χρησιμοποιηθούν. Ωστόσο, καμία από αυτές τις μεθόδους δεν αφορά την ανάκτηση της πραγματικής θέσης στο χώρο. Αντιθέτως, η τρισδιάστατη ανίχνευση έχει ως στόχο τη συνεχή ανάκτηση και των έξι βαθμών ελευθερίας (DOF, degrees of freedom) που καθορίζουν την θέση της κάμερας και τον προσανατολισμό σε σχέση με τη σκηνή, ή, ισοδύναμα, την τρισδιάστατη μετατόπιση ενός αντικειμένου σε σχέση με την κάμερα.

Η βιβλιογραφία για την τρισδιάστατη ανίχνευση είναι ιδιαίτερα ογκώδης καθώς υπάρχουν τόσες πολλές προσεγγίσεις όσες οι εφαρμογές που προσπαθούν να υλοποιηθούν και επειδή είναι πιθανές πολλές διαφορετικές προσεγγίσεις για την επίλυση του ίδιου προβλήματος.

#### **4.1.1 Τρισδιάστατη ανίχνευση στο πεδίο Computer Vision**

Πολλές τεχνολογίες πέρα από το πεδίο Computer Vision έχουν προσπαθήσει να επιτύχουν την τρισδιάστατη ανίχνευση, αλλά όλες είχαν αδυναμίες. Οι μηχανικοί ανιχνευτές είναι αρκετά ακριβείς αλλά δεσμεύουν τον χρήστη σε ένα περιορισμένο όγκο εργασίας. Οι μαγνητικοί ανιχνευτές είναι ευάλωτοι σε παραμορφώσεις του μετάλλου στο περιβάλλον, οι οποίες είναι πολύ συνηθισμένες και επίσης περιορίζουν το εύρος των μετατοπίσεων. Υπέρηχοι ανιχνευτές υποφέρουν από τον θόρυβο και τείνουν να είναι ανακριβείς σε μεγάλα εύρη, λόγω μεταβολών στη θερμοκρασία περιβάλλοντος.

Αντίθετα, το πεδίο Computer Vision έχει τη δυνατότητα να παράγει ακριβείς και χαμηλού κόστους λύσεις στο πρόβλημα, υπό την προϋπόθεση ότι κάποιος είναι πρόθυμος να επενδύσει την απαιτούμενη προσπάθεια ώστε να αναπτύξει “ισχυρούς” αλγόριθμους. Σε μερικές περιπτώσεις, είναι δυνατόν να προστεθούν fiducials, όπως LEDs ή ειδικοί δείκτες, στη σκηνή ή το αντικείμενο στόχου, για να διευκολυνθεί το έργο καταγραφής. Φυσικά, αυτό προϋποθέτει ότι ένα ή περισσότερα fiducials είναι ορατά ανά πάσα στιγμή. Σε αντίθετη περίπτωση, η καταγραφή καταρρέει. Ωστόσο, δεν είναι πάντοτε δυνατόν να τοποθετηθούν fiducials. Για παράδειγμα, οι χρήστες εφαρμογών Augmented Reality προτιμούν να τα αποφεύγουν επειδή είναι ορατά στη σκηνή και δεν είναι πάντα εφικτό να τροποποιηθεί το περιβάλλον προτού τρέξει η εφαρμογή. Συνεπώς, είναι πολύ πιο επιθυμητό να βασίζονται σε φυσικά χαρακτηριστικά, όπως τα άκρα, γωνίες, ή η υφή. Φυσικά, αυτό καθιστά την ανίχνευση πολύ πιο δύσκολη: Η εύρεση και η παρακολούθηση χαρακτηριστικών στοιχείων ή άκρων μπορεί να είναι δύσκολη επειδή υπάρχουν πολύ λίγα από αυτά σε τυπικά αντικείμενα.

Ολική ή ακόμα και μερική απόφραξη των αντικειμένων οδηγεί σε αποτυχία ανίχνευσης. Η κάμερα μπορεί να κινείται πολύ γρήγορα, με αποτέλεσμα οι εικόνες να παράγουν μια θολή κίνηση. Ο φωτισμός μπορεί να αλλάζει σημαντικά. Αντανεκλάσεις μπορεί να προκαλέσουν σύγχυση στον ανιχνευτή. Επιπλέον, ένα αντικείμενο μπορεί να αλλάζει όψη πολύ γρήγορα λόγω μετατόπισης. Για παράδειγμα, αυτό συμβαίνει όταν μια κάμερα κινηματογραφεί ένα κτίριο και πηγαίνει προς την γωνία, προκαλώντας τον ένα

τοίχο να εξαφανιστεί και ένα νέο να εμφανιστεί. Στις περιπτώσεις αυτές, τα χαρακτηριστικά που πρέπει να παρακολουθηθούν αλλάζουν συνεχώς και ο ανιχνευτής πρέπει να σχεδιαστεί ώστε να χειρίζεται χαρακτηριστικά μέσα και έξω από την εικόνα.

#### 4.1.2 Μαθηματικά εργαλεία

##### *Αναπαράσταση κάμερας (Camera Representation)*

Εστιάζουμε στο μοντέλο της τυπικής κάμερας pinhole, συμπεριλαμβανομένων των δυνητικών αποκλίσεων από αυτό, το οποίο είναι κατάλληλο για τις περισσότερες κάμερες που χρησιμοποιούνται σήμερα για σκοπούς ανίχνευσης. Αξίζει να σημειωθεί ωστόσο, ότι οι νέοι σχεδιασμοί για κάμερες, όπως οι πανκατευθυντικές κάμερες που βασίζονται σε υπερβολικά ή παραβολικά κάτοπτρα για την επίτευξη οπτικών πεδίων με μεγάλο εύρος, γίνονται όλο και πιο δημοφιλείς.

##### *Το μοντέλο προοπτικής προβολής (The Perspective Projection Model)*

Μαθηματικά, η διαμόρφωση της εικόνας μπορεί να οριστεί ως η προβολή από τον τρισδιάστατο χώρο στο επίπεδο εικόνας. Οι συντεταγμένες ενός τρισδιάστατου σημείου  $M = [X, Y, Z]^T$ , εκφρασμένες σε ένα Ευκλείδειο σύστημα παγκόσμιων συντεταγμένων και το αντίστοιχο δισδιάστατο σημείο  $m = [u, v]^T$  σχετίζονται με την εξίσωση

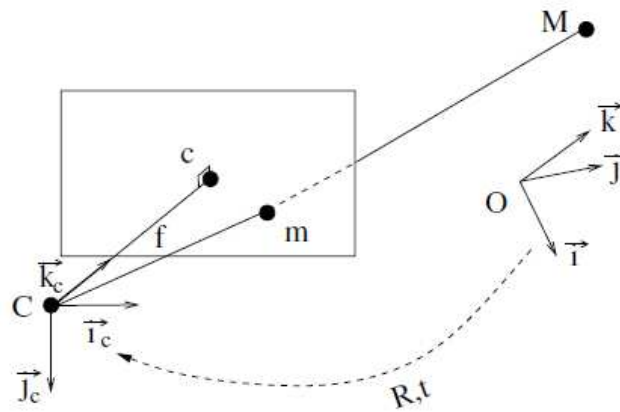
$$sm = PM \quad (\text{σχέση 1})$$

όπου  $s$  είναι ο παράγοντας κλίμακας,  $m = [u, v, 1]^T$  και  $M = [X, Y, Z, 1]^T$  είναι οι ομογενείς συντεταγμένες των σημείων  $m$  και  $M$  και  $P$  ένας  $3 \times 4$  πίνακας προβολής. Ο  $P$  ορίζεται με ένα συντελεστή κλίμακας και εξαρτάται από 11 παραμέτρους. Συνήθως, θεωρείται σαν πίνακας προοπτικής προβολής. Ένας τέτοιος πίνακας προοπτικής προβολής μπορεί να αναλυθεί ως

$$P = K[R|t],$$

όπου  $K$  είναι ο  $3 \times 3$  πίνακας διαρρύθμισης που εξαρτάται από τις εσωτερικές παραμέτρους της κάμερας, όπως το μήκος εστίασης,  $[R | t]$  είναι ο  $3 \times 4$  πίνακας εξωτερικών παραμέτρων και αντιστοιχεί στον ευκλείδειο μετασχηματισμό από ένα παγκόσμιο σύστημα συντεταγμένων στο τοπικό σύστημα συντεταγμένων της κάμερας: το  $R$  αναπαριστά έναν πίνακα περιστροφής  $3 \times 3$  και το  $t$  μια μετατόπιση.





Εικόνα 40 : Το μοντέλο προοπτικής προβολής:  $(O, \vec{i}, \vec{j}, \vec{k})$  είναι το παγκόσμιο σύστημα συντεταγμένων,  $(C, \vec{i}_c, \vec{j}_c, \vec{k}_c)$  είναι το σύστημα συντεταγμένων της κάμερας, M ένα τρισδιάστατο σημείο και m είναι η προβολή του στο επίπεδο εικόνας.

Ο πίνακας διαρρύθμισης κάμερας (The Camera Calibration Matrix)

Ο πίνακας διαρρύθμισης κάμερας K περιέχει τις εσωτερικές παραμέτρους της κάμερας.

Έχουμε,

$$K = \begin{pmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{σχέση 2})$$

όπου  $\alpha_u$  και  $\alpha_v$  είναι αντίστοιχα ο παράγοντας κλίμακας στις u και v κατευθύνσεις συντεταγμένων. Είναι ανάλογες με το μήκος εστίασης f της κάμερας:  $\alpha_u = k_u f$  και  $\alpha_v = k_v f$ , όπου  $k_u$  και  $k_v$  είναι ο αριθμός των pixels ανά μονάδα απόστασης στις κατευθύνσεις u και v. Το  $c = [u_0, v_0]^T$  αναπαριστά τις συντεταγμένες εικόνας της τομής του οπτικού άξονα και του επιπέδου εικόνας, που καλείται και σαν κύριο σημείο. Το s (αναφέρεται και ως skew) είναι μη μηδενικό μόνο όταν οι u και v κατευθύνσεις δεν είναι κάθετες, το οποίο είναι εξαιρετικά σπάνιο στις καινούριες κάμερες.

Μια πολύ λογική προσέγγιση είναι να λαμβάνουμε το κύριο σημείο c να βρίσκεται στο κέντρο της εικόνας. Ομοίως, εάν έχουμε υποθέσει ότι τα pixels είναι τετράγωνα, τα  $\alpha_u$  και  $\alpha_v$  μπορούν θεωρηθούν ίσα. Επομένως, πλέον, μια κάμερα θεωρείται ρυθμισμένη όταν οι εσωτερικές παράμετροι είναι γνωστές.

*Ο πίνακας εξωτερικών παραμέτρων (The External Parameters Matrix)*

Ο πίνακας εξωτερικών παραμέτρων  $[R|t]$  προσδιορίζει τον προσανατολισμό και τη θέση της κάμερας. Σχηματίζεται από ένα πίνακα περιστροφής  $R$  και ένα διάνυσμα μετακίνησης  $t$ . Συχνά αναφερόμαστε σε αυτόν σαν “στάση” κάμερας. Οι εφαρμογές ανίχνευσης συνήθως υποθέτουν ότι ο πίνακας διαρρύθμισης  $K$  είναι γνωστός και εστιάζουν στον υπολογισμό των  $R$  και  $t$ , ή ισοδύναμα στη θέση και προσανατολισμό του αντικειμένου-στόχου ως προς την κάμερα. Ή διαφορετικά, αντιστοιχεί στην Ευκλείδεια μετατροπή από ένα παγκόσμιο σύστημα συντεταγμένων στο σύστημα συντεταγμένων της κάμερας : ένα τρισδιάστατο σημείο που απεικονίζεται με το διάνυσμα  $M_w$  στο παγκόσμιο σύστημα συντεταγμένων αναπαρίσταται με το διάνυσμα  $M_c = RM_w + t$  στο σύστημα συντεταγμένων της κάμερας. Από αυτή τη σχέση μπορούμε να ανακτήσουμε την έκφραση του κέντρου της κάμερας ή του οπτικού κέντρου  $c$  στο παγκόσμιο σύστημα συντεταγμένων. Θα πρέπει να ικανοποιεί την σχέση  $0=Rc+t$  η οποία συνεπάγεται  $c = -R^{-1}t = -R^T t$ .

*Εκτίμηση του πίνακα διαρρύθμισης κάμερας*

Στις περισσότερες μεθόδους τρισδιάστατης παρακολούθησης, γίνεται η υπόθεση πως οι εσωτερικές παράμετροι είναι σταθερές και γνωστές, πράγμα που σημαίνει ότι η κάμερα δεν μπορεί να ζουμάρει, επειδή είναι δύσκολο να γίνει διάκριση μεταξύ μεταβολής στο εστιακό μήκος και μετατόπισης κατά μήκος του άξονα  $Z$  της κάμερας. Αυτές οι παράμετροι μπορούν να υπολογιστούν κατά τη διάρκεια ενός offline σταδίου διαρρύθμισης της κάμερας, από τις ίδιες τις εικόνες. Υποθέτουμε πως οι κλασικές μέθοδοι διαρρύθμισης κάνουν χρήση ενός προτύπου διαρρύθμισης γνωστού μεγέθους εντός του οπτικού πεδίου. Μερικές φορές είναι ένα τρισδιάστατο πλέγμα διαρρύθμισης, στο οποίο είναι σχεδιασμένα σχέδια, όπως για παράδειγμα μαύροι δίσκοι.

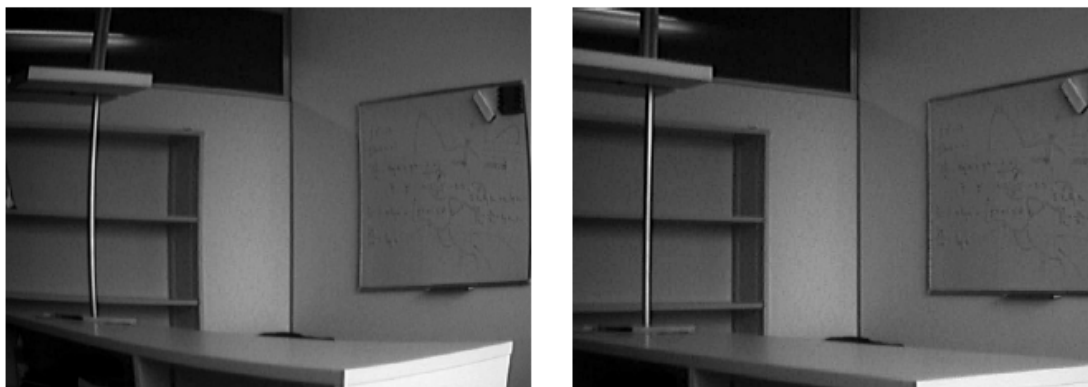


*Εικόνα 41* : Ένα τρισδιάστατο πλέγμα που χρησιμοποιείται για την εκτίμηση του πίνακα διαρρύθμισης κάμερας. Σε αυτό, οι τρεις διαστάσεων συντεταγμένες των γωνιών των άσπρων τετραγώνων ως προς την πάνω αριστερή γωνία του πλέγματος είναι γνωστές. Είναι σχετικά εύκολο να βρει κανείς αυτές τις γωνίες της εικόνας και από την αντιστοιχία μεταξύ των τρισδιάστατων σημείων και των δισδιάστατων σημείων εικόνας να υπολογιστούν οι παράμετροι του πίνακα προβολής.

#### *Χειρισμός παραμόρφωσης φακών*

Το μοντέλο της προοπτικής προβολής δεν είναι πάντα επαρκές για να αναπαραστήσει όλες τις πτυχές του σχηματισμού της εικόνας, διότι δεν λαμβάνει υπόψη την πιθανή στρέβλωση από τους φακούς της κάμερας, η οποία μπορεί να μην είναι αμελητέα, ειδικά για τις κάμερες ευρείας γωνίας. Ευτυχώς, η παραμόρφωση των φακών μπορεί να μοντελοποιηθεί ως δισδιάστατη παραμόρφωση της εικόνας.

Δοθείσας μιας εκτίμησης για τις παραμέτρους παραμόρφωσης, το αποτέλεσμα παραμόρφωσης μπορεί να αφαιρεθεί αποτελεσματικά από την εικόνα που λαμβάνεται από την κάμερα στο χρόνο εκτέλεσης, χρησιμοποιώντας έναν look-up πίνακα, και το μοντέλο προοπτικής προβολής μπορεί να εφαρμοστεί στη συνέχεια. Μεταξύ άλλων, αυτό επιτρέπει τη χρήση καμερών με ευρείες γωνίες για σκοπούς ανίχνευσης, το οποίο μπορεί να είναι επιθυμητό καθώς καθιστούν ευκολότερο το να κρατηθούν τα αντικείμενα-στόχος εντός του οπτικού πεδίου.



*Εικόνα 42* : Αριστερά: Εξαιτίας της προοπτικής παραμόρφωσης, η προβολή ευθείας γραμμής είναι κυρτή. Δεξιά: στην μη παραμορφωμένη εικόνα, η προβολή ευθείας γραμμής είναι πλέον ευθεία.

Μια συνηθισμένη αναπαράσταση των παραμέτρων στρέβλωσης είναι η εξής: Έστω  $\tilde{\mathbf{u}}=[\tilde{u},\tilde{v}]^T$  οι συντεταγμένες της παραμορφωμένης εικόνας και  $\tilde{\mathbf{x}}=[\tilde{x},\tilde{y}]^T$  οι αντίστοιχες κανονικοποιημένες συντεταγμένες έτσι ώστε  $\tilde{u}=u_0+\alpha_u\tilde{x}$  και  $\tilde{v}=v_0+\alpha_v\tilde{y}$ , όπου  $u_0, v_0, \alpha_u, \alpha_v$  είναι οι εσωτερικές παράμετροι της σχέσης 2. Έστω  $\mathbf{u}=(u,v)$  και  $\mathbf{x}=(x,y)$  οι αντίστοιχες μη παραμορφωμένες τιμές. Η παραμόρφωση εκφράζεται ως το άθροισμα δύο στοιχείων έτσι ώστε  $\tilde{\mathbf{x}}=\mathbf{x}+d\mathbf{x}_{\text{radial}}+d\mathbf{x}_{\text{tangential}}$ . Η παραμόρφωση radial μπορεί να προσεγγιστεί ως  $d\mathbf{x}_{\text{radial}}=(1+k_1*r^2+k_2*r^4+\dots)\mathbf{x}$ , όπου  $r=||\mathbf{x}||=\sqrt{(x^2+y^2)}$ .

Η παραμόρφωση tangential έχει πολύ μικρότερη επίδραση. Μπορεί να εκφραστεί ως

$$d\mathbf{x}_{\text{tangential}} = \begin{pmatrix} 2*p_1*x*y + p_2*(r^2+2*x^2) \\ p_1*(r^2+2*y^2) + 2*p_2*x*y \end{pmatrix}$$

αν και συνήθως αγνοείται.

### 4.1.3 Παραμετροποίηση της “στάσης” της κάμερας (Camera Pose Parameterization)

Για σκοπούς βελτιστοποίησης της εκτίμησης των αποτελεσμάτων, η “στάση” της κάμερας πρέπει να ρυθμιστεί κατάλληλα. Αν και η παρουσίαση μετατοπίσεων γίνεται χωρίς ιδιαίτερο πρόβλημα, η παραμετροποίηση περιστροφής στον χώρο  $R^3$  αντιμετωπίζει προβλήματα. Μια περιστροφή στον  $R^3$  έχει μόνο 3 βαθμούς ελευθερίας και θα ήταν περίεργο να χρησιμοποιήσει κανείς απευθείας τα εννιά στοιχεία του  $3 \times 3$  πίνακα περιστροφής. Έξι επιπλέον μη γραμμικοί περιορισμοί – οι τρεις αναγκάζουν τις τρεις στήλες να είναι μοναδιαίου μήκους και οι άλλοι τρεις να τις διατηρούν αμοιβαία ορθογώνιους (mutual orthogonal) – προστίθενται για να βεβαιώσουν πως ο πίνακας είναι ορθοκανονικός (orthonormal).

Έχει αποδειχθεί πως τρεις είναι οι μέθοδοι παραμετροποίησης με αποδοτικά αποτελέσματα για την τρισδιάστατη ανίχνευση : οι γωνίες του Euler, τα quaternions και οι εκθετικοί χάρτες (exponential maps).

#### *Οι γωνίες του Euler*

Ένας πίνακας περιστροφής  $R$  μπορεί πάντα να γραφτεί σαν γινόμενο τριών πινάκων που αντιπροσωπεύουν περιστροφές γύρω από τους άξονες  $x, y$  και  $z$ . Υπάρχουν ορισμένες συμβάσεις όσον αφορά την σειρά με την οποία πραγματοποιούνται αυτές οι περιστροφές. Για παράδειγμα, ας θεωρήσουμε πως τα  $\alpha, \beta$  και  $\gamma$  είναι οι γωνίες

περιστροφής γύρω από τον άξονα z, y και x αντίστοιχα. Τότε ο πίνακας περιστροφής γίνεται

$$R = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix}$$

Η αντίστροφη διαδικασία, δηλαδή η εξαγωγή των γωνιών του Euler για έναν συγκεκριμένο πίνακα περιστροφής, μπορεί να εκτελεστεί εύκολα αναγνωρίζοντας τους συντελεστές του πίνακα και την αναλυτική έκφραση τους.

Αν και οι γωνίες του Euler προσφέρουν αρκετά για ένα μεγάλο αριθμό καμερών έχουν ένα μειονέκτημα. Όταν δύο από τους τρεις άξονες περιστροφής ευθυγραμμιστούν κάποια περιστροφή δεν έχει καμία επίδραση.

### Quaternions

Η περιστροφή στον τρισδιάστατο χώρο μπορεί να παρουσιαστεί και ως ένα μοναδιαίο quaternion. Τα quaternions είναι αριθμοί που μπορούν να γραφτούν με γραμμικό συνδυασμό

$a + bi + cj + dk$ , με την ιδιότητα  $i^2 = j^2 = k^2 = ijk = -1$ .

Μια περιστροφή γύρω από το μοναδιαίο διάνυσμα  $\omega$  υπό γωνία  $\theta$  αντιπροσωπεύεται από το μοναδιαίο quaternion

$$q = (\cos((1/2)\theta), \omega \sin((1/2)\theta)).$$

Για την περιστροφή ενός σημείου  $M$  χρησιμοποιούμε το quaternion  $p = (0, M)$  και παίρνουμε το περιστρεφόμενο σημείο  $p' = q \cdot p \cdot q^{-1}$ , όπου  $\cdot$  είναι το σύμβολο πολλαπλασιασμού των quaternions και  $q^{-1} = (\cos((1/2)\theta), -\omega \sin((1/2)\theta))$  ο συζυγής του  $q$ .

Η τεχνική αυτή δεν αντιμετωπίζει το πρόβλημα της προηγούμενης αλλά πρέπει να φροντίζει ώστε να διατηρεί την νόρμα του  $q$  ίση με 1. Εάν το  $q$  υπολογίστηκε με χρήση τεχνικών βελτιστοποίησης, αυτό μπορεί να γίνει προσθέτοντας τον όρο

$\kappa(1 - \|q\|^2)$  στην συνάρτηση εισόδου, όπου το  $\kappa$  είναι μία σταθερά. Το γεγονός αυτό, όμως, αυξάνει την πολυπλοκότητα του αλγορίθμου.

### Εκθετικοί χάρτες

Σε αντίθεση με την αναπαράσταση των quaternions, οι εκθετικοί χάρτες χρειάζονται τρεις μόνο παραμέτρους για την περιγραφή μιας περιστροφής.

Ας θεωρήσουμε ένα διάνυσμα τριών διαστάσεων και ας το ορίσουμε ως

$\omega = [\omega_x, \omega_y, \omega_z]$ . Η νόρμα του μπορεί να εκφραστεί ως  $\theta = \|\omega\|$ . Μία περιστροφή γωνίας

$\theta$  με βάση το διάνυσμα  $\omega$  μπορεί να αναπαρασταθεί ως

$$\exp(\Omega) = I + \Omega + (1/2!)\Omega^2 + (1/3!)\Omega^3 + \dots$$

όπου  $\Omega$  είναι ο πίνακας

$$\Omega = \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix}$$

Μπορεί να επιλυθεί χρησιμοποιώντας τον τύπο του Rodrigues

$$R(\Omega) = \exp(\Omega) = I + \sin\theta\tilde{\Omega} + (1 - \cos\theta)\tilde{\Omega}^2,$$

όπου  $\tilde{\Omega}$  είναι ο πίνακας που αντιστοιχεί στο μοναδιαίο διάνυσμα ( $\omega/\|\omega\|$ ).

Εν ολίγοις, η αναπαράσταση με την βοήθεια των εκθετικών χαρτών χρησιμοποιεί για την περιγραφή μιας περιστροφής διανύσματα τριών διαστάσεων που δίνουν του άξονες και το μέγεθος. Δεν αντιμετωπίζει το πρόβλημα των γωνιών του Euler και δεν υπακούει σε περιορισμούς όπως τα quaternions.

#### 4.1.4 Εκτίμηση του πίνακα εξωτερικών παραμέτρων

Η συνεχής εκτίμηση του πίνακα εξωτερικών παραμέτρων από μια ροή βίντεο διευκολύνεται αισθητά από μια εκ των προτέρων θέση της κάμερας που δίνεται από προηγουμένως υπολογισθείσες θέσεις κάμερας.

Παρ' όλα αυτά, ακολουθούν προσεγγίσεις για την εκτίμηση των παραμέτρων αυτών χωρίς καμία προηγούμενη γνώση της θέσης της κάμερας, αλλά δοθέντων αντιστοιχιών μεταξύ των σημείων τριών διαστάσεων στο παγκόσμιο σύστημα συντεταγμένων και των προβολών τους στο επίπεδο εικόνας. Είναι συγχρόνως μια καλή εισαγωγή για άλλες προσεγγίσεις ανίχνευσης. Η εκτίμηση των εξωτερικών παραμέτρων, όταν είναι γνωστές οι εσωτερικές παράμετροι αναφέρεται συχνά ως πρόβλημα εκτίμησης "στάσης".

Υποθέτουμε ότι διαθέτουμε  $n$  αντιστοιχίες μεταξύ τρισδιάστατων σημείων  $M_i$  και των προβολών τους  $m_i$ . Θέλουμε να βρούμε τον πίνακα προοπτικής προβολής  $P$  που προβάλλει τα σημεία  $M_i$  στα  $m_i$ . Με άλλα λόγια θέλουμε να ισχύει  $PM_i \equiv m_i$  για όλα τα  $i$ , όπου  $\equiv$  υποδηλώνει την ισότητα με συντελεστή κλίμακας.

*Πόσες αντιστοιχίες απαιτούνται*

Όταν οι εσωτερικές παράμετροι είναι γνωστές,  $n = 3$  γνωστές αντιστοιχίες  $M_i \leftrightarrow m_i$  παράγουν 4 πιθανές λύσεις. Όταν  $n = 4$  ή  $n = 5$  αντιστοιχίες είναι γνωστές, υπάρχουν τουλάχιστον 2 λύσεις, αλλά όταν τα σημεία είναι ομοεπίπεδα και δεν υπάρχει τριάδα συγγραμμικών σημείων, η λύση είναι μοναδική για  $n \geq 4$ . Για  $n \geq 6$  η λύση είναι μοναδική.

*Ο άμεσος γραμμικός μετασχηματισμός (The Direct Linear Transformation ) DLT*

Μπορεί να χρησιμοποιηθεί για να υπολογίσει ολόκληρο τον πίνακα  $P$  της σχέσης 1, λύνοντας ένα γραμμικό σύστημα ακόμα και όταν οι εσωτερικές παράμετροι δεν είναι γνωστές. Κάθε αντιστοιχία  $M_i \leftrightarrow m_i$  αυξάνει 2 γραμμικά ανεξάρτητες εξισώσεις έτσι ώστε:

$$(P_{11}X_i + P_{12}Y_i + P_{13}Z_i + P_{14}) / (P_{31}X_i + P_{32}Y_i + P_{33}Z_i + P_{34}) = u_i$$

$$(P_{21}X_i + P_{22}Y_i + P_{23}Z_i + P_{24}) / (P_{31}X_i + P_{32}Y_i + P_{33}Z_i + P_{34}) = v_i$$

Αυτές οι εξισώσεις μπορούν να ξαναγραφούν στη μορφή  $Ap=0$ , όπου  $p$  ένα διάνυσμα με συντελεστές  $P_{ij}$ . Η προφανής λύση  $p=0$  δεν μας απασχολεί ενώ η σωστή λύση μπορεί να βρεθεί από SVD ανάλυση του  $A$ .

Οι εσωτερικές και εξωτερικές παράμετροι μπορούν στη συνέχεια να εξαχθούν από τον  $P$ , αλλά θα ήταν δύσκολο να εφαρμοστεί αυτή η προσέγγιση για τρισδιάστατη ανίχνευση. Στην πράξη, η ταυτόχρονη εκτίμηση εσωτερικών και εξωτερικών παραμέτρων εξαρτάται σε μεγάλο βαθμό από τη γεωμετρία και τον αριθμό των αντιστοιχιών. Σε ευνοϊκές περιπτώσεις, 15 ως 20 αντιστοιχίες ενδέχεται να αρκούν, ενώ σε δυσμενείς, ακόμη και εκατοντάδες μπορεί να μην είναι αρκετές. Σε τέτοιες περιπτώσεις, είναι πάντα προτιμότερο να εκτιμηθούν οι εσωτερικές παράμετροι ξεχωριστά. Πιο συγκεκριμένα, για τρισδιάστατη παρακολούθηση, η χρήση ρυθμισμένης κάμερας και η εκτίμηση μόνο του προσανατολισμού και της θέσης

$[R / t]$  δίνει καλύτερα αποτελέσματα.

Αφού ο πίνακας  $K$  διαρρύθμισης της κάμερας είναι γνωστός, ο πίνακας εξωτερικών παραμέτρων μπορεί να εξαχθεί από τον  $P$  με συντελεστή κλίμακας:  $[R | t] \sim K^{-1}P$ . Ο  $3 \times 3$  πίνακας κατασκευασμένος από τις 3 πρώτες στήλες αυτού του πίνακα δεν είναι απαραίτητα ένας πίνακας περιστροφής, αλλά μπορεί εύκολα να γίνει διόρθωση.

Σε πραγματικές συνθήκες, οι θέσεις των pixel  $m_i$  είναι συνήθως θορυβώδεις. Επειδή το πραγματικό σφάλμα ελαχιστοποιημένο με τον αλγόριθμο DLT αντιστοιχεί σε

αλγεβρικό λάθος, οι παράμετροι της κάμερας που υπολογίζονται με τη μέθοδο αυτή θα πρέπει να βελτιωθούν μέσω επαναληπτικής βελτιστοποίησης του μη γραμμικού σφάλματος επαναπροβολής.

#### *Το πρόβλημα προοπτικής $n$ σημείων*

Η μέθοδος DLT στοχεύει στον υπολογισμό και των 11 παραμέτρων του πίνακα προβολής. Μπορεί επίσης να εφαρμοστεί για την εκτίμηση θέσης, αλλά στηρίζεται σε μια υπερπαραμετροποίηση αν οι εσωτερικές παράμετροι είναι γνωστές. Αυτό έχει ως αποτέλεσμα τη μειωμένη σταθερότητα και απαιτεί περισσότερες αντιστοιχίες σημείων από τις απολύτως απαραίτητες.

Το πρόβλημα “στάσης” 3 σημείων, το οποίο είναι η εκτίμηση της “στάσης” της κάμερας από 3 αντιστοιχίες σημείων, γνωστό και ως πρόβλημα προοπτικής 3 σημείων (perspective-3-point problem (P3P)) έχει πάνω από 4 πραγματικές λύσεις στις περισσότερες περιπτώσεις. Για 4 ή περισσότερα σημεία η λύση είναι εν γένει μοναδική, ανεξάρτητα από το αν είναι ομοεπίπεδα, εφ' όσον αυτά δεν έχουν ευθυγραμμιστεί. Ωστόσο, σε δυσμενείς περιπτώσεις μπορεί να υπάρχουν πολλές λύσεις ανεξάρτητα του πόσα σημεία παρέχονται.

Διάφορες προσεγγίσεις για το πρόβλημα P3P έχουν προταθεί στην κοινότητα Computer Vision. Προσπαθούν να υπολογίσουν τις αποστάσεις  $x_i = \|CM_i\|$  ανάμεσα στο κέντρο της κάμερας  $c$  και τα 3D σημεία  $M_i$ , από περιορισμούς που δίνονται από τα τρίγωνα  $CM_iM_j$ . Καθώς τα  $x_i$  είναι γνωστά, τα  $M_i$  εκφράζονται στο πλαίσιο της κάμερας σαν  $M_i^c$ . Τότε το  $[R / t]$  είναι η μετατόπιση που ευθυγραμμίζει τα σημεία  $M_i$  στο  $M_i^c$  και η λύση μπορεί να βρεθεί χρησιμοποιώντας quaternions ή SVD ανάλυση. Η επίλυση για  $x_i$  ουσιαστικά σημαίνει την εύρεση των ριζών ενός πολυωνύμου τετάρτου βαθμού.

Για να αφαιρεθεί η ασάφεια, μία ακόμα αντιστοιχία απαιτείται. Μια απλή προσέγγιση είναι να επιλύσουμε το πρόβλημα P3P για υποσύνολα των τριών από τα τέσσερα σημεία και να διατηρηθεί η κοινή λύση.

Το POSIT είναι ένας ακόμα δημοφιλής τρόπος να λυθεί το πρόβλημα εκτίμησης θέσης για  $n \geq 4$ . Περιλαμβάνει πρώτα τον υπολογισμό μιας προσεγγιστικής λύσης υποθέτοντας μια κλιμακωτή ορθογραφική προβολή για το μοντέλο κάμερας, το οποίο σημαίνει ότι μια πρώτη εκτίμηση της “στάσης” της κάμερας μπορεί να βρεθεί με την επίλυση ενός γραμμικού συστήματος. Ένα βάρος αποδίδεται σε κάθε σημείο με βάση την εκτιμώμενη “στάση” και εφαρμόζεται στις συντεταγμένες του. Μια νέα προβολή υπολογίζεται από αυτές τις διαβαθμισμένες συντεταγμένες, και η διαδικασία επαναλαμβάνεται μέχρι τη σύγκλιση. Αυτή η μέθοδος είναι πολύ απλή να εφαρμοστεί, αλλά ευαίσθητη στο θόρυβο.



Μια άλλη προσέγγιση είναι για την περίπτωση  $n \geq 5$ , η οποία περιλαμβάνει την επίλυση ενός γραμμικού συστήματος της μορφής  $AX=0$ , όπου  $X$  είναι ένα διάνυσμα φτιαγμένο από τις πρώτες δυνάμεις της απόστασης  $\|CM\|$  μεταξύ του κέντρου της κάμερας  $C$  και ενός από τα 3D σημεία  $M$  και το  $A$  φτιάχνεται από τους συντελεστές πολυωνύμων τετάρτου βαθμού. Και αυτή η προσέγγιση δεν εκφυλίζεται για την ομοεπίπεδη περίπτωση.

#### Εκτίμηση "στάσης" από ένα τρισδιάστατο επίπεδο

Η "στάση" της κάμερας μπορεί να υπολογιστεί από μια επίπεδη δομή, όταν είναι γνωστές οι εσωτερικές παράμετροι. Αυτή η ιδιότητα χρησιμοποιείται συχνά στην τρισδιάστατη ανίχνευση επειδή οι προβολές των επίπεδων δομών είναι εύκολο να ανακτηθούν από εικόνες.

Η σχέση μεταξύ ενός τρισδιάστατου επιπέδου και της προβολής εικόνας του μπορούν να αναπαρασταθούν από έναν ομογενή πίνακα  $3 \times 3$ , που λέγεται ομογραφικός πίνακας. Ας θεωρήσουμε το επίπεδο  $Z=0$ . Η έκφραση του ομογραφικού  $H$  που αντιστοιχεί ένα σημείο  $M=(X,Y,0)^T$  σε αυτό το επίπεδο και το αντίστοιχο δισδιάστατο σημείο  $m$  υπό προοπτική προβολή  $P=K[R|t]$  μπορεί να ανακτηθεί γράφοντας:

$$\tilde{m}=P\tilde{M}=K(R^1R^2R^3t) \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = K(R^1R^2t) \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = H \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

όπου  $R^1$ ,  $R^2$  και  $R^3$  η πρώτη, δεύτερη και τρίτη στήλη του πίνακα περιστροφής  $R$ .

Αντιστρόφως, καθώς τα  $H$  και  $K$  είναι γνωστά, η θέση της κάμερας μπορεί να ανακτηθεί. Ο πίνακας  $H$  μπορεί να υπολογιστεί από 4 αντιστοιχίες  $M_i \leftrightarrow m_i$  χρησιμοποιώντας έναν αλγόριθμο DLT παρόμοιο με αυτό που αναλύθηκε παραπάνω. Καθώς  $H_w^t = K(R^1R^2t)$ , το διάνυσμα μετατόπισης  $t$  και οι δύο πρώτες στήλες του πίνακα περιστροφής  $R$  της θέσης της κάμερας μπορούν να ανακτηθούν από το γινόμενο  $K^{-1}H_w^t$ . Η τελευταία στήλη  $R^3$ , δίνεται από το εξωτερικό γινόμενο  $R^1 \times R^2$  καθώς οι στήλες του  $R$  πρέπει να είναι ορθοκανονικές. Η θέση της κάμερας μπορεί να βελτιστοποιηθεί μέσω μη γραμμικής ελαχιστοποίησης όπως περιγράφεται παρακάτω.

*Μη γραμμικό σφάλμα επαναπροβολής*

Βασικό χαρακτηριστικό των προηγούμενων μεθόδων είναι ότι δεν απαιτούν αρχικούς υπολογισμούς και είναι γρήγοροι. Αλλά είναι ευαίσθητοι σε θόρυβο και χάνουν σε ακρίβεια. Συγκεκριμένα, ο αλγόριθμος DLT, παρέχει λύση που ελαχιστοποιεί το αλγεβρικό λάθος, το οποίο είναι καλύτερα να εξεταστεί σε σχέση με ένα γεωμετρικό σφάλμα. Στους αλγορίθμους PnP, δεν είναι προφανές ποιο σφάλμα ελαχιστοποιείται.

Αν οι μετρήσεις  $m_i$  είναι θορυβώδεις, η “στάση” της κάμερας θα μπορούσε να βελτιστοποιηθεί ελαχιστοποιώντας το άθροισμα των σφαλμάτων επαναπροβολής, που είναι η απόσταση στο τετράγωνο ανάμεσα στην προβολή των τρισδιάστατων σημείων και των μετρημένων δύο διαστάσεων συντεταγμένων:

$$[R|t] = \operatorname{argmin}_{[R|t]} \sum_i \operatorname{dist}^2(P\tilde{c}_i, m_i),$$

το οποίο είναι βέλτιστο αν υποτεθεί ότι τα σφάλματα μέτρησης είναι ανεξάρτητα και Γκαουσιανά. Αυτή η ελαχιστοποίηση δεν μπορεί να λυθεί σε κλειστή μορφή, αλλά απαιτεί ένα επαναληπτικό σχήμα βελτιστοποίησης. Ένα τέτοιο σχήμα απαιτεί συνήθως μια αρχική εκτίμηση “στάσης” που παρέχεται με βάση τις προηγούμενες μεθόδους ή υπό συνθήκες ανίχνευσης, από ένα μοντέλο κίνησης που εφαρμόζεται στην προηγούμενως εκτιμηθείσα “στάση”.

**4.1.5 Τεχνικές βελτιστοποίησης ελαχίστων τετραγώνων**

Οι περισσότεροι τρισδιάστατοι αλγόριθμοι υπολογίζουν την “θέση” της κάμερας χρησιμοποιώντας την τεχνική βελτιστοποίησης ελαχίστων τετραγώνων. Γενικότερα, δηλαδή, η “θέση” της κάμερας μπορεί να βρεθεί ελαχιστοποιώντας το άθροισμα των λαθών  $r_i$

$$p = \operatorname{arg}(\min_p (\sum_i r_i^2)).$$

Μπορεί, επίσης, να γραφτεί και σε μορφή πίνακα

$$p = \operatorname{arg}(\min_p (||f(p) - b||^2)),$$

όπου  $p$  ένα διάνυσμα παραμέτρων που ορίζει την “θέση” της κάμερας,  $b$  ένα διάνυσμα που προέκυψε από μετρήσεις και  $f$  μια συνάρτηση που συνδέει τις μετρήσεις αυτές με την “θέση” της κάμερας.

*Γραμμική μέθοδος ελαχίστων τετραγώνων*

Σε μερικές περιπτώσεις, η συνάρτηση  $f$  είναι γραμμική, οι παράμετροι, δηλαδή, της “θέσης” της κάμερας  $p$  μπορούν να γραφτούν ως άγνωστοι σε ένα σύνολο γραμμικών εξισώσεων και μπορούν να εκφραστούν σε μορφή πίνακα ως  $Ap = b$ . Συνήθως, το μέγεθος του διανύσματος  $b$  είναι μεγαλύτερο από τον αριθμό των παραμέτρων του  $p$ , και το  $p$  μπορεί να υπολογιστεί ως  $p = A^+ b$ , όπου

$$A^+ = (A^T A)^{-1} A^T \text{ είναι ο ψευδός αντίστροφος του } A.$$

Εδώ οι μετρήσεις έχουν όλες την ίδια επιρροή, αλλά κάποια μπορεί να θέλει να συνεισφέρει με διαφορετικό βάρος, για παράδειγμα με βάση την ποιότητα της. Σε αυτήν την περίπτωση πρόκειται για πρόβλημα ελαχίστων τετραγώνων με βάρη και λύνεται

$$p = (A^T W A)^{-1} A^T W b,$$

όπου  $W$  ένας πίνακας βαρών που συνήθως είναι διαγώνιος.

*Αλγόριθμοι βελτιστοποίησης βασιζόμενοι στον Newton*

Όταν η συνάρτηση  $f$  δεν είναι γραμμική, που τις περισσότερες φορές δεν συμβαίνει εξαιτίας της μη γραμμικής συμπεριφοράς της προοπτικής προβολής, χρησιμοποιείται ο αλγόριθμος των Gauss-Newton ή των Levenberg-Marquardt. Και οι δύο αλγόριθμοι ξεκινούν από μία αρχική εκτίμηση για το ελάχιστο σημείο  $p_0$  και την ανανεώνουν επαναληπτικά :

$$p_{i+1} = p_i + \Delta_i,$$

όπου το βήμα  $\Delta_i$  υπολογίζεται ανάλογα με την μέθοδο που χρησιμοποιείται. Στον αλγόριθμο του Gauss-Newton το  $\Delta_i$  επιλέγεται έτσι ώστε να ελαχιστοποιεί το σφάλμα στην επανάληψη  $i+1$  και υπολογίζεται εκτιμώντας την  $f$ . Γράφοντας

$f(p_i + \Delta) \approx f(p_i) + J\Delta_i$ , όπου  $J$  ο Ιακωβιανός πίνακας της  $f$  στο σημείο  $p_i$  οδηγούμαστε στη λύση

$$\Delta_i = \arg \min ||f(p_i + \Delta) - b|| = \arg \min ||f(p_i) + J\Delta - b|| = \arg \min ||\varepsilon_i + J\Delta|| = -J^+ + \varepsilon_i,$$

όπου  $\varepsilon_i = f(p_i) - b$  προσδιορίζει το σφάλμα στην επανάληψη  $i$  και  $J^+$  είναι ο ψευδός αντίστροφος του  $J$ . Επομένως μπορούμε να γράψουμε

$$\Delta_i = -(J^T J)^{-1} J^T \varepsilon_i.$$

Στον αλγόριθμο του Levenberg-Marquardt (LM) το  $\Delta_i$  υπολογίζεται με ελαφρώς διαφορετικό τρόπο και ισούται με

$$\Delta_i = -(J^T J + \lambda I)^{-1} J^T \varepsilon_i.$$

Ο επιπλέον όρος  $\lambda I$  σταθεροποιεί τη συμπεριφορά του αλγορίθμου Gauss-Newton. Εάν η τιμή του  $\Delta_i$  οδηγεί σε μείωση του σφάλματος γίνεται αποδεκτή και η τιμή του  $\lambda$

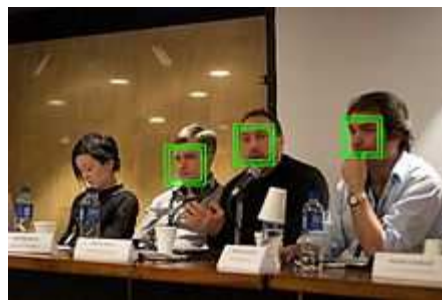
μειώνεται. Διαφορετικά η τιμή του  $\lambda$  αυξάνεται και υπολογίζεται νέα τιμή για το  $\Delta_i$ . Η διαδικασία αυτή επαναλαμβάνεται μέχρι να βρεθεί τιμή για το  $\Delta_i$  που να ελαχιστοποιεί την συνάρτηση εισόδου κάθε φορά. Όταν το  $\lambda$  είναι μικρό ο αλγόριθμος συμπεριφέρεται όπως και ο αλγόριθμος των Gauss-Newton. Αντίθετα, όταν το  $\lambda$  είναι μεγάλο συμπεριφέρεται σαν ένας συμβατικός αλγόριθμος gradient descent προκειμένου να πετύχει το επιθυμητό αποτέλεσμα. [3]

## 4.2 Face tracking

### 4.2.1 Εισαγωγή

Με τον όρο face Detection (αναγνώριση προσώπου/ων) εννοούμε την τεχνολογία που αυτόματα προσδιορίζει τη θέση και τα σχετικά μεγέθη ανθρώπινων προσώπων σε τυχαία δείγματα -ψηφιακών- εικόνων. Αναγνωρίζει τα χαρακτηριστικά των προσώπων αγνοώντας άλλα στοιχεία, όπως πράγματα, φυτά, δένδρα, κτήρια κλπ. Μπορεί να θεωρηθεί ως παρακλάδι της Object Class Detection, της αντικειμενοστραφούς αναγνώρισης δηλαδή, εκεί όπου μπαίνει σαν σκοπός η εύρεση των θέσεων των αντικειμένων που ανήκουν σε μια συγκεκριμένη κατηγορία.

Η υλοποίηση αυτής της θαυμαστής τεχνολογίας ήλθε σαν παρεπόμενο πολυετούς μαθηματικής έρευνας, εφαρμογών θεωρίας νευρωνικών δικτύων και προηγμένων αλγορίθμων που εκπόνησαν επιστήμονες όπως οι Haar, Viola και Jones. Οι παλιότεροι αλγόριθμοι εστίαζαν στην αναγνώριση μπροστινών ανθρώπινων προσώπων, ενώ οι πιο καινούριοι αλγόριθμοι προσπαθούν να επιλύσουν το πιο γενικό και δύσκολο πρόβλημα της πολλαπλής αναγνώρισης προσώπων. Επιθυμούν, δηλαδή, την αναγνώριση προσώπων στραμμένων κατά μήκος του άξονα που ενώνει το πρόσωπα με τον παρατηρητή ή βρίσκονται κάθετα στον άξονα αυτό. Η θεωρία και πράξη της αναγνώρισης προσώπων χρησιμεύει στην βιομετρική, τα συστήματα ασφαλείας (παρακολούθηση με βιντεοκάμερες) αλλά και στις ψηφιακές φωτογραφικές μηχανές. [26, (όρος face detection)]



Εικόνα 43 : Αναγνώριση προσώπων με την χρήση της OpenCV

Σε πολλά συστήματα αναγνώρισης προσώπων, η είσοδος είναι μία φωτογραφία ή ένα βίντεο που περιέχει ένα ή περισσότερα πρόσωπα. Η ανίχνευση προσώπου μπορεί να πάρει διάφορες κατευθύνσεις ανάλογα με την μέθοδο που χρησιμοποιείται, ανίχνευση κεφαλιού, ανίχνευση χαρακτηριστικών, ανίχνευση βασιζόμενη σε εικόνες ή ανίχνευση βασιζόμενη σε μοντέλα. Η έξοδος ενός συστήματος ανίχνευσης προσώπου μπορεί να είναι είτε η θέση, δύο διαστάσεων, του προσώπου στην εικόνα/βίντεο (δισδιάστατη ανίχνευση), είτε η τρισδιάστατη πόζα του προσώπου (τρειςδιάστατη ανίχνευση), είτε η θέση των χαρακτηριστικών του προσώπου. Κάποιοι άλλοι ανιχνευτές δίνουν ως έξοδο και άλλες παραμέτρους που σχετίζονται με τον φωτισμό ή την έκφραση. Η μεγαλύτερη πρόκληση που αντιμετωπίζουν τα συστήματα είναι η ακρίβεια τους (όταν αλλάζει πόζα το πρόσωπο), οι διαφορετικές συνθήκες φωτισμού, οι περιορισμοί του προσώπου (όπως το χρώμα δέρματος) και ο “συνωστισμός” στις εικόνες/βίντεο με αποτέλεσμα να είναι δύσκολη η διάκριση του προσώπου από τα άλλα αντικείμενα.

Υπάρχουν διάφοροι τρόποι κατηγοριοποίησης των αλγορίθμων ανίχνευσης προσώπων. Ένας από αυτούς βασίζεται στο εάν ολόκληρο το πρόσωπο ανιχνεύεται σαν μια μοναδική οντότητα (μερικές φορές χρησιμοποιείται και ο όρος ανίχνευση κεφαλιού ή head tracking) ή αν ανιχνεύονται ξεχωριστά ορισμένα χαρακτηριστικά του. Μερικές φορές χρησιμοποιείται και συνδυασμός αυτών. Μία άλλη μέθοδος κατηγοριοποίησης βασίζεται στο εάν η ανίχνευση γίνεται σε μια δισδιάστατη εικόνα ή σε ένα τρισδιάστατο περιβάλλον. Για την πρώτη, η έξοδος (συνολική θέση κεφαλιού ή θέση ξεχωριστών χαρακτηριστικών) είναι μια περιοχή στην δισδιάστατη εικόνα και δεν περιέχει πληροφορίες σχετικά με τυχόν αλλαγές στην θέση του κεφαλιού. Τέτοιες μέθοδοι δεν είναι συνήθως ακριβείς αλλά είναι πιο εύκολες στην χρήση. Εναλλακτικά, οι τρισδιάστατες μέθοδοι ανίχνευσης, οι οποίες λειτουργούν με βάση κάποιο μοντέλο σε κάθε βίντεο, μπορούν να εκτιμήσουν και την πόζα του προσώπου. Ωστόσο, έχουν μεγαλύτερες υπολογιστικές απαιτήσεις. Εξάλλου, πολλοί αλγόριθμοι είναι σε θέση, πλέον, να αντιμετωπίζουν τέτοιες καταστάσεις, όπως αλλαγές φωτισμού και παραμορφώσεις προσώπων.

#### 4.2.2 Βασικό μαθηματικό υπόβαθρο

Έστω  $p \in \mathbb{R}^p$  ένα διάνυσμα που αντιπροσωπεύει την έξοδο ενός ανιχνευτή προσώπου. Ορίζουμε μία συνάρτηση  $f: \mathbb{R}^2 \times \mathbb{R}^p \rightarrow \mathbb{R}^2$ , η οποία παίρνει ένα pixel

$v \in \mathbb{R}^2$  την χρονική στιγμή  $(t - 1)$  και το μετατρέπει σε  $f(v,p)$  την χρονική στιγμή  $t$ . Για έναν δισδιάστατο ανιχνευτή, αυτή η συνάρτηση θα μπορούσε να είναι ο μετασχηματισμός μεταξύ δύο εικόνων σε δύο συνεχόμενες χρονικές στιγμές. Για έναν τρισδιάστατο ανιχνευτή που βασίζεται σε μοντέλα, θα μπορούσε να είναι μία διαδικασία rendering του αντικειμένου με μια συγκεκριμένη πόζα  $p$  σε καθορισμένο χρονικό πλαίσιο και σύμφωνα με τις συντεταγμένες του pixel  $v$ . Δίνεται μια εικόνα  $I(v)$ , γίνεται προσπάθεια “ευθυγράμμισης” της με το αποτέλεσμα και αποκτούμε το

$p^* = \arg(\min(g(f(v,p) - I(v))))$  που δηλώνει το εκτιμώμενο διάνυσμα για την εικόνα εισόδου.

Η προσέγγιση αυτή θυμίζει την πολύ γνωστή ιδέα ανίχνευσης των Lucas – Kanade, μία αποδοτική και ακριβής προσέγγιση που κάνει χρήση του αντίστροφου υπολογισμού (inverse compositional approach). Ανάλογα με την επιλογή των  $v$  και  $p$ , η μέθοδος είναι εφαρμόσιμη σε ολόκληρη την εικόνα του προσώπου, σε μια συλλογή ξεχωριστών χαρακτηριστικών ή σε ένα τρισδιάστατο μοντέλο προσώπου. Η συνάρτηση κόστους  $g$  υλοποιείται συνήθως ως μια νόρμα  $L_2$ , δηλαδή το άθροισμα των τετραγώνων των λαθών σε μια περιοχή ενδιαφέροντος. Συνήθως, ένας ανιχνευτής προσώπου παρουσιάζεται ως πρόβλημα ελαχιστοποίησης ελαχίστων τετραγώνων (least-squares optimization problem).

Ας υποθέσουμε ότι έχουμε να αντιμετωπίσουμε το πρόβλημα της αλλαγής,  $\Delta p_t = m_t$ , στο διάνυσμα εισόδου μεταξύ δύο συνεχόμενων πλαισίων,  $I_t(v)$  και  $I_{t-1}(v)$ . Το πρόβλημα ορίζεται ως  $m_t^* = \arg(\min_m(\sum(f(v, p_{t-1}^* + m) - I_t(v))^2)$  και  $p_t^* = p_{t-1}^* + m_t^*$ . Η βελτιστοποίηση αυτού του προβλήματος επιτυγχάνεται θεωρώντας ότι η παράμετρος  $m$  είναι γνωστή και η εξίσωση έχει υπολογιστεί επανειλημμένα κατά τις αυξήσεις  $\Delta m$  έτσι ώστε  $\sum_v(f(v, p_{t-1}^* + m + \Delta m) - I_t(v))^2$ .

Αν και η ιδέα των αλγορίθμων ανίχνευσης προσώπων είναι απλή, η πρόκληση που αντιμετωπίζουν είναι να εκτελεστεί η βελτιστοποίηση αποδοτικά και με μεγάλη ακρίβεια. Η συνάρτηση  $f$  θα είναι γενικά μη γραμμική, καθώς περιλαμβάνει πληροφορίες από την προβολή της κάμερας, την τρισδιάστατη πόζα του μοντέλου, την επίδραση του φωτισμού, την αντανάκλαση της επιφάνειας και άλλους παράγοντες.

Το πρόβλημα που αντιμετωπίζουν οι περισσότεροι αλγόριθμοι είναι η ανάγκη για μια καλή αρχικοποίηση στην διαδικασία ανίχνευσης. Δυστυχώς η συνάρτηση  $f$  είναι δύσκολο να οριστεί ακριβώς σε μερικά στιγμιότυπα. Τέτοια παραδείγματα είναι η

μερική αφαίρεση χαρακτηριστικών του προσώπου, παραμορφώσεις λόγω αλλαγών έκφρασης και οι διαφοροποιήσεις που δημιουργεί το φως ακόμη και η παραγωγή σκιών. Πρέπει να δοθεί ιδιαίτερη σημασία ώστε να αντιμετωπιστούν οι καταστάσεις αυτές, διαφορετικά η βελτιστοποίηση της συνάρτησης κόστους θα δώσει λάθος αποτέλεσμα.

Ένα ακόμη σημαντικό ζήτημα στην σχεδίαση των αλγορίθμων είναι η ταχύτητα υπολογισμού. Μέθοδοι τοπικής βελτιστοποίησης (local optimization methods), όπως η διαδικασία gradient descent μπορούν να δώσουν καλό αποτέλεσμα εφόσον το αρχικό σημείο είναι κοντά στην επιθυμητή λύση. Ωστόσο, η διαδικασία είναι λιγάκι αργή αφού απαιτεί επαναυπολογισμό των παραγόντων σε κάθε επανάληψη. Πρόσφατα, έχει προταθεί μία αποδοτική και ακριβής μέθοδος εκτέλεσης της βελτιστοποίησης χρησιμοποιώντας την προσέγγιση αντίστροφου υπολογισμού που δεν απαιτεί ίδιους υπολογισμούς σε κάθε βήμα. Εκτιμήσεις έχουν δείξει ότι σε βίντεο πραγματικού χρόνου, η προσέγγιση αυτή οδηγεί σε αποτέλεσμα με καθυστέρηση μόνο μερικών κλασμάτων του δευτερολέπτου. [8]

### 4.3 Eye tracking

Η ανίχνευση κίνησης του ματιού, ή αλλιώς eye tracking, είναι η διαδικασία παρακολούθησης του βλέμματος (“προς τα πού κοιτάζουμε”) ή της κίνησης του ματιού αναφορικά με το κεφάλι. Ανιχνεύονται, δηλαδή, οι κινήσεις του ματιού του χρήστη και παρατηρούνται οι κόρες των ματιών του όσο ο ίδιος κοιτά σε κάτι συγκεκριμένο. Με αυτόν τον τρόπο, μπορούμε να καταλάβουμε τι παρακολουθεί κάποιος. Στην πράξη, όμως, αυτές οι μετρήσεις είναι δύσκολο να πραγματοποιηθούν και απαιτούν όργανα υψηλής ακρίβειας αλλά και προσεκτική ανάλυση δεδομένων. Επομένως, ο υπολογισμός της θέσης ή της κίνησης των ματιών πραγματοποιείται με την βοήθεια συσκευών γνωστών και ως ανιχνευτές ματιού (eye trackers). Χρησιμοποιούν υπέρυθρη ακτινοβολία και εντοπίζουν την κόρη του ματιού αντανακλώντας φως από τον αμφιβληστροειδή χιτώνα του ματιού ή ανιχνεύουν μέσω webcam κάνοντας ανάλυση της εικόνας του βίντεο.

Μια μεγάλη ποικιλία εφαρμογών χρησιμοποιεί τεχνικές ανίχνευσης κίνησης ματιού, όπως για παράδειγμα οι έρευνες για marketing, αλληλεπίδραση ανθρώπου - υπολογιστή αλλά και ιατρικές έρευνες (κυρίως νευρολογικές διαγνώσεις). Κάποιες εφαρμογές περιλαμβάνουν την κίνηση ματιού για διάβασμα γλωσσικών κειμένων,

μουσικής καθώς και για την αναγνώριση συγκεκριμένων προσώπων. Με την σωστή ιδέα δεν υπάρχει όριο στις εφαρμογές της ανίχνευσης των ματιών. Προς το παρόν, η κυριότερη χρήση της ανίχνευσης είναι ακαδημαϊκή έρευνα, για παράδειγμα στην γνωστική επιστήμη (cognitive science), στην ψυχολογία και στην φαρμακευτική, έρευνα αγοράς και χρηστικές μελέτες, όπως εκτιμήσεις διαφημίσεων ή σχεδιασμού πακέτων λογισμικού αλλά και χρηστικότητας ιστοσελίδων. Η ανίχνευση ματιών μπορεί, επίσης, να χρησιμοποιηθεί και για αλληλεπίδραση – οι άνθρωποι μπορούν να χειριστούν έναν υπολογιστή και να προκαλέσουν γεγονότα απλά κοιτώντας τον. Ο έλεγχος του ματιού μπορεί να χρησιμοποιηθεί μόνος του ως μέσο αλληλεπίδρασης ή σε συνδυασμό με το πληκτρολόγιο, το ποντίκι και την φωνή. Ο έλεγχος του ματιού χρησιμοποιείται ως μέσο επικοινωνίας για ανθρώπους με ειδικές ανάγκες αλλά και σε διάφορες βιομηχανικές και φαρμακευτικές εφαρμογές. [26, (όρος Eye tracking)][24]

#### 4.3.1 Μέθοδοι

Η κύρια πρόκληση για την ανίχνευση του ματιού δεν είναι μόνο η ανάπτυξη των κατάλληλων αλγορίθμων και η εύρεση των κατάλληλων αισθητήρων, που είναι απαραίτητες προϋποθέσεις για ένα υψηλό επίπεδο ακρίβειας, αλλά ο τρόπος με τον οποίο οι χρήστες αλληλεπιδρούν με μια συγκεκριμένη συσκευή ανίχνευσης ματιού. Οι ανιχνευτές ματιών πρέπει να μπορούν να λειτουργούν με όλους τους τύπους ματιών και σε όλες τις συνθήκες όπως η χρήση γυαλιών, φακών επαφής, κινήσεις κεφαλιού και συνθήκες φωτισμού. Οι χρήστες πρέπει, επίσης, να έχουν την δυνατότητα να διατηρούν προσωπικές ρυθμίσεις και να κοιτούν πέρα του ανιχνευτή ματιού χωρίς να υπάρχει ανάγκη διαρρύθμισης (calibration) κάθε φορά από την αρχή. [24]

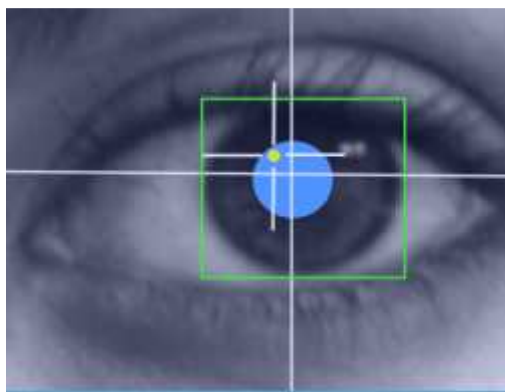
Υπάρχουν πολλές μέθοδοι για την διεξαγωγή της διαδικασίας ανίχνευσης του ματιού. Η πιο γνωστή από αυτές χρησιμοποιεί ανιχνευτές ματιού που δουλεύουν με βίντεο. Μία κάμερα εστιάζει σε ένα ή και στα δύο μάτια και καταγράφει την κίνηση τους καθώς ο θεατής αντιδρά σε κάποιο οπτικό ερέθισμα. Οι πιο σύγχρονοι ανιχνευτές ματιού δημιουργούν αντιθέσεις για τον εντοπισμό του κέντρου της κόρης του ματιού και χρησιμοποιούν υπέρυθρες (infrared) αλλά και μια άλλη κατηγορία υπέρυθρων (near-infrared) δεσμίδων φωτός για να δημιουργήσουν αντανάκλαση του κερατοειδή χιτώνα του ματιού. Το διάνυσμα μεταξύ των δύο αυτών χαρακτηριστικών μπορεί να χρησιμοποιηθεί για την εύρεση της τομής του βλέμματος με μια επιφάνεια.

Ένα σημαντικό χαρακτηριστικό της ανίχνευσης ματιών είναι το κουτί ανίχνευσης (track box). Πρόκειται για το φανταστικό κουτί, στα πλαίσια του οποίου ο χρήστης μπορεί να



κινεί το κεφάλι του αλλά ταυτόχρονα να είναι ανιχνεύσιμος από τον ανιχνευτή. Με ένα μεγαλύτερο “κουτί ανίχνευσης” ο χρήστης θα έχει περισσότερη ελευθερία κίνησης και θα αποκτήσει καλύτερη εμπειρία.

Χρησιμοποιούνται δύο γενικοί τύποι τεχνικών για την ανίχνευση ματιού : φωτεινή κόρη ματιού (Bright Pupil) και σκοτεινή κόρη ματιού (Dark Pupil). Η διαφορά τους έγκειται στην θέση στην οποία βρίσκεται η πηγή φωτισμού. Εάν η πηγή φωτισμού και το οπτικό μονοπάτι έχουν κοινό άξονα, το μάτι δρα ως συσκευή ανάκλασης (retroreflector), καθώς το φως ανακλάται στον αμφιβληστροειδή χιτώνα δημιουργώντας μια φωτεινή κόρη (φαινόμενο παρόμοιο με το φαινόμενο του κόκκινου ματιού ή αλλιώς red-eye effect). Εάν η πηγή φωτισμού δεν έχει σχέση με το οπτικό μονοπάτι, η κόρη του ματιού εμφανίζεται σκούρα επειδή η αντανάκλαση από τον αμφιβληστροειδή κατευθύνεται μακριά από την κάμερα.



Εικόνα 44 : εντοπισμός της κόρης του ματιού

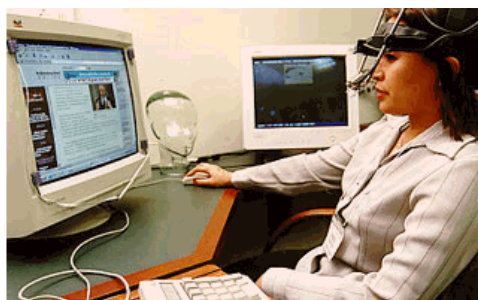
Η τεχνική του Bright Pupil δημιουργεί καλύτερη αντίθεση ίριδας/κόρης ματιού, μετατρέποντας τους ανιχνευτές ματιού σε πιο σταθερά συστήματα, για όλους τους χρωματισμούς της ίριδας, και μειώνοντας την επίδραση της ανάμειξης των βλεφαρίδων και άλλων τέτοιων χαρακτηριστικών που εμποδίζουν στην ορθή ανίχνευση της κίνησης. Επιτρέπει, επίσης, την ανίχνευση σε διάφορες συνθήκες φωτός από τελείως σκοτάδι σε τελείως φωτισμένο χώρο. Όμως, οι τεχνικές αυτές δεν είναι αποτελεσματικές σε υπαίθριους χώρους καθώς οι εξωτερικές πηγές υπεριώθρων παρεμβαίνουν στην ανίχνευση.

Μέχρι πρόσφατα, διαφορετικοί τύποι ματιών απαιτούσαν διαφορετικές τεχνικές ανίχνευσης. Η τεχνική του Dark Pupil δούλευε καλύτερα για ανθρώπους με σκούρα μάτια, ενώ η τεχνική του Bright Pupil δούλευε καλύτερα για παιδιά και ανθρώπους με μπλε μάτια. Πρόσφατα, και οι δύο αυτές οι τεχνικές συνδυάστηκαν για να εξαλείψουν την ανάγκη δύο ξεχωριστών ανιχνευτών ματιών. [26, (όρος Eye tracking)]

### 4.3.2 Τεχνολογίες

Οι τεχνολογίες που χρησιμοποιούνται για την ανίχνευση διαφέρουν αρκετά, μερικές ενσωματώνονται στο κεφάλι (είναι δηλαδή head-mounted), κάποιες απαιτούν το κεφάλι να παραμένει σταθερό και άλλες λειτουργούν απομονωμένα και εντοπίζουν αυτόματα το κεφάλι κατά την κίνηση. Οι περισσότερες δουλεύουν στα 30 Hz τουλάχιστον. Αν και τα 50/60 Hz είναι πιο συνηθισμένα, στις μέρες μας πολλοί ανιχνευτές ματιού που βασίζονται σε κάμερα λειτουργούν στα 240, 350 ή ακόμη και στα 1000/1250 Hz, τα οποία είναι απαραίτητα προκειμένου να “αιχμαλωτιστεί” η λεπτομέρεια μιας πολύ γρήγορης κίνησης του ματιού κατά τη διάρκεια ανάγνωσης ή κατά τη διάρκεια νευρολογικών ερευνών.

Οι head-mounted ανιχνευτές αποτελούνται από ένα κράνος με συρμάτινο πλαίσιο που τοποθετείται στο κεφάλι του χρήστη προκειμένου να σταθεροποιήσει κινήσεις κεφαλιού και ματιού.



Εικόνα 45 : χρήστης με ανιχνευτή head-mounted

Η τεχνική Gaze-detection (ανίχνευση βλέμματος), σε αντίθεση, με την τεχνική head mounted υπάρχει από τα μέσα του 1990 και χρησιμοποιείται κυρίως ως διεπαφή σε πεδία εικονικής πραγματικότητας. Ωστόσο, βελτιώνεται συνεχώς με την ανάπτυξη της τεχνολογίας. [18]

### 4.3.3 Ανιχνεύοντας κινήσεις ματιού

Χρησιμοποιώντας έναν ανιχνευτή ματιού, είναι δυνατόν να καταγράψουμε ακριβώς τι κοιτά ο χρήστης όσο αυτός αλληλεπιδρά με κάποιο οπτικό ερέθισμα. Για παράδειγμα, αν πούμε ότι ένας χρήστης περιηγείται σε μια ιστοσελίδα ψάχνοντας να αγοράσει κάτι από το διαδίκτυο. Με την ανίχνευση του ματιού του μπορούμε να εξακριβώσουμε σε ποιο σημείο ψάχνει ο χρήστης πληροφορίες για το συγκεκριμένο προϊόν, πως συγκρίνει διαφορετικά προϊόντα, και τι του τράβηξε ιδιαίτερα την προσοχή. Το σημείο του

βλέμματος του χρήστη καταγράφεται σε ένα βίντεο όσο ο χρήστης βρίσκεται στο διαδίκτυο και αλληλεπιδρά. Με αυτόν τον τρόπο, οι κατασκευαστές ιστοσελίδων μπορούν να επεξεργαστούν την διαμόρφωση της ιστοσελίδας τους μέσα από τα μάτια του χρήστη. Είναι, επίσης, πιθανό να αντιληφθούμε αντικειμενικά εάν ο χρήστης έχει δυσκολία στον εντοπισμό πληροφοριών, εάν χάνει κρίσιμες πληροφορίες ή έχει γενικότερα κάποιο πρόβλημα στην περιήγηση.

Μεγάλη πρόοδος έχει σημειωθεί στο υλικό και στην ανάπτυξη αλγορίθμων επεξεργασίας εικόνων για τους ανιχνευτές ματιών. Αρκετές ιδιωτικές εταιρίες παράγουν τώρα ανιχνευτές ματιών και σχετικό λογισμικό έχοντας περιορισμένο καταναλωτικό κοινό : έρευνα, ανάλυση και τεχνολογία για ανθρώπους με ειδικές ανάγκες. Όπως συμβαίνει σε όλους τους τομείς της τεχνολογίας, οι δημιουργίες που ήταν κάποτε ογκώδεις, ακριβές και θεωρούνταν υψηλής τεχνολογίας, μερικά χρόνια αργότερα μπορούν να βρεθούν σε μεγάλες ποσότητες και να γίνουν φθηνά καταναλωτικά αγαθά. Στις μέρες μας είναι πιθανό να χρησιμοποιήσουμε τεχνολογία, που υπάρχει εδώ και καιρό, και πηγές ανοιχτού κώδικα προκειμένου να δημιουργήσουμε χαμηλού κόστους ανιχνευτή και να κάνουμε την τεχνολογία αυτή αρκετά πιο προσιτή. Παρόλα αυτά, η ποιότητα και η σταθερότητα αυτών των συστημάτων δεν συγκρίνεται με αυτά που δημιουργούνται για εμπορικούς σκοπούς. Μερικά χρόνια πριν, τα περισσότερα συστήματα χρησιμοποιούσαν ειδικό υλικό (hardware) για επεξεργασία εικόνων, σήμερα όμως η επεξεργαστική ισχύς ενός κανονικού υπολογιστή είναι αρκετή για τους αλγορίθμους ανάλυσης εικόνων στην διαδικασία ανίχνευσης κινήσεων ματιών. Ατενίζοντας τον ορίζοντα, η εξέλιξη του ψηφιακού βίντεο υψηλής ανάλυσης στην καταναλωτική αγορά διευκολύνει την περαιτέρω ανάπτυξη ανιχνευτών ματιών χαμηλού κόστους.

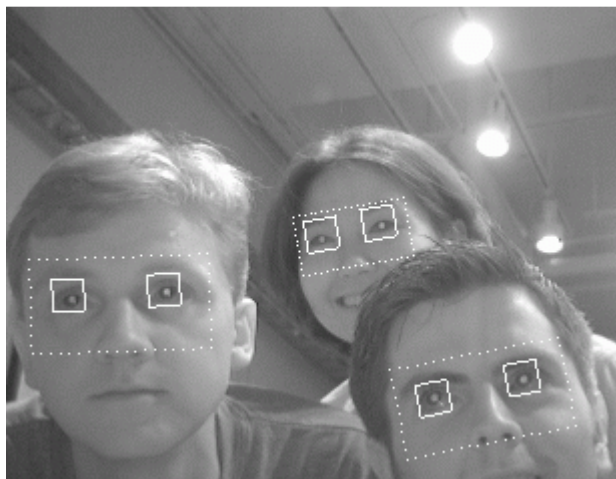
Η νέα γενιά των συστημάτων που βασίζονται σε ανιχνευτές ματιών δείχνει πόσο προσιτή έχει γίνει η τεχνολογία. Μπορεί να μην είναι τόσο γρήγορα ή ακριβά όσο ο εξοπλισμός ενός εργαστηρίου αλλά ένα μεγάλο ποσοστό χρηστών μπορούν να τα ρυθμίσουν και να τα χρησιμοποιήσουν μέσα σε λίγα δευτερόλεπτα. Επιπλέον, τα συστήματα αυτά κρατούν πληροφορίες για την τοποθεσία του προσώπου και επιτρέπουν περιορισμένες κινήσεις του κεφαλιού, ενώ τα πιο γρήγορα συστήματα απαιτούν από τους συμμετέχοντες να ξεκουράζουν το πηγούνι τους στην συσκευή προκειμένου να σταθεροποιήσουν την εικόνα του ματιού.

Η ανίχνευση ματιού δουλεύει αντανάκλωντας άορατο υπέρυθρο φως στα μάτια, καταγράφοντας το μονοπάτι αντανάκλασης με ένα σύστημα αισθητήρων, και έπειτα υπολογίζοντας το ακριβές σημείο του βλέμματος χρησιμοποιώντας ένα γεωμετρικό μοντέλο. Μόλις καθοριστεί το σημείο αυτό, μπορεί να οπτικοποιηθεί και να προβληθεί

σε κάποια οθόνη. Το σημείο του βλέμματος μπορεί να χρησιμοποιηθεί, επίσης, και για τον έλεγχο και την αλληλεπίδραση με διάφορα μηχανήματα. Αυτή η τεχνική αναφέρεται σαν έλεγχος ματιού.

Τα συστήματα που βασίζονται σε ανιχνευτές ματιών συνήθως αποτελούνται από μια κάμερα ικανή να αιχμαλωτίζει εικόνες στο φάσμα υπέρυθρου φωτός. Η κάμερα τοποθετείται κάτω από την οθόνη και περιβάλλεται από ένα σύνολο διόδων υπέρυθρου φωτός (infrared light emitting diodes ή IR LEDs). Η κάμερα, συνήθως ανάλυσης 1-2 Megapixel, αιχμαλωτίζει το πρόσωπο του χρήστη κάθε 30-120 milliseconds, χρόνος που εξαρτάται από το σύστημα. Συγκριτικά, το ανώτερο όριο υψηλής ταχύτητας του εργαστηρίου αιχμαλωτίζει εικόνες 1000-1500 Megapixels το δευτερόλεπτο. Το προφανές πλεονέκτημα των συστημάτων είναι ότι επιτρέπει έναν συγκεκριμένο αριθμό κινήσεων του κεφαλιού, ενώ τα υψηλής ταχύτητας συστήματα βασίζονται σε μια συγκεκριμένη θέση, τοποθετώντας το πηγούνι σε ειδική τοποθεσία. Χωρίς αμφιβολία, τα συστήματα υψηλής ταχύτητας έχουν μεγάλη ακρίβεια αλλά αυτό δεν αποτελεί λύση για καθημερινή αλληλεπίδραση με την βοήθεια του βλέμματος. Για να επιτευχθεί υψηλής ακρίβειας ανίχνευση ματιού και οι δύο αυτοί τύποι χρησιμοποιούν βιομηχανικές κάμερες CCD αφού οι πιο πολλές καταναλωτικές κάμερες καταγράφουν με μικρότερη ανάλυση και μικρότερο ρυθμό πλαισίων (15-30 εικόνες το δευτερόλεπτο). Ωστόσο, με την βελτίωση των καταναλωτικών υψηλής ανάλυσης εφαρμογών, οι κάμερες που υποστηρίζουν 1920x1080 στα 25 πλαίσια ανά δευτερόλεπτο γίνονται ολοένα και πιο προσιτές. Υποθετικά, το γεγονός αυτό θα προκαλούσε μια κατάσταση όπου θα παράγονταν μαζικά συσκευές ανίχνευσης ματιών με αρκετές πληροφορίες για την αλληλεπίδραση με βάση το βλέμμα και τιμή υλικού που ολοένα και θα έπεφτε.

Υπάρχουν αρκετά στάδια στην ανίχνευση του ανθρώπινου ματιού. Μέχρι να αιχμαλωτιστεί το πρόσωπο διεξάγονται βήματα επεξεργασίας εικόνων. Το πρόσωπο ανιχνεύεται από τα χαρακτηριστικά του όπως το στόμα, η μύτη, τα μάτια και τα λοιπά. Έπειτα, δημιουργείται μια περιοχή ενδιαφέροντος γύρω από τα μάτια που είναι η εικόνα που χρειάζεται περαιτέρω επεξεργασία. Σήμερα, όπως έχει ήδη ειπωθεί, οι περισσότεροι ανιχνευτές ματιών βασίζονται στην μέθοδο αντανάκλασης του κερατοειδή (corneal reflection method), όπου υπέρυθρο φως στέλνεται κατευθείαν στο πρόσωπο. Η αντανάκλαση που δημιουργεί το φως στο μάτι χρησιμοποιείται για τον υπολογισμό του διανύσματος του βλέμματος σε σχέση με την λάμψη και την θέση της κόρης του ματιού. Το φάσμα υπέρυθρου φωτός που χρησιμοποιείται εμποδίζει το φως που φτάνει στο πρόσωπο να γίνει οπτικά αντιληπτό. Ένα εμφανές πλεονέκτημα χρήσης τέτοιου συστήματος είναι ότι ο χρήστης δεν χρειάζεται να φορά συγκεκριμένο εξοπλισμό ή να τοποθετήσει στο κεφάλι του συγκεκριμένη συσκευή.



*Εικόνα 46* : εντοπισμός των ματιών χρηστών όσο αυτοί κοιτούν την οθόνη του υπολογιστή τους

Οι περισσότεροι ανιχνευτές που βασίζονται στην αντανάκλαση υπέρυθρου φωτός είναι ευαίσθητοι σε μεγάλες ποσότητες ηλιακού φωτός καθώς έχει αποδειχθεί ότι αναμειγνύεται με τις αντανάκλασεις του κερατοειδούς χιτώνα. Προκαλεί προβλήματα στην αλληλεπίδραση με βάση το βλέμμα σε υπαίθριο χώρο, όπου μεγάλες ποσότητες ηλιακού φωτός καλύπτουν το υπέρυθρο φως που εκπέμπεται από τα LEDs των ανιχνευτών ματιών. Οι κάμερες στα συστήματα καλύπτουν ένα συγκεκριμένο οπτικό πεδίο μπροστά από τον ανιχνευτή ματιού. Το πλάτος και το βάθος του κουτιού ανίχνευσης θέτει περιορισμό στην ευελιξία στάσης που μπορεί να πάρει ο χρήστης. Τα περισσότερα συστήματα είναι ανεκτικά, μέχρι κάποιον βαθμό, όσον αφορά την κίνηση του κεφαλιού και μπορούν συνεχώς να ανιχνεύουν την θέση του κεφαλιού και των ματιών.

Προτού χρησιμοποιήσουμε έναν ανιχνευτή ματιού, αυτός πρέπει να διαρρυθμιστεί σε σχέση με την οθόνη. Παρουσιάζοντας ένα σύνολο σημείων στην οθόνη μπορεί να εντοπιστεί συσχέτιση μεταξύ της θέσης της κόρης του ματιού και των συντεταγμένων X και Y της οθόνης. Χρησιμοποιώντας ένα μεγαλύτερο πλήθος σημείων κατά την διαρρύθμιση πετυχαίνεται μεγαλύτερη ακρίβεια στον προσδιορισμό της θέσης του βλέμματος.

Υπάρχουν, ωστόσο, παράγοντες που με την πάροδο του χρόνου επηρεάζουν την ακρίβεια της αρχικής διαρρύθμισης. Ένα θέμα που προκύπτει είναι οι ιδιότητες του ματιού που αλλάζουν όταν το μάτι γίνεται πιο υγρό μετά από πολύωρη ενασχόληση με την οθόνη του υπολογιστή. Αυτό επηρεάζει την αντανάκλαση του κερατοειδούς χιτώνα. Επιπλέον, αλλαγές στη στάση και στην απόσταση από την κάμερα στον χρόνο μειώνουν την ποιότητα της αρχικής διαρρύθμισης. Αυτοί οι παράγοντες δημιουργούν μια

απόκλιση (offset) στην διαρρύθμιση που γίνεται περισσότερο εμφανής στις άκρες της οθόνης του υπολογιστή. Η ικανότητα των ανιχνευτών ματιών να αντισταθμίζουν τους παράγοντες αυτούς είναι σημαντική για την εμπειρία αλληλεπίδρασης. [13]

#### 4.3.4 Μελλοντική αξία

Η ανίχνευση ματιού έχει αναπτυχθεί εδώ και έναν αιώνα, αλλά τα τελευταία πέντε χρόνια η βελτίωση της τεχνολογίας έχει προσφέρει νέες πιο εκπληκτικές δυνατότητες. Η σύγχρονη ανίχνευση ματιού μπορεί να χρησιμοποιηθεί όχι μόνο σε εργαστηριακές έρευνες αλλά και σε σπίτια, σχολεία και επιχειρήσεις, οπουδήποτε δηλαδή βοηθά έρευνες, αναλύσεις και αλληλεπιδράσεις με υπολογιστές, φίλους και οικογένεια.

Οι σκληρές και πολύπλοκες τεχνικές ανίχνευσης ματιού του παρελθόντος έχουν αντικατασταθεί από μεθόδους προσδιορισμένες ξανά με μεγαλύτερη προσοχή και πιο φιλικές προς τον χρήστη, οι οποίες παράγουν αξιόλογα αποτελέσματα και υπόσχονται ακόμη περισσότερα για το μέλλον. Η ανίχνευση αλλά και ο έλεγχος ματιού έχουν απεριόριστη διάρκεια ζωής στο μέλλον. Θα αξιοποιηθούν με διάφορους τρόπους, αδιανόητοι προς το παρόν, σε πολλούς τομείς όπως η αυτοματοποιημένη βιομηχανία, οι φαρμακευτικές έρευνες, η εκπαίδευση αλλά και η προσωπική χρήση υπολογιστών.

Καταλήγοντας, η ανίχνευση ματιού έχει ένα πολλά υποσχόμενο μέλλον. Καθώς βελτιώνεται η τεχνολογία βελτιώνονται κατά συνέπεια και οι εφαρμογές, οι μέθοδοι, οι έρευνες και οι πληροφορίες που σχετίζονται με την ανίχνευση ματιών. Έχει ειπωθεί, άλλωστε, ότι η ανίχνευση κίνησης του ματιού είναι το επόμενο μεγάλο βήμα σε έρευνες για την συμπεριφορά του ανθρώπου, “πολύ κοντά στην εύρεση της σκέψης κάποιου”. Ωστόσο κάποιοι αντιλέγουν υποστηρίζοντας ότι πρόκειται για μια ακριβή διαδικασία που δεν αποκαλύπτει τίποτα περισσότερο απ’ όσα είναι ήδη γνωστά. Συνεχίζουν, λέγοντας πως δεν είναι όλα ρόδινα για την διαδικασία αυτή. Τα δύο μεγαλύτερα πρακτικά προβλήματα είναι η διαρρύθμιση και η πολυπλοκότητα στην καταγραφή και στην ανάλυση πληροφοριών. [17][24]

## 4.4 Hand tracking

### 4.4.1 Εισαγωγή

Το πιο σημαντικό “εργαλείο” του ανθρώπινου σώματος για αλληλεπίδραση είναι τα χέρια. Είναι απαραίτητα σε σχεδόν όλες τις ανθρώπινες δραστηριότητες από απλές χειρονομίες μέχρι πιο πολύπλοκες εργασίες, όπως τον έλεγχο ενός υπολογιστή, την χειραγώγηση πραγματικών και εικονικών αντικειμένων και την πλοήγηση σε ένα εικονικό περιβάλλον. Συνεπώς, η ανίχνευση χεριού θα μπορούσε να εμπλουτίσει πολυάριθμους τομείς. Σε κάποιες εφαρμογές ήδη χρησιμοποιείται με μεγάλη επιτυχία. Σε άλλες, μπορούν να προκύψουν εντυπωσιακά αποτελέσματα ακόμα και με την ανίχνευση ενός μικρού υποσυνόλου των βαθμών ελευθερίας του χεριού. Το χέρι έχει 20 τοπικούς βαθμούς ελευθερίας και 6 καθολικούς βαθμούς. Για παράδειγμα, ένας αλγόριθμος που μπορεί να ταξινομήσει κάποιες χειρονομίες θα μπορούσε να χρησιμοποιηθεί για την κατανόηση της νοηματικής γλώσσας ή σαν συσκευή εισόδου σε έναν υπολογιστή για να αντικαταστήσει το ποντίκι. Η δισδιάστατη θέση του χεριού θα μπορούσε να αντικαταστήσει την κίνηση του ποντικιού, ενώ οι χειρονομίες του χεριού τα πλήκτρα του ποντικιού. Κάποιος θα μπορούσε να χρησιμοποιήσει την ανίχνευση χεριού για να εμπλουτίσει μια παρουσίαση σε “τοίχο” μετακινώντας ή προσθέτοντας απευθείας εικονικά στοιχεία στην παρουσίαση. Ακόμα και για απλές δισδιάστατες αλληλεπιδράσεις, η ανίχνευση χεριού μπορεί να αντικαταστήσει αλληλεπιδράσεις που βασίζονται σε συσκευές. Για παράδειγμα, η ανίχνευση δυο δαχτύλων θα μπορούσε να αντικαταστήσει τα touchpads. Το πληκτρολόγιο θα μπορούσε να αντικατασταθεί από ένα εικονικό πληκτρολόγιο, ανιχνεύοντας τα δάχτυλα και των δύο χεριών. Φυσικά, η αξιοποίηση της ανίχνευσης όλων των βαθμών ελευθερίας θα μπορούσε να συμβάλει σε ακόμα πιο επιδέξιους χειρισμούς και αλληλεπιδράσεις. Η ανίχνευση χεριού θα μπορούσε να γίνει μια ακόμα πιο ισχυρή συσκευή εισόδου σε σχέση με τις κοινές συσκευές εισόδου (π.χ. οθόνες αφής ή το ποντίκι).

Σαν παράδειγμα εικονικής προσομοίωσης, μπορεί να θεωρηθεί ένας μηχανικός που αλληλεπιδρά με την CAD εφαρμογή του σε ένα εικονικό περιβάλλον. Με την ανίχνευση χεριού, θα μπορούσε κουνώντας ελεύθερα τα χέρια του να πλοηγηθεί διαισθητικά, να ελέγξει και να χειριστεί την CAD εφαρμογή του, χωρίς πρόσθετες συσκευές εισόδου. Μία κίνηση, όπως το άνοιγμα της πόρτας ενός αυτοκινήτου, είναι η ίδια ενέργεια όπως στον πραγματικό χρόνο, απλά διαφέρει ως προς την παραδοσιακή αλληλεπίδραση

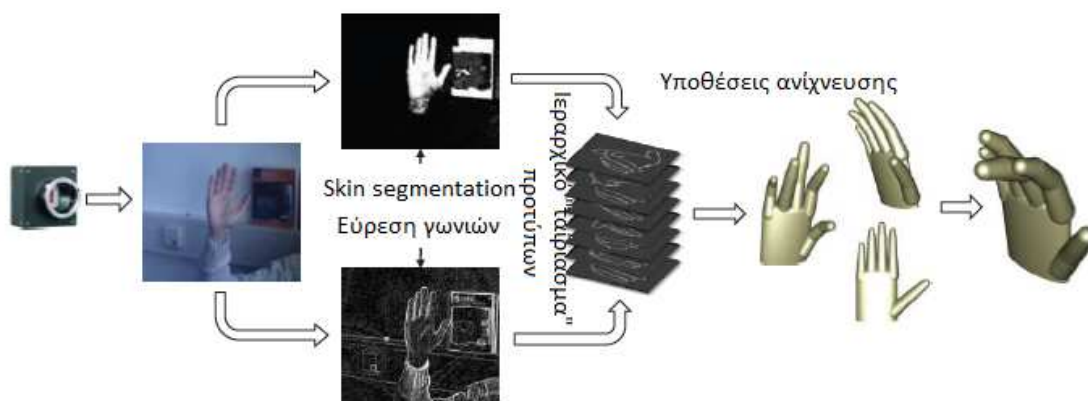
μέσω ποντικιού, όπου θα έπρεπε να εξομοιωθεί μέσω μιας ακολουθίας από “κλικ” ποντικιού.

Μια εντελώς διαφορετική εφαρμογή της ανίχνευσης χεριού είναι τα βίντεο παιχνίδια και τα παιχνίδια σε υπολογιστή. Μπορεί κανείς να φανταστεί έναν σκοποβόλο ο οποίος χρησιμοποιεί το χέρι του για να εστιάσει στον στόχο. Κάτι τέτοιο είναι πολύ πιο διαισθητικό και διαδραστικό από την χρήση του ποντικιού, γιατί θα μπορούσε εύκολα να ελέγξει τους 6 βαθμούς ελευθερίας (μετατόπιση και περιστροφή) που απαιτούνται σε τρισδιάστατο περιβάλλον. Σε ένα παιχνίδι περιπέτειας, ο χρήστης θα μπορούσε να πιάνει και να αφήνει αντικείμενα με πιο διαισθητικό τρόπο σε σχέση με την χρήση συντομεύσεων πληκτρολογίου.

Αυτές είναι μερικές μόνο από τις πολυάριθμες εφαρμογές στις οποίες μπορεί να χρησιμοποιηθεί η ανίχνευση χεριού. Οι περισσότερες από αυτές χρειάζονται ακριβή, σε πραγματικό χρόνο ανίχνευση του χεριού με 26 βαθμούς ελευθερίας, με αποτέλεσμα οι αντίστοιχοι αλγόριθμοι να αποτελούν τεχνολογία ευρείας διάδοσης για αυτό το είδος παραδειγματικής αλληλεπίδρασης. [10]

#### 4.4.2 Η διαδικασία ανίχνευσης χεριού

Η διαδικασία ανίχνευσης του ανθρώπινου χεριού μπορεί να διαιρεθεί σε 4 βασικά στάδια: Το πρώτο είναι η μετατροπή των εικόνων που παράγει η κάμερα σε μορφή RGBA, το δεύτερο η αφαίρεση του “φόντου” (background subtraction), το τρίτο η κατάτμηση δέρματος και το τέταρτο η εφαρμογή της μεθόδου “connected component labelling”.



*Εικόνα 47* : Το χέρι ανιχνεύεται από την κάμερα, τα χαρακτηριστικά εξαγονται, οι πιο πιθανές παράμετροι χεριού υπολογίζονται και τελικά μία παράμετρος επιλέγεται σαν αποτέλεσμα.



Η κάμερα δεν μπορεί να δώσει εικόνες σε μορφή RGB απευθείας. Αυτό είναι ένα χαρακτηριστικό κάθε κάμερας χαμηλού κόστους. Αντίθετα δίνει εικόνες σε μορφή YCbCr 4:2:2, όπου Y είναι η συνιστώσα φωτεινότητας και τα Cb και Cr είναι οι συνιστώσες μπλε και κόκκινης διαφοράς. Το YCbCr συχνά συγχέεται με τη μορφή YUV, αλλά το YUV εφαρμόζεται μόνο σε αναλογικό βίντεο ενώ το YCbCr εφαρμόζεται σε ψηφιακό βίντεο. Οι παρακάτω εξισώσεις χρησιμοποιούνται για την μετατροπή του YCbCr σε RGB:

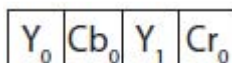
$$R=1.164(Y - 16) + 1.596(Cr - 128)$$

$$G=1.164(Y - 16) - 0.813(Cr - 128) - 0.391(Cb - 128)$$

$$B=1.164(Y - 16) + 2.018(Cb - 128)$$

όπου οι τιμές R,G,B κυμαίνονται από 0 ως 255.

Η μορφή YCbCr 4:2:2 “πακετάρει” δύο pixels εικόνας σε ένα macro-pixel.



Εικόνα 48 : Η μορφή ενός macro-pixel

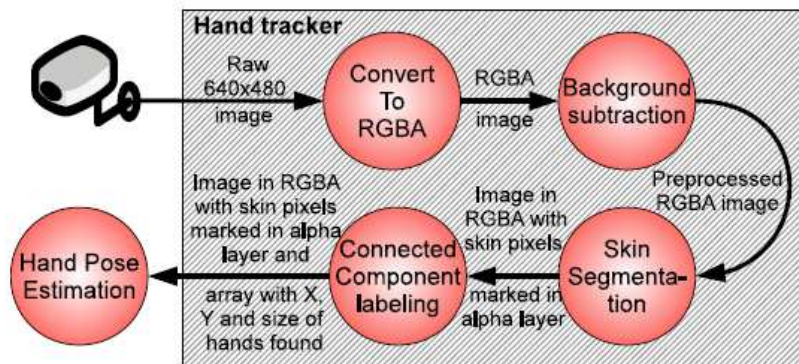
Δυο συνιστώσες φωτεινότητας  $Y_0$  και  $Y_1$  μοιράζονται το ίδιο σύνολο συνιστωσών χρώματος  $Cb_0$  και  $Cr_0$ . Σε αντίθεση με το RGB, το ήμισυ του εύρους ζώνης χρειάζεται για την μεταφορά της εικόνας από την κάμερα στον υπολογιστή. Η συνιστώσα μετατροπής χώρου χρώματος διατηρεί ένα επιπλέον byte για το επίπεδο alpha, το οποίο χρησιμοποιείται για διαφάνεια, αλλά ο ανιχνευτής χεριού χρησιμοποιεί αυτό το επίπεδο για να “μαρκάρει” τα pixels ως δέρμα ή background.

Μετά την μετατροπή της η εικόνα τροφοδοτεί τη διαδικασία αφαίρεσης “φόντου”.

Γενικά, το “φόντο” της εικόνας δεν θα αλλάζει πολύ συχνά, πιθανώς μόνο με τις μεταβαλλόμενες συνθήκες φωτισμού. Όπως αναφέρθηκε, η κάμερα χρησιμοποιεί μια στατική ρύθμιση. Συνεπώς, με τη διαδικασία αφαίρεσης background, γίνεται μια πρόχειρη επιλογή για το ποια pixels απαιτούν περαιτέρω εξέταση. Η διαδικασία αυτή επομένως, συγκρίνει απλά κάθε pixel με το αντίστοιχο pixel στην προηγούμενη αποθηκευμένη εικόνα background. Εάν η διαφορά είναι μικρότερη από ένα ορισμένο όριο T θα υπολογίζεται ως background. Αν το pixel υπολογίζεται ως background η αντίστοιχη θέση στο επίπεδο alpha ορίζεται σε μηδέν, ειδάλλως ορίζεται σε 255. Το πλεονέκτημα αυτής της προσέγγισης είναι ότι η αρχική εικόνα διατηρείται και απαιτείται λίγη επιπλέον μνήμη.

Αφού ολοκληρωθεί η διαδικασία αφαίρεσης background, η εικόνα περνά στο στάδιο κατάτμησης δέρματος (skin segmentation). Σε αυτό το στάδιο, συγκρίνονται μόνο τα pixels που είναι μαρκαρισμένα σαν non-background. Συγκεκριμένα, κάθε non-

background pixel συγκρίνεται με έναν προκαθορισμένο πίνακα και όταν η διαφορά είναι μεγαλύτερη από ένα όριο  $T$ , το pixel θεωρείται ως δέρμα. Αν ένα pixel υπολογίζεται ως δέρμα, η αντίστοιχη τιμή στο επίπεδο alpha τίθεται 255. Μετά την κατάτμηση του δέρματος το τελικό βήμα είναι η εύρεση όλης της περιοχής μαρκαρισμένου δέρματος μέσα στην εικόνα. Χρησιμοποιώντας τη τεχνική connected component labelling οι περιοχές αυτές ανιχνεύονται και τα  $X, Y$ , πλάτος και ύψος των περιοχών, αποθηκεύονται σε πίνακα. Αυτός ο πίνακας μαζί με εικόνα RGBA περνούν στον εκτιμητή θέσης χεριού.



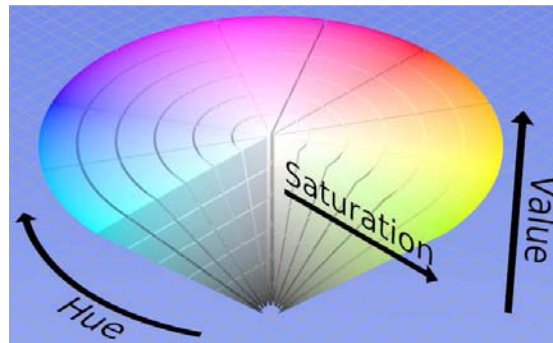
Εικόνα 49 : Τα στάδια της διαδικασίας ανίχνευσης χεριού

### *Skin segmentation*

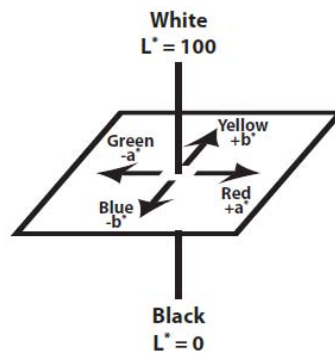
Στο πεδίο της ανίχνευσης χεριού έχουν προταθεί διάφορες μέθοδοι. Μία σχετικά απλή τεχνική είναι η χρήση ιστογραμμάτων για τον προσδιορισμό της κατανομής του χρώματος του δέρματος. Ένα ιστόγραμμα στο πλαίσιο αυτό δείχνει την συχνότητα όλων των χρωμάτων, όπου κάθε χρώμα ανήκει σε μια κατηγορία. Αυτές οι κατηγορίες είναι σειρές χρωμάτων και ονομάζονται κάδοι. Οι περισσότεροι αλγόριθμοι μετατρέπουν κόκκινες, πράσινες και μπλε (RGB) τιμές σε άλλο χώρο χρώματος για την ελαχιστοποίηση της επίδρασης των μεταβαλλόμενων συνθηκών φωτισμού και της πολυμορφίας του χρώματος του δέρματος. Ως επί το πλείστον χρησιμοποιούνται ιστογράμματα δύο διαστάσεων.

Οι χώροι χρώματος που χρησιμοποιούνται περισσότερο είναι η απόχρωση (hue), ο κορεσμός (saturation) και η αξία (value), ή η ένταση (intensity) (HSV/I) και το  $L^*a^*b$ . Το HSV/I περιγράφει τα χρώματα ως σημεία σε έναν κώνο. Η “απόχρωση” αντιπροσωπεύει το καθαρό χρώμα, χωρίς σκιές ή αναμειξίς, και κυμαίνεται από 0 ως 360. Ο “κορεσμός” αναφέρεται στην καθαρότητα του χρώματος και κυμαίνεται από 0 ως 100%. Το τρίτο στοιχείο είναι η “αξία” η οποία κυμαίνεται από 0 ως 100% αλλά αυτό

το συστατικό αντιπροσωπεύει την φωτεινότητα του χρώματος. Το  $L^*a^*b$  παρουσιάζεται σε σχήμα κύβου. Το  $L$  κυμαίνεται από 0 ως 100, όπου η μέγιστη τιμή 100 αντιπροσωπεύει το άσπρο και το 0 το μαύρο. Μια θετική τιμή για το  $a$  αντιπροσωπεύει το κόκκινο ενώ μία αρνητική το πράσινο. Όσον αφορά το  $b$ , μια θετική τιμή αντιπροσωπεύει το κίτρινο και μία αρνητική το μπλε. Τα  $a$  και  $b$  δεν έχουν ειδικές μέγιστες τιμές.



Εικόνα 50 : HSV μοντελο χρώματος που παρουσιάζεται σαν κώνος



Εικόνα 51 : Αναπαράσταση του χώρου χρωμάτων  $L^*a^*b$

### Background Subtraction

Για να βελτιωθεί η ανίχνευση χεριού είναι δυνατόν να συνδυαστεί η κατάτμηση χεριού με την αφαίρεση “φόντου” (background subtraction). Σαν αφαίρεση φόντου, ορίζεται η κατάτμηση των περιοχών ενδιαφέροντος, οι οποίες είναι οι μετακινούμενες περιοχές σε μια σκηνή, υποθέτοντας πως η πλειοψηφία των μετακινούμενων περιοχών είναι τα χέρια. Η ρύθμιση της κάμερας είναι στατική, οπότε το φόντο δεν αλλάζει σημαντικά με την πάροδο του χρόνου. Σκοπός επομένως είναι να αφαιρεθούν τα περισσότερα από τα pixel του background, ώστε να βελτιωθεί η κατάτμηση δέρματος ως προς την ακρίβεια, αλλά και ως προς τον υπολογιστικό χρόνο. Μόνο τα πιθανά εκτός background pixel πρέπει να ταξινομηθούν από τον ανιχνευτή χεριού. Έχουν μελετηθεί αρκετές τεχνικές αφαίρεσης background. Οι περισσότερες χρησιμοποιούν μια τιμή για ένα pixel. Στην

πραγματικότητα, ένα pixel αποτελείται από περισσότερα συστατικά, συνήθως ένα κόκκινο, πράσινο και μπλε συστατικό. Φυσικά είναι δυνατόν να χρησιμοποιηθούν και τα τρία συστατικά χρωμάτων αλλά για εξοικονόμηση χώρου στη μνήμη, κάθε καρτέ μετατρέπεται σε αποχρώσεις του γκρι πριν αποθηκευτεί σαν καρτέ background. Για μετατροπή από RGB σε διαβαθμίσεις του γκρι χρησιμοποιείται η ακόλουθη εξίσωση:

$$\text{Grayscale} = 0.3 \_ \text{Red} + 0.59 \_ \text{Green} + 0.11 \_ \text{Blue},$$

όπου Grayscale είναι η προκύπτουσα τιμή σε κλίμακα του γκρι.

### *Connected component labeling*

Η διαδικασία αυτή στοχεύει στην σύνδεση περιοχών από μια δυαδική εικόνα. Ένα συνδεδεμένο στοιχείο ορίζεται ως υπογράφημα σε ένα μη κατευθυνόμενο γράφο. Τα στοιχεία συνδέονται αν δύο ή περισσότερες κορυφές συνδέονται μέσω μονοπατιών.

Στόχος αυτού του βήματος είναι η εξαγωγή των διαστάσεων όλων των περιοχών της εικόνας, τις δυαδικές, δηλαδή, περιοχές που είναι σημειωμένες στο στρώμα άλφα. Οι περιοχές κόβονται σε σχήμα ορθογωνίου και αποθηκεύονται οι συντεταγμένες του πάνω αριστερά pixel, το πλάτος και ύψος του ορθογωνίου. Η είσοδος σε αυτό το βήμα είναι μια δυαδική εικόνα όπου όλα τα pixels δέρματος είναι σημειωμένα ως 1, ενώ τα background pixels σαν 0. Μέσω μεθόδου γνωστή ως connected component labelling, υπολογίζονται τα περιγράμματα. Τα περιγράμματα χρησιμοποιούνται στο επόμενο στάδιο εκτίμησης της θέσης του χεριού. Αυτή είναι μια ευρέως χρησιμοποιούμενη τεχνική για ανάλυση δυαδικών εικόνων. [58]

## Κεφάλαιο 5

# Το ελεύθερο λογισμικό Sweet Home 3D

### 5.1 Εισαγωγή

Το Sweet Home 3D είναι ένα πρόγραμμα σχεδιασμού, αναπτυγμένο σε γλώσσα Java, που μπορεί κανείς να κατεβάσει δωρεάν και να πειραματιστεί εύκολα στη διαρρύθμιση ενός εσωτερικού χώρου. Είναι ένα αρκετά ενδιαφέρον ερασιτεχνικό εργαλείο για την διακόσμηση τόσο εσωτερικών όσο και εξωτερικών χώρων. Περιλαμβάνει τη μετατροπή μιας κάτοψης δύο διαστάσεων σε τρισδιάστατο όγκο, την τοποθέτηση των κατάλληλων επίπλων γραφείου, δωματίου, κουζίνας ή οποιουδήποτε άλλου χώρου μελέτης, καθώς επίσης και την εισαγωγή νέων επίπλων.



Εικόνα 52: το εσωτερικό ενός σπιτιού διαμορφωμένο στο Sweet Home 3D

Στην ιστοσελίδα του λογισμικού υπάρχει μεγάλη συλλογή από έπιπλα, καθώς οι χρήστες κατασκευάζουν συνεχώς καινούρια, και τα προσθέτουν στον κατάλογο. Οι περισσότερες λειτουργίες εκτελούνται με πολύ απλό και γρήγορο τρόπο. Έτσι η προσθήκη και η μετακίνηση των αντικειμένων γίνεται αρκετά εύκολα.

## 5.2 Προετοιμασία συστήματος

### 5.2.1 Ελάχιστες Προδιαγραφές

Το λογισμικό τρέχει στις πλατφόρμες Windows 98 μέχρι Windows 7, Mac OS X 10.4 (Tiger) έως 10.6 (Snow Leopard), σε διάφορες διανομές Linux αλλά και στο Solaris. Για να λειτουργήσει σωστά το πρόγραμμα πρέπει ο υπολογιστής να διαθέτει επεξεργαστή με ταχύτητα τουλάχιστον 400MHZ, 256MB RAM και με τους οδηγούς της κάρτας γραφικών ενημερωμένους στην τελευταία τους έκδοση. Εάν κάποιος διαθέτει λειτουργικό περιβάλλον προγενέστερο των Windows 98 (SP1 ή SP2) πρέπει να έχει εγκατεστημένη την Java, προκειμένου να εγκαταστήσει και να δουλέψει με το Sweet Home 3D.

### 5.2.2 Εγκατάσταση προγράμματος

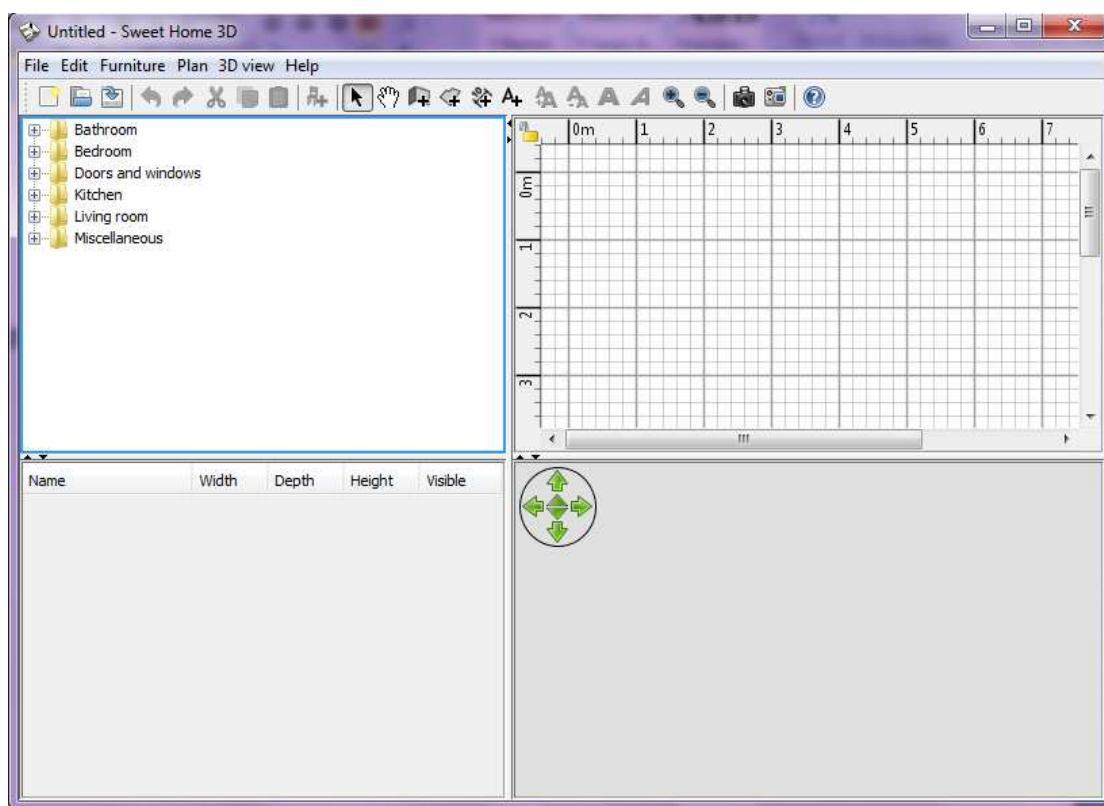
Ανάλογα με την πλατφόρμα που διαθέτουμε κατεβάζουμε την αντίστοιχη έκδοση αρχείου και διαβάζουμε τις οδηγίες. Υπάρχει, επίσης, η δυνατότητα να δουλέψουμε με το πρόγραμμα, χωρίς να το εγκαταστήσουμε στον υπολογιστή μας, απευθείας μέσω διαδικτύου. Προσφέρονται και εδώ οι ίδιες δυνατότητες απλά τα αρχεία που κατασκευάζονται αποθηκεύονται στον server της εφαρμογής, αφού εγγραφεί ο χρήστης, και όχι στον υπολογιστή.

## 5.3 Περιγραφή προγράμματος

Κάθε καινούριο παράθυρο του Sweet Home 3D αποτελείται από μια γραμμή εργαλείων που βοηθούν στην επεξεργασία των χώρων και χωρίζεται σε τέσσερα τμήματα, το μέγεθος των οποίων μπορεί να αλλάξει ανάλογα με τις προτιμήσεις μας. Όταν, λοιπόν, ανοίγουμε το πρόγραμμα βλέπουμε : ένα πλέγμα όπου εισάγονται και μετακινούνται οι τοίχοι και τα έπιπλα (το σχέδιο του χώρου), ένα άλλο που εμφανίζει αυτά σε τρεις διαστάσεις (η τρισδιάστατη απεικόνιση), μία λίστα με όλα τα μοντέλα των επίπλων που μπορούν να εισαχθούν στην τρέχουσα σκηνή (ο κατάλογος επίπλων) και ένας πίνακας με τα ήδη υπάρχοντα στην σκηνή (η λίστα επίπλων που χρησιμοποιούνται στον χώρο).

Ο σχεδιαστής ξεκινά πάρα πολύ εύκολα, με το να εισάγει τους τοίχους στην κάτοψη, και στην συνέχεια εισάγει τα “έπιπλα”. Οι πόρτες και τα παράθυρα ανήκουν στην κατηγορία “έπιπλα”. Όταν εισαχθεί ένα έπιπλο, μπορεί ο χρήστης με δεξί κλικ να επέμβει στις ιδιότητές του.

Κάθε τμήμα μπορεί να έχει τον έλεγχο (δηλαδή να λαμβάνει είσοδο από το πληκτρολόγιο) και μερικές λειτουργίες εξαρτώνται από το εστιασμένο κομμάτι το οποίο αναγνωρίζεται κάθε φορά από το χρωματιστό ορθογώνιο που το περιβάλλει. Για να μεταφερθεί ο έλεγχος σε άλλο τμήμα πατάμε το πλήκτρο Tab και στην συνέχεια τα πλήκτρα Shift και Tab ταυτόχρονα ή πατάμε με το ποντίκι το τμήμα που θέλουμε να εστιαστεί.



Εικόνα 53 : το αρχικό παράθυρο του Sweet Home 3D

### Ο κατάλογος επίπλων

Αυτός ο κατάλογος είναι οργανωμένος σε κατηγορίες και περιέχει όλα τα έπιπλα και τα αντικείμενα που είναι δυνατόν να τοποθετηθούν σε έναν εξωτερικό ή εσωτερικό χώρο. Τα έπιπλα μιας κατηγορίας μπορούν να εμφανιστούν πατώντας με το ποντίκι στο σύμβολο + που υπάρχει δίπλα στο όνομα.

### Η λίστα επίπλων που χρησιμοποιούνται στον χώρο

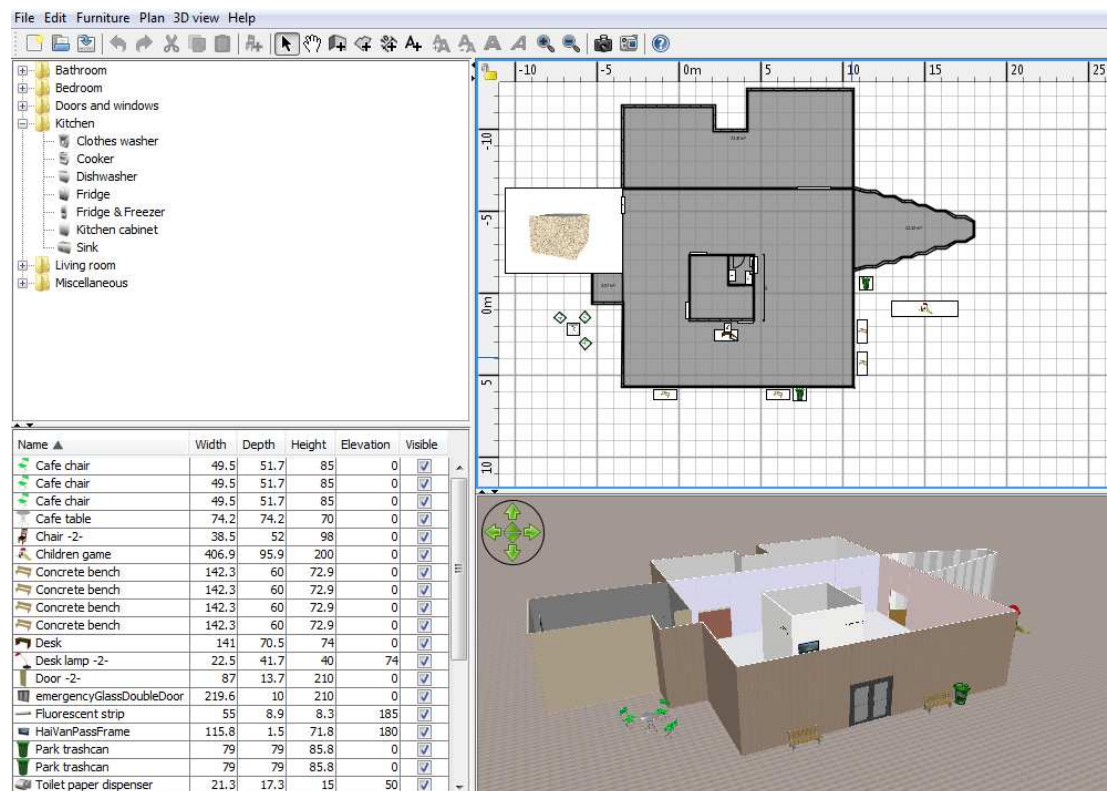
Αυτή η λίστα περιέχει τα έπιπλα που έχουν τοποθετηθεί στο σχέδιο δίνοντας πληροφορίες για το όνομα τους, το μέγεθος τους και άλλα χαρακτηριστικά. Μπορούμε να δούμε τις πληροφορίες αυτές πατώντας με το ποντίκι στο αντικείμενο που θέλουμε.

### Το σχέδιο του χώρου

Το τμήμα αυτό παρουσιάζει σε πλέγμα τον χώρο που έχει σχεδιαστεί και περιβάλλεται από χάρακες. Σε αυτό σχεδιάζονται οι τοίχοι του σπιτιού με την βοήθεια του ποντικιού και τοποθετούνται τα έπιπλα.

### Η τρισδιάστατη απεικόνιση του σπιτιού

Αυτό το τμήμα παρουσιάζει το σπίτι στις 3 διαστάσεις. Το σπίτι εμφανίζεται είτε από πάνω είτε από την οπτική γωνία ενός εικονικού επισκέπτη.

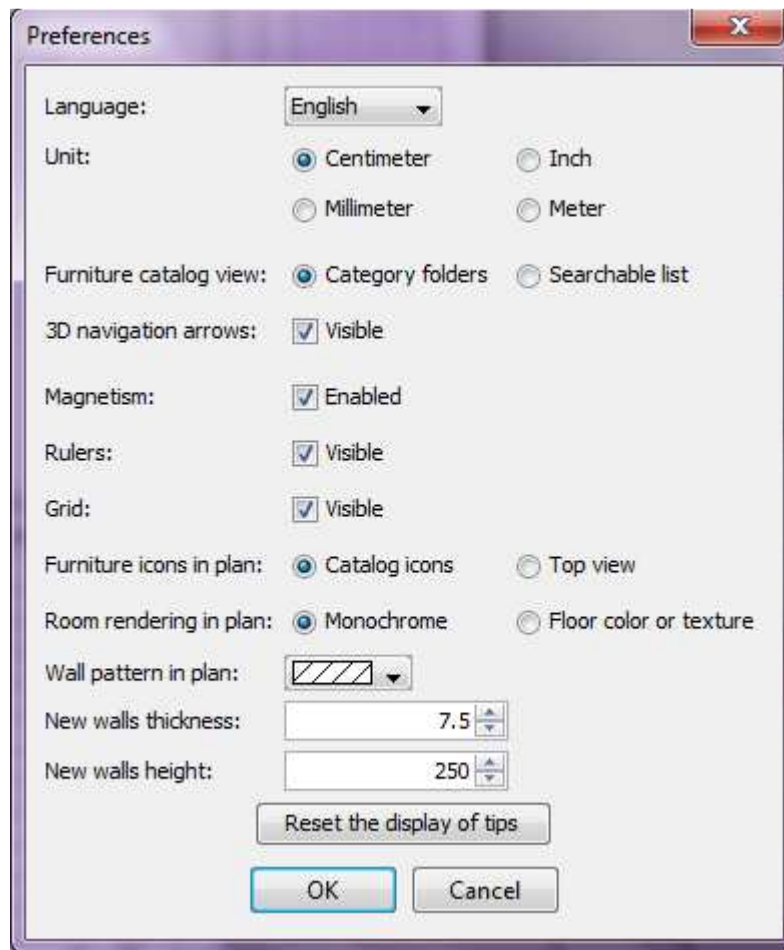


Εικόνα 54 : το παράθυρο του προγράμματος με τα 4 τμήματα. Ο έλεγχος βρίσκεται στο σχέδιο του χώρου, η τρισδιάστατη απεικόνιση δείχνει το σπίτι από πάνω και στον κατάλογο επίπλων φαίνονται τα έπιπλα της κατηγορίας kitchen.



## 5.4 Σχεδίαση ενός νέου χώρου

Προτού ξεκινήσουμε την δημιουργία ενός νέου χώρου πρέπει να επιλέξουμε τις ρυθμίσεις που επιθυμούμε. Να καθορίσουμε, δηλαδή, τις μονάδες μέτρησης, το ύψος και το πάχος των τοίχων και άλλες τέτοιες προτιμήσεις. Η δυνατότητα αυτή δίνεται από την επιλογή του μενού File->Preferences.



Εικόνα 55 : Το κουτί διαλόγου Preferences

Παράθυρο του Sweet Home 3D για την δημιουργία νέου χώρου ανοίγει είτε πατώντας το εργαλείο New home από την γραμμή εργαλείων (το πρώτο εικονίδιο που συναντάμε σε αυτή) είτε από την επιλογή του μενού File->New.

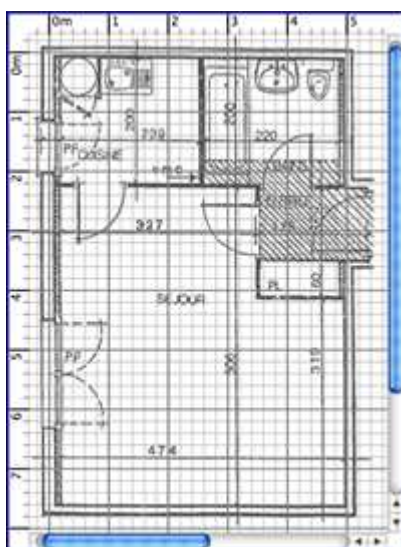


Εικόνα 56 : Το εργαλείο New home

Αρχικά σχεδιάζουμε την κάτοψη του χώρου που θέλουμε να δημιουργήσουμε στο τμήμα του σχεδίου. Εάν έχουμε ήδη σε χαρτί το προσχέδιο του χώρου μπορούμε να το εισαγάγουμε στο πρόγραμμα, αφού το σκανάρουμε, και να το χρησιμοποιήσουμε ως

εικόνα υποβάθρου. Με τον τρόπο αυτό επιταχύνεται ο σχεδιασμός των τοίχων και του χώρου γενικότερα. Εάν δεν υπάρχει το σχέδιο το δημιουργούμε μόνοι μας κατασκευάζοντας τοίχους και δωμάτια. Η φόρτωση ήδη έτοιμου σχεδίου γίνεται από την επιλογή του μενού Choose Plan > Import background image και ένας οδηγός ανοίγει για να διευκολύνει την ολοκλήρωση της διαδικασίας. Το Sweet Home 3D μπορεί να φορτώσει αρχεία τύπου BMP, JPEG, GIF και PNG.

Μόλις κλείσει ο οδηγός η εικόνα θα εμφανιστεί στο πλέγμα του σχεδίου με την κλίμακα και το μέγεθος που ορίστηκαν κατά την διαδικασία φόρτωσης. Εάν δεν είμαστε ευχαριστημένοι με το αποτέλεσμα μπορούμε να επεξεργαστούμε ξανά την εικόνα επιλέγοντας από το μενού Plan > Modify background image.



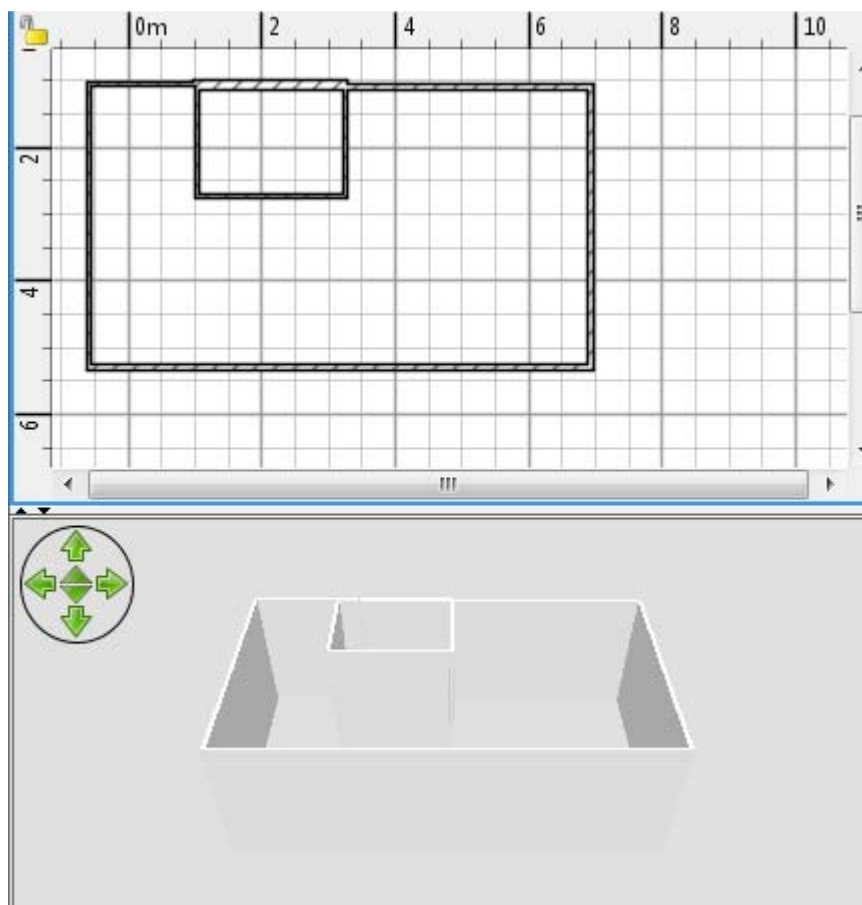
Εικόνα 57 : Η κάτοψη ενός σπιτιού ως εικόνα υποβάθρου στο χώρο του σχεδίου

Για τον σχεδιασμό τοίχων πατάμε το εργαλείο Create walls από την γραμμή εργαλείων είτε επιλέγουμε από το μενού Plan > Create Walls, αφού έχουμε αποφασίσει για την θέση στην οποία θα σχεδιαστεί. Στο σχέδιο κλικάρουμε με το ποντίκι στο σημείο από το οποίο θα ξεκινά ο νέος τοίχος και στην συνέχεια κλικάρουμε στο σχέδιο στο σημείο στο οποίο θέλουμε να τελειώνει. Από την στιγμή που κλικάρουμε το αρχικό σημείο και δεν πατήσουμε το πλήκτρο Escape κάθε νέο κλικ υποδεικνύει το τελικό σημείο του τρέχοντος τοίχου και το αρχικό του επόμενου.

Για να τερματιστεί η διαδικασία σχεδίασης τοίχων πατάμε το εργαλείο Select από την γραμμή εργαλείων.



Εικόνα 58 : Το εργαλείο Create Walls και δίπλα το εργαλείο Select



Εικόνα 59 : Μερικοί τοίχοι όπως αυτοί φαίνονται στο σχέδιο και στην τρισδιάστατη απεικόνιση

Αφού γίνει η σχεδίαση των τοίχων μπορούμε να τους επεξεργαστούμε ακολουθώντας τον οδηγό που εμφανίζεται επιλέγοντας από το μενού Plan > Modify walls ή πατώντας διπλό κλικ στον τοίχο που θέλουμε. Κατά την διάρκεια της διαδικασίας αυτής μπορούμε να αλλάξουμε το χρώμα, ή να βάλουμε textures της επιλογής μας, τόσο στην αριστερή πλευρά όσο και στην δεξιά, και να ρυθμίσουμε το πάχος και το ύψος των τοίχων.

Για να μπορέσουμε να επεξεργαστούμε αντικείμενα ή τοίχους αρκεί να πατήσουμε στο σχέδιο πάνω σε αυτό που θέλουμε έχοντας ήδη επιλέξει το εργαλείο Select. Μπορούμε επίσης να επιλέξουμε ομάδα αντικειμένων σχεδιάζοντας ένα ορθογώνιο επιλογής γύρω από αυτά ή πατώντας σε καθένα από αυτά ενώ έχουμε πατημένο το πλήκτρο Shift. Τα επιλεγμένα αντικείμενα μπορούμε να τα αλλάξουμε θέση αν τα σύρουμε και τα αφήσουμε στο σημείο που θέλουμε να πάνε (απλό drag and drop).

Αφού σχεδιάστηκαν οι τοίχοι και έχουν τοποθετηθεί στο κατάλληλο σημείο μπορούμε να προσθέσουμε πόρτες, παράθυρα και έπιπλα. Για την εισαγωγή τους επιλέγουμε αυτό που θέλουμε από τον κατάλογο επίπλων και είτε το σέρνουμε και το αφήνουμε στο σχέδιο ή στην λίστα των επίπλων που χρησιμοποιούνται στον χώρο, είτε επιλέγουμε το εργαλείο Add furniture από την γραμμή εργαλείων. Το αντικείμενο που επιλέξαμε

εμφανίζεται στο σχέδιο, στην λίστα επίπλων που χρησιμοποιούνται στο χώρο αλλά και στην τρισδιάστατη απεικόνιση. Στην συνέχεια προσαρμόζουμε το μέγεθος, τη θέση και τη γωνία του.



Εικόνα 60 : Το εργαλείο Add furniture

Εάν θέλουμε να επεξεργαστούμε κάποιο άλλο αντικείμενο το επιλέγουμε στο σχέδιο και εμφανίζονται πάνω του, σε κάθε γωνία, 4 δείκτες με βάση τους οποίους ρυθμίζουμε τα χαρακτηριστικά του. Ο δείκτης περιστροφής υποδεικνύει την γωνία που μπορούμε να σύρουμε προκειμένου να περιστραφεί το αντικείμενο, ο δείκτης ανύψωσης την γωνία για να αλλάξουμε την απόσταση του από το έδαφος, ο δείκτης ύψους την γωνία για να μεταβάλλουμε το ύψος του και τέλος ο δείκτης μεγέθους υποδεικνύει την γωνία την οποία μπορούμε να σύρουμε προκειμένου να αλλάξουμε το βάθος και το πλάτος του επιλεγμένου αντικειμένου.



Εικόνα 61 : Ένα επιλεγμένο αντικείμενο με τους 4 δείκτες σε κάθε γωνία

Η επεξεργασία ενός αντικειμένου μπορεί, επίσης, να γίνει από την επιλογή του μενού Furniture > Modify. Εμφανίζεται ένα κουτί διαλόγου που βοηθά στην επεξεργασία του ονόματος του επιλεγμένου επίπλου, της γωνίας περιστροφής του, της θέσης του, της απόστασης του από το έδαφος, του μεγέθους του, του χρώματος του και της ορατότητας του. Τα έπιπλα με επιλογή μηδενικής ορατότητας δεν εμφανίζονται στο σχέδιο ούτε και στο τμήμα τρισδιάστατης απεικόνισης αλλά εξακολουθούν να εμφανίζονται στην λίστα επίπλων ώστε να ξαναγίνουν ορατά όποτε επιθυμούμε.

Το Sweet Home 3D μας δίνει την δυνατότητα να εισάγουμε και δικά μας αντικείμενα, εάν κάποιο από αυτά που ήδη διαθέτει δεν μας αρέσει. Δέχεται αρχεία τύπου OBJ, DAE, 3DS, LWS ή αρχείο ZIP που περιέχει αρχεία τέτοιων τύπων.

Για την εισαγωγή, λοιπόν, δικών μας τρισδιάστατων μοντέλων αρκεί να πατήσουμε την επιλογή του μενού Furniture > Import Furniture και θα ανοίξει ο οδηγός που θα μας βοηθήσει. Μόλις κλείσει ο οδηγός το εισαγόμενο μοντέλο θα εμφανιστεί στον κατάλογο επίπλων και/ή στο σχέδιο και στην λίστα επίπλων ανάλογα με το τι έχουμε επιλέξει. Στην συνέχεια το χρησιμοποιούμε όπως οποιοδήποτε άλλο αντικείμενο του καταλόγου.

Στην συνέχεια μπορούμε να δημιουργήσουμε και δωμάτια στον χώρο. Αυτό γίνεται απλά πατώντας την επιλογή του μενού Plan > Create rooms ή επιλέγοντας το εργαλείο Create rooms από την γραμμή εργαλείων. Για την δημιουργία τους εστιάζουμε στο χώρο του σχεδίου και είτε πατάμε διπλό κλικ οπουδήποτε μέσα σε μια κλειστή επιφάνεια (δηλαδή επιφάνεια που περικλείεται από τοίχους) είτε κλικάρουμε σε κάθε γωνία του δωματίου που θέλουμε να δημιουργηθεί και διπλό κλικ στο τελευταίο σημείο ή πατάμε το πλήκτρο Escape.



Εικόνα 62 : Το εργαλείο Create rooms

Μόλις σχεδιαστεί το δωμάτιο μπορούμε να τροποποιήσουμε το όνομα του το χρώμα και το texture του πατώματος και του ταβανιού επιλέγοντας από το μενού Plan > Modify rooms. Όταν ένα δωμάτιο είναι επιλεγμένο μπορούμε να μετακινήσουμε κάθε σημείο του με το ποντίκι απλά σέρνοντας το και αφήνοντάς το στην καινούρια θέση.

Εάν το επιθυμούμε μπορούμε να εμφανίσουμε στο σχέδιο και διαστάσεις, αποστάσεις, δηλαδή, από αντικείμενα. Γίνεται πολύ απλά επιλέγοντας το εργαλείο Create dimensions ή την επιλογή του μενού Plan > Create dimensions. Έπειτα κλικάρουμε στο αρχικό σημείο που θέλουμε να μετρήσουμε και διπλό κλικ στο τελικό σημείο.

Μπορούμε να εμφανίσουμε και κείμενο στο σχέδιο σε όποιο σημείο θέλουμε εμείς (εάν για παράδειγμα θέλουμε να δώσουμε όνομα σε ένα αντικείμενο). Γίνεται και αυτό πολύ απλά επιλέγοντας το εργαλείο Add texts από την γραμμή εργαλείων ή την επιλογή του μενού Plan > Add texts. Φυσικά η γραμματοσειρά και η μορφή των γραμμάτων είναι της επιλογής μας χρησιμοποιώντας τα εργαλεία γραμμάτων από την γραμμή εργαλείων.

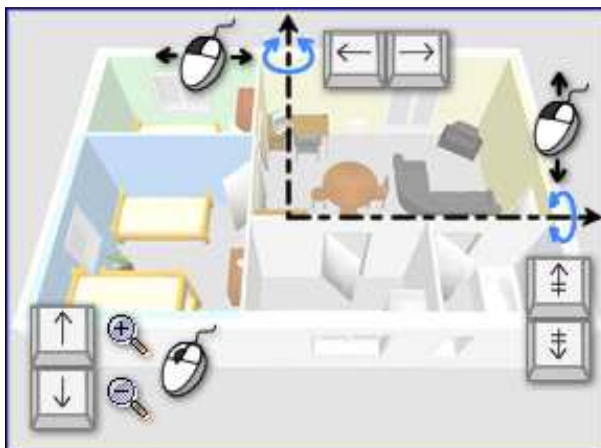


Εικόνα 63 : Το εργαλείο Add texts και τα εργαλεία γραμμάτων

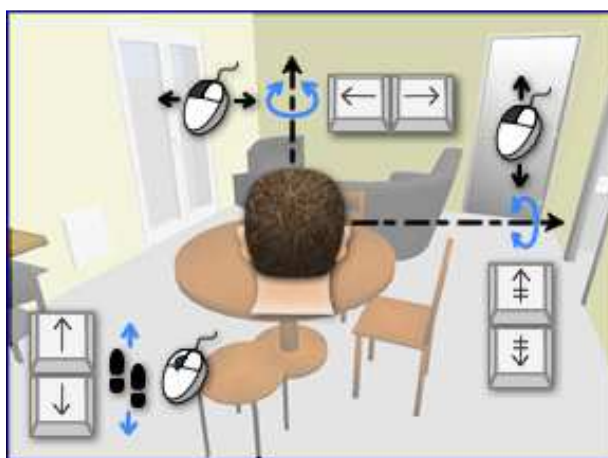


Εικόνα 64 : Το εργαλείο Create dimensions

Καθ' όλη την διάρκεια σχεδίασης του νέου χώρου μπορούμε να βλέπουμε το αποτέλεσμα των πράξεων μας στην τρισδιάστατη απεικόνιση. Υπάρχουν δύο διαφορετικοί τρόποι για να δούμε το σπίτι : με εικονική επίσκεψη (virtual visit) και από ψηλά (aerial view). Ο προεπιλεγμένος τρόπος είναι από ψηλά, αλλά μπορούμε να επιλέξουμε αυτό που θέλουμε εμείς από την επιλογή του μενού 3D view > Aerial view ή 3D view > Virtual visit. Και με τους δύο τρόπους μπορούμε να χρησιμοποιήσουμε τα πλήκτρα βελάκια του πληκτρολογίου ή το ποντίκι για να αλλάξουμε την οπτική γωνία.



Εικόνα 65: Πως αντιδρά το ποντίκι και το πληκτρολόγιο όταν είμαστε στη Aerial View



Εικόνα 66 : Πως αντιδρά το ποντίκι και το πληκτρολόγιο όταν είμαστε στη Virtual View

Όταν είναι ενεργοποιημένη η επιλογή Virtual visit ένας εικονικός επισκέπτης κάνει την εμφάνιση του στο σπίτι. Η θέση του και η γωνία του ενημερώνεται συνεχώς με κάθε βήμα του, τόσο στο σχέδιο όσο και στην τρισδιάστατη απεικόνιση. Αυτός ο επισκέπτης περιβάλλεται από τέσσερις δείκτες για την προσαρμογή της οπτικής του γωνίας. Ο δείκτης “γωνία κεφαλιού” υποδεικνύει την γωνία που μπορούμε να αλλάξουμε προκειμένου να κινηθεί το κεφάλι του επισκέπτη πάνω ή κάτω, ο δείκτης “οπτικό πεδίο” την γωνία με την οποία βλέπουμε στην τρισδιάστατη απεικόνιση, ο δείκτης “γωνία σώματος” την γωνία για να κινηθεί το σώμα του αριστερά ή δεξιά και ο δείκτης “ύψωση ματιών” υποδεικνύει το σημείο το οποίο μπορούμε να σύρουμε για να αλλάξει η οπτική γωνία του επισκέπτη.

Το οπτικό πεδίο του επισκέπτη αλλάζει και από την επιλογή του μενού 3D view > Modify. Από εδώ μπορούμε να αλλάξουμε επίσης το ύψος το χρώμα ή το texture του εδάφους και του ουρανού την φωτεινότητα και την διαφάνεια του τοίχου.



Εικόνα 67 : Ο εικονικός επισκέπτης και οι 4 δείκτες του

## 5.5 Άλλα χαρακτηριστικά του προγράμματος

### *Εκτύπωση*

Αφού τελειώσουμε με το σχεδιασμό του σπιτιού μπορούμε να το εκτυπώσουμε εάν θέλουμε με την επιλογή του μενού File > Print ή File > Print to PDF. Μπορούμε επίσης να δούμε μια προεπισκόπηση του τελικού σχεδίου με την επιλογή του μενού File > Print preview. Το Sweet Home 3D από προεπιλογή εκτυπώνει την λίστα επίπλων, το σχέδιο και την τρέχουσα τρισδιάστατη απεικόνιση του σπιτιού χρησιμοποιώντας προεπιλεγμένο μέγεθος χαρτιού και περιθώρια. Επιλέγοντας από το μενού File > Page setup μπορούμε να τροποποιήσουμε τις επιλογές αυτές.

### *Δημιουργία εικόνων από την τρισδιάστατη απεικόνιση*

Για να δημιουργήσουμε μια εικόνα της τρισδιάστατης απεικόνισης του χώρου που σχεδιάσαμε και να την αποθηκεύσουμε στον υπολογιστή μας επιλέγουμε το εργαλείο Create photo ή πατάμε την επιλογή του μενού 3D view > Create photo. Το αρχείο που σώζεται είναι της μορφής PNG.

### *Δημιουργία τρισδιάστατων βίντεο*

Για να δημιουργήσουμε ένα βίντεο του χώρου που σχεδιάσαμε επιλέγουμε το εργαλείο Create video ή πατάμε την επιλογή του μενού 3D view > Create video. Το αρχείο που σώζεται είναι τύπου Quicktime. Επιλέγουμε στην τρισδιάστατη απεικόνιση την αρχική εικονική θέση της βιντεοκάμερας και πατάμε το κόκκινο κουμπί στον οδηγό που εμφανίζεται. Έπειτα, μετακινούμαστε στην επόμενο σημείο που θέλουμε να έρθει η εικονική βιντεοκάμερα και πατάμε ξανά το κόκκινο κουμπί. Τα βήματα αυτά επαναλαμβάνονται μέχρι να τοποθετήσουμε την κάμερα σε όλα τα σημεία από τα οποία

θέλουμε να περάσει. Αφού τελειώσουμε με τα σημεία αυτά πατάμε το κουμπί Create του οδηγού και τα καρτέ του βίντεο ετοιμάζονται.

#### *Εξαγωγή του χώρου που δημιουργήσαμε σε αρχείο .obj*

Προκειμένου να ανοίξουμε τον χώρο που δημιουργήσαμε και σε άλλα προγράμματα τρισδιάστατης αναπαράστασης, όπως το Blender (για να βελτιώσουμε την περιήγηση μας σε αυτόν για παράδειγμα), το Sweet Home 3D προσφέρει την δυνατότητα δημιουργίας αρχείου .obj με τον χώρο μας. Αυτό γίνεται με την επιλογή του μενού 3D view > Export to OBJ format. Η επιλογή αυτή θα αποθηκεύσει σε αρχείο .obj τις περιγραφές όλων των αντικειμένων που εμφανίζονται στην τρισδιάστατη απεικόνιση, θα δημιουργήσει ένα αρχείο mtl όπου θα περιγράφονται τα χρώματα των αντικειμένων και θα σώσει τις εικόνες των textures που έχουν χρησιμοποιηθεί για τον χώρο αυτό. [23]

## **5.6 Αξιολόγηση του προγράμματος**

Το λογισμικό Sweet Home 3D είναι μια πολύ ενδιαφέρουσα και πολύ εύκολη εφαρμογή που βοηθά στην βελτίωση των σχεδιαστικών ικανοτήτων. Είναι τόσο απλή στην χρήση της που ακόμη και ένας αρχάριος σε ηλεκτρονικούς υπολογιστές αλλά και στην διακόσμηση χώρων μπορεί να την χρησιμοποιήσει χωρίς να ζητήσει βοήθεια από κάποιον. Υπάρχει μεγάλη πληθώρα τρισδιάστατων μοντέλων που μπορούν να εισαχθούν στο πρόγραμμα και να κάνουν το χώρο να φαίνεται ακόμη πιο ρεαλιστικός. Το πρόγραμμα διατίθεται δωρεάν και συνεχώς εξελίσσεται ενώ υποστηρίζει πέντε γλώσσες : αγγλικά, γαλλικά, γερμανικά, ιταλικά και πορτογαλικά.

Παρόλο που είναι ιδανικό για τους αρχάριους και τους ερασιτέχνες δεν είναι το κατάλληλο για τους επαγγελματίες, εξαιτίας της απλότητας του και της έλλειψης αναπτυγμένων δυνατοτήτων. Αν και η ποικιλία των τρισδιάστατων μοντέλων είναι μεγάλη δεν είναι ικανή να καλύψει τις ανάγκες ενός αρχιτέκτονα ή ενός διακοσμητή. Το πρόγραμμα δυστυχώς, τουλάχιστον ακόμη, δεν προσφέρει την δυνατότητα κατασκευής κτιρίων με πολλούς ορόφους παρά μόνο έναν.



## Κεφάλαιο 6

### Το ελεύθερο λογισμικό EaZD

#### 6.1 Εισαγωγή

Το EaZD αναπτύχθηκε αρχικά ως μια πειραματική πλατφόρμα ανάπτυξης διαδικασιών deferred rendering για την εξέταση διαφόρων αλγορίθμων σκίασης και επίδρασης φωτισμού (shading and illumination effects algorithms) χρησιμοποιώντας την OpenGL σε πραγματικό χρόνο. Στην πορεία εξελίχθηκε σε ένα καλό ανεξάρτητο λογισμικό και αποφασίστηκε από τους δημιουργούς του να γίνει διαθέσιμο στο διαδίκτυο για όλους τους ενδιαφερόμενους κατασκευαστές γραφικών υπολογιστών.

Η βιβλιοθήκη είναι χωρισμένη σε δύο τμήματα, την βασική διεπαφή renderer (core renderer API) και το γράφημα σκηνής (scene graph) που εκμεταλλεύεται τις δυνατότητες του deferred renderer και των αποτελεσμάτων του. Ωστόσο, η μέθοδος deferred renderer μπορεί να χρησιμοποιηθεί είτε ξεχωριστά, περνώντας μόνο δύο callbacks κλήσεις σε μια διαδικασία rendering και σε μια συνάρτηση αρχικοποίησης κάμερας αντίστοιχα, είτε σε συνδυασμό με το γράφημα σκηνής (SceneGraph3D). Παρ' όλο που λειτουργεί σωστά δεν έχει ακόμη ολοκληρωθεί και υπάρχουν κάποιοι περιορισμοί, όσον αφορά τις δυνατότητες του και τα αρχεία που υποστηρίζει. Εξελίσσεται όμως διαρκώς και τα χαρακτηριστικά του βελτιώνονται.

Η βιβλιοθήκη συνοδεύεται από ένα σύνολο παραδειγμάτων που υποδεικνύουν την χρήση της διεπαφής και την σωστή σειρά αρχικοποίησης των στοιχείων της.



Εικόνα 68 : Μία σκηνή που δημιουργήθηκε με το EaZD

## 6.2 Προετοιμασία συστήματος

### 6.2.1 Ελάχιστες προδιαγραφές

Το πρόγραμμα EaZD βασίζεται στην OpenGL, έκδοσης 2.0 ή παραπάνω, για την διαδικασία rendering και απαιτεί πλήρη υποστήριξη για την ανάπτυξη προγραμμάτων στην γλώσσα προγραμματισμού GLSL shading language (αυτό σημαίνει ότι σε ορισμένες κάρτες ATI δεν θα λειτουργήσει το πρόγραμμα). Το γράφημα σκηνής χρησιμοποιεί για τις συσκευές εισόδου το Direct Input, επομένως η κατάλληλη έκδοση του DirectX SDK πρέπει να είναι εγκατεστημένη στον υπολογιστή. Τα αρχεία σκηνής είναι αναπτυγμένα σε μορφή XML που υποστηρίζεται στο πρόγραμμα από την βιβλιοθήκη libXML 2.0. Για μεγαλύτερη ευκολία παρέχονται με το πρόγραμμα όλες οι υπόλοιπες βιβλιοθήκες που χρειάζονται.

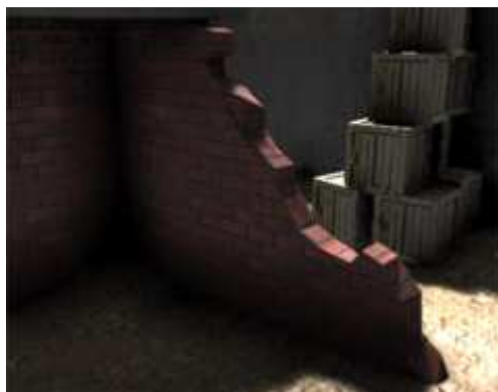
### 6.2.2 Εγκατάσταση προγράμματος

Το πρόγραμμα τρέχει σε πλατφόρμες Windows αλλά και σε διάφορες διανομές Linux με ορισμένους, όμως, περιορισμούς. Στα Linux, για παράδειγμα, δεν λειτουργεί ακόμη το joystick και η μόνη είσοδος που δέχεται είναι από το ποντίκι και το πληκτρολόγιο. Και στα δύο λειτουργικά χρειάζεται το πρόγραμμα CMake, έκδοσης μεγαλύτερης της 2.6, για να μπορέσει να χτιστεί και να δοκιμαστεί το λογισμικό. Επίσης, στα Windows το EaZD δουλεύει μέσω του προγράμματος Microsoft Visual Studio.

## 6.3 Η διεπαφή deferred renderer

Κατά την διαδικασία deferred rendering οι γεωμετρικές πληροφορίες, που χρησιμοποιούνται για την διαδικασία rasterization, αναμειγνύονται με τις πληροφορίες φωτισμού και χρώματος της τελικής σκηνής. Για να γίνει η σωστή ανάμειξη, όλες οι πληροφορίες που είναι απαραίτητες για τους υπολογισμούς αποθηκεύονται σε ενδιάμεσους αποθηκευτικούς χώρους (buffers) χρησιμοποιώντας την τεχνική της διαδικασίας rendering πολλαπλών στόχων (multiple targets). Στην συνέχεια, όλοι οι

υπολογισμοί εκτελούνται στον χώρο της σκηνής απομακρύνοντας την πολυπλοκότητα της σκηνής από τον χρόνο της διαδικασίας παραγωγής της εικόνας.



Εικόνα 69 : Το αποτέλεσμα μιας σκηνής του EaZD

### 6.3.1 Εισαγωγή του deferred renderer σε εφαρμογές

Το πρόγραμμα EaZD συνοδεύεται από ένα αρχείο στατικής βιβλιοθήκης, το `DeferredRenderer.lib`, και δύο αρχεία επικεφαλίδας (header files), το `DeferredRenderer.h` και το `DeferredRendererShaders.h`. Όταν δημιουργεί κάποιος μια εφαρμογή, που βασίζεται στο EaZD, πρέπει να την συνδέσει με την στατική βιβλιοθήκη, να συμπεριλάβει το `DeferredRenderer.h` και να της παρέχει το μονοπάτι για αυτό το header αρχείο. Το δεύτερο header αρχείο δεν είναι απαραίτητο για σύνδεση και για την διαδικασία μεταγλώττισης (compiling) της εφαρμογής. Απαιτείται μόνο σε περίπτωση που η βιβλιοθήκη πρέπει να μεταγλωττιστεί ξανά.

#### *Χρήση deferred renderer ξεχωριστά*

Όπως, ήδη, έχει αναφερθεί αρχικά πρέπει να αποκτήσει το λειτουργικό περιβάλλον την υποστήριξη της OpenGL. Πιο συγκεκριμένα, απαιτείται η χρήση της βιβλιοθήκης OpenGL Extensions Wrangler (GLEW).

Εκτελείται, επίσης, και η κλήση δύο callback συναρτήσεων, της `draw_callback` και της `camera_callback`. Η πρώτη περιέχει τον κώδικα για την εξέταση της γεωμετρίας αναφορικά με το παγκόσμιο σύστημα συντεταγμένων. Η δεύτερη παρέχει όλους τους γεωμετρικούς μετασχηματισμούς και την προβολή που είναι απαραίτητη για την μεταφορά όλης της σκηνής σε συντεταγμένες οθόνης, συμπεριλαμβανομένου και της αντίστοιχης οπτικής γωνίας.

Εάν πρέπει να ληφθούν υπ' όψιν και οι σκιές, που προκαλούνται από μία τουλάχιστον πηγή φωτός, χρειάζεται να κληθεί η μέθοδος `update Shadows()` πριν την κλήση της `draw` προκειμένου να προετοιμαστούν τα στοιχεία για τις σκιές.

#### *Χρήση deferred renderer σε συνδυασμό με το SceneGraph3D*

Ο τρόπος αυτός είναι αρκετά πιο εύκολος. Φυσικά και εδώ απαιτείται η μέριμνα για την βιβλιοθήκη GLEW. Για την χρήση του SceneGraph3D πρέπει να συμπεριληφθεί το αρχείο SceneGraph.h στην εφαρμογή και να γίνει σύνδεση με την στατική βιβλιοθήκη SceneGraph.lib. Σε αυτήν την περίπτωση δεν είναι απαραίτητη η κλήση callback συνάρτησης αφού όλοι οι μετασχηματισμοί και η εξέταση της γεωμετρίας διαχειρίζεται εσωτερικά από την SceneGraph3D. Το μόνο που χρειάζεται είναι η αρχικοποίηση του γραφήματος σκηνής και η επιλογή του αντικειμένου renderer που θα χρησιμοποιηθεί για έξοδο. Αυτό, φυσικά, σημαίνει πως μπορούν να παρέχονται διαφορετικές εκδόσεις ενός renderer, να γίνονται αλλαγές μεταξύ αυτών ή να ενσωματώνεται ο ίδιος σε πολλαπλά γραφήματα σκηνής.

Από τον ορισμό της μεθόδου `setRenderer()` διαπιστώνεται πως παίρνει ως είσοδο ένα αντικείμενο `GenericRenderer`, δηλαδή ένα στιγμιότυπο υπερκλάσης του `DeferredRenderer`. Αυτό γίνεται για την ανάπτυξη του γραφήματος σκηνής ακόμη και σε πλατφόρμες όπου δεν μπορούν να χρησιμοποιηθούν shaders ή δεν ικανοποιούν τις προδιαγραφές του EaZD.

Αφού αρχικοποιηθεί το EaZD προχωρά στην φόρτωση της σκηνής και στην προετοιμασία του γραφήματος για την διαδικασία rendering. Όταν τερματιστεί η εφαρμογή το αντικείμενο του γραφήματος σκηνής πρέπει να απελευθερωθεί.

### **6.3.2 Γεωμετρικοί “κάδοι” (Geometry bins)**

Το πρόγραμμα EaZD χρησιμοποιεί τέσσερις κάδους, στους οποίους αποθηκεύονται οι πληροφορίες γεωμετρίας είτε δυναμικά είτε μόλις δημιουργηθούν. Κατά την διαδικασία rendering διατρέχεται το γράφημα σκηνής σε διαφορετικές χρονικές στιγμές και κάθε στιγμή ενεργοποιείται μία κατάσταση “κάδου” (bin state). Εξαρτάται από την θέληση του προγραμματιστή το αν θα πραγματοποιηθεί έλεγχος συνθηκών για να επιτρέπεται στις γεωμετρικές πληροφορίες να εμφανίζονται σε κάθε στιγμή σύμφωνα με τον “κάδο” στον οποίο ανατέθηκαν. Όλοι οι κόμβοι ακολουθούν αυτήν την αρχή. Για παράδειγμα, οι γεωμετρικοί κόμβοι χωρίζουν τα πρωταρχικά σχέδια με βάση την διαφάνεια του

υλικού και τις πηγές φωτός, που σχεδιάζονται ως τμήματα (particles) που λάμπουν, όταν ένας “κάδος” αυτών των τμημάτων ενεργοποιείται και ούτω καθεξής. Μέχρι στιγμής υπάρχουν τέσσερις κάδοι.

#### *SCENE\_GRAPH\_RENDER\_MODE\_HIDDEN*

Την χρονική στιγμή που εκτελείται “hidden” εμφανίζεται η γεωμετρία που είναι μαρκαρισμένη ως μη ορατή. Αυτή η στιγμή ενεργοποιεί την διαδικασία rendering για σκιές. Δεν χρησιμοποιείται πολύ συχνά.

#### *SCENE\_GRAPH\_RENDER\_MODE\_NORMAL*

Η κλασική (και πρώτη στη σειρά) χρονική στιγμή που εμφανίζει αδιαφανή γεωμετρία. Εδώ λαμβάνονται υπ’ όψιν και τα πολύγωνα με αδιαφανή textures.

#### *SCENE\_GRAPH\_RENDER\_MODE\_TRANSPARENCY*

Αυτός ο “κάδος” δέχεται την διαδικασία rendering μετά τον κανονικό “κάδο” και απενεργοποιεί την ανανέωση του buffer βάθους. Τα πολύγωνα με αδιαφανή υλικό λαμβάνονται υπ’ όψιν.

#### *SCENE\_GRAPH\_RENDER\_MODE\_PARTICLES*

Αντιπροσωπεύει την διαδικασία rendering των τμημάτων και είναι ο τελευταίος στη σειρά που εμφανίζεται. Είναι ιδιαίτερος “κάδος” καθώς οι εντολές της OpenGL χρησιμοποιούνται με πολύ συγκεκριμένο τρόπο για την κατασκευή ενός γεωμετρικού shader που θα παράγει τμήματα.

### **6.3.3 Φώτα και σκιές**

Το κομμάτι του EaZD deferred renderer δημιουργεί τις δικές του πηγές φωτός και τα φώτα της OpenGL απορρίπτονται εντελώς. Η διαχείριση φωτός του renderer (παρούσα και στην κλάση GenericRenderer για να λειτουργήσει σωστά η προσομοίωση της συμπεριφοράς της στο πεδίο της OpenGL) ορίζει μέχρι και 128 εσωτερικά φώτα που μπορούν να δημιουργηθούν, να ρυθμιστούν, να ανάψουν ή να σβήσουν, να μεταβληθούν ή να μετακινηθούν ανά πάσα χρονική στιγμή. Οι πηγές φωτός μπορεί να είναι φώτα σημείων, φώτα περιοχής (δηλαδή φώτα σημείων με μη μηδενική ακτίνα) και φώτα σποτάκια με ή χωρίς απόσβεση ή σκιές. Τα φώτα έχουν και εύρος επίδρασης (και κοντινό εύρος για να ορίζουν την πιο κοντινή στο φως απόσταση για τον χάρτη

σκιάς (shadow map)), όπως επίσης και τιμές χρώματος σε RGB που μπορούν να οριστούν και πιο πάνω από (1,1,1). Εσωτερικά, ο φωτισμός μπορεί να αντέξει σε έκθεση και μέχρι 200% χωρίς να κορεστεί, αν και το τελικό αποτέλεσμα εξαρτάται από τον buffer χρώματος.

Για να προστεθεί κάποιο φως στην εφαρμογή (όχι μέσω του γραφήματος σκηνής, που τα χειρίζεται ως κανονικούς κόμβους γραφήματος και εκτελεί τις υπόλοιπες λειτουργίες εσωτερικά) πρέπει να κληθεί η μέθοδος του renderer `int createLight()`. Αυτή η μέθοδος αρχικοποιεί μία καινούρια πηγή φωτός (αρχικά είναι απενεργοποιημένη) και επιστρέφει ένα αναγνωριστικό σε αυτή. Επιπλέον κλήσεις για επεξεργασία της πηγής φωτός γίνονται μέσω αυτού του αναγνωριστικού. Η μέθοδος επιστρέφει -1 εάν μια πηγή φωτός δεν μπορεί να δημιουργηθεί.

Η θέση του φωτός καθορίζεται αναφορικά με το παγκόσμιο σύστημα συντεταγμένων. Η σειρά των εντολών για τα φώτα δεν είναι περιοριστική, εφόσον το φως έχει δημιουργηθεί αρχικά και οποιαδήποτε κλήση μπορεί να γίνει ανά πάσα στιγμή. Αλλάζοντας, όμως, μερικά χαρακτηριστικά, όπως η μέθοδος δειγματοληψίας για την δημιουργία σκιών, απαιτείται η εξέταση των shaders ξανά με αποτέλεσμα να υπάρχει μια μικρή καθυστέρηση.

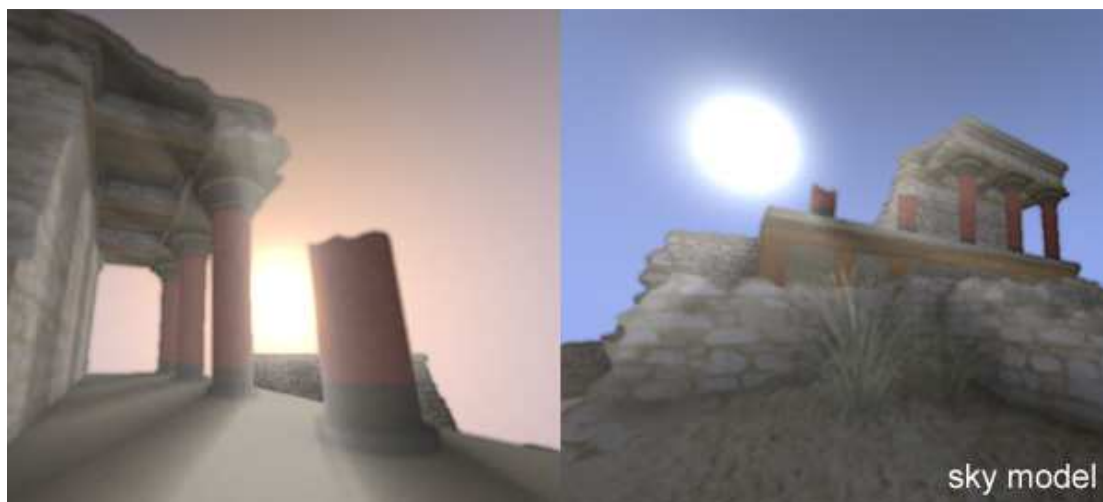


Εικόνα 70 : Ένα κομμάτι σκηνής με ενεργοποίηση των σκιών

### 6.3.4 Το μοντέλο του ουρανού

Το EaZD παρέχει ένα μοντέλο φωτισμού με βάση τον ουρανό. Ο χρήστης καθορίζει ποια πηγή φωτός είναι ο ήλιος και σχεδιάζει τον ουρανό και την υπόλοιπη σκηνή με βάση την θέση του φωτός. Το μοντέλο δεν είναι 100% φυσικό αλλά είναι πολύ γρήγορο στους υπολογισμούς. Για να βελτιώσουμε την ρεαλιστικότητα του αποτελέσματος ο ήλιος μπορεί να αντιμετωπιστεί ως ογκώδες φως (volumetric light) διαχέοντας κατάλληλες ακτίνες φωτός (και καθορίζοντας αντίστοιχα τις σκιές). Ο αριθμός των δειγμάτων σε κάθε ακτίνα είναι προσαρμόσιμος. Για την ενεργοποίηση του μοντέλου

ουρανού χρησιμοποιείται η μέθοδος `void enableSkyLight(bool e)`. Καλώντας την συνάρτηση `SkyModel * getSkyModel()` παρέχεται δείκτης στην δομή του μοντέλου έτσι ώστε να ρυθμιστούν οι παράμετροι του.



Εικόνα 71 : σκηνές με ενεργοποιημένο το μοντέλο του ουρανού

### 6.3.5 Απόδοση

Εάν η απόδοση είναι αργή, εξαιτίας υψηλής ανάλυσης αποτελέσματος ή για κάποιον άλλο λόγο, υπάρχουν ορισμένες ενέργειες που μπορούν να γίνουν για να αυξηθούν οι ρυθμοί πλαισίου (frame rates). Αρχικά, μπορεί να μειωθεί το μέγεθος του εσωτερικού buffer είτε καθορίζοντας μία αναλογία μεταξύ του πραγματικού μεγέθους παραθύρου και του μεγέθους του buffer πλαισίων, κάνοντας χρήση της μεθόδου `setBufferScale(float scale)`, είτε παρέχοντας ένα καθορισμένο μέγεθος για τον εσωτερικό buffer ασχέτως μεγέθους παραθύρου, καλώντας την μέθοδο `setFixedBufferSize(int width, int height)`. Ο αρχικός παράγοντας μεγέθους στην `setBufferScale` ορίζεται ως  $scale = (\text{μέγεθος εσωτερικού buffer}) / (\text{πραγματικό μέγεθος παραθύρου})$ .

Εάν είναι ενεργοποιημένα στην σκηνή πολλά φώτα που δημιουργούν σκιές, ο ρυθμός των υπολογισμών του χάρτη σκιών μειώνεται καλώντας την μέθοδο `void setLightSkipFrames(int light, int sf)` (και φυσικά μειώνοντας το μέγεθος τους).

Τελειώνοντας, το EaZD παρέχει την δυνατότητα δυναμικής προσαρμογής του μεγέθους του buffer απεικόνισης μέσω μιας συνάρτησης του ρυθμού πλαισίων. Χρησιμοποιώντας έναν συνδυασμό των μεθόδων `getFrameRate()` και `setBufferScale()` προσαρμόζονται δυναμικά οι εσωτερικοί buffers και διατηρείται μία σταθερή απόδοση.

## 6.4 Η διεπαφή SceneGraph3D

Η διεπαφή SceneGraph3D λειτουργεί ως ένα εργαλείο διαχείρισης γραφήματος σκηνής για την διεπαφή renderer του EaZD. Αν και μπορεί να δουλέψει και να σταθεί μόνη της σαν βιβλιοθήκη έχει σχεδιαστεί για να λειτουργεί μαζί με το deferred renderer του EaZD. Προσφέρει, προς το παρόν, μόνο τους πρωταρχικούς κόμβους για την σχεδίαση απλών γραφημάτων σκηνής και λόγω του τρόπου δημιουργίας του είναι αρκετά εύκολο να επεκταθεί και να παρέχει πιο εξειδικευμένους κόμβους.

Όπως έχει ήδη αναφερθεί, το γράφημα σκηνής αποτελείται από έναν αριθμό κόμβων, ο καθένας από τους οποίους αντιπροσωπεύει ένα λειτουργικό κομμάτι στον τρισδιάστατο χώρο. Ο κόμβος - ρίζα του γραφήματος ονομάζεται World3D. Κάθε κόμβος στην ιεραρχία παράγεται από την γενική κλάση Node3D, η οποία διαχειρίζεται όλες τις διαδικασίες γεγονότων και μηνυμάτων. Μία μεγάλη κατηγορία υποκλάσεων της κλάσης αυτής είναι οι κόμβοι ομάδας (Group3D). Εκτός από την βασική λειτουργία διαχείρισης μιας ομάδας κόμβων και εκτέλεσης των εντολών του γραφήματος που τους αφορούν, ο κόμβος Group3D παρέχει επίσης και συναρτήσεις κλειδιά, όπως αναζητήσεις με βάση το όνομα ή τον τύπο του κόμβου. Ο κόμβος World3D είναι ένας ειδικός κόμβος Group3D που ασχολείται με την εισαγωγή ολόκληρου του γραφήματος και την σύνδεση του στο rendering σύστημα. Από την δημιουργία τους, όλοι οι κόμβοι διατηρούν αναφορές στο μοναδικό στιγμιότυπο World3D του γραφήματος σκηνής αλλά και στον κόμβο γονέα τους.

Το γράφημα σκηνής μπορεί να δημιουργηθεί είτε χειρονακτικά, προσθέτοντας νέα στιγμιότυπα διαφόρων κόμβων στους ήδη υπάρχοντες κόμβους που παράγονται από τον Group3D, είτε αφού φορτωθεί από ένα αρχείο XML, γνωστό ως αρχείο σκηνής.

Υπάρχουν εννέα τύποι κόμβων σε ένα γράφημα σκηνής.

### *Node3D*

Από αυτόν παράγονται όλοι οι άλλοι κόμβοι. Παρέχει βασικές καταστάσεις, όπως ορατότητα (visible) και ενεργητικότητα (active), τον μηχανισμό ουράς μηνυμάτων και την δυνατότητα κλήσεων όλων των βασικών λειτουργιών στο γράφημα σκηνής.

### *World3D*

Είναι ο κόμβος - ρίζα του γραφήματος σκηνής και φυσικά υποκλάση του Group3D. Πραγματοποιεί την σύνδεση με το renderer και δημιουργεί αναφορές στην ενεργή κάμερα της σκηνής (αφού πολλοί κόμβοι Camera3D μπορούν να αλλάξουν μεταξύ τους κατά την διάρκεια ζωής του γραφήματος σκηνής).



### *User3D*

Αυτός ο τύπος κόμβου δημιουργεί μια αυτόνομη οντότητα που μπορεί να περιηγηθεί στον τρισδιάστατο κόσμο με την βοήθεια ενός ελεγκτή που παρέχει είσοδο και οδηγεί τον κόμβο. Δεν είναι ο μοναδικός κόμβος, πολλά στιγμιότυπα User3D μπορούν να υπάρχουν στον εικονικό κόσμο, ο καθένας από τους οποίους αντιπροσωπεύει και έναν διαφορετικό χρήστη (για παράδειγμα σε παιχνίδια με πολλούς παίκτες) ή χαρακτήρες ελεγχόμενοι από το πληκτρολόγιο, ζώα ή κινούμενα αντικείμενα γενικότερα. Οι κόμβοι αυτοί δεν ορίζουν ένα ορατό σύστημα, αυτό είναι δουλειά του κόμβου Camera3D (οι κάμερες και οι χρήστες είναι εντελώς διαφορετικές οντότητες στο SceneGraph3D). Ο User3D είναι υποκλάση του Group3D (στην πραγματικότητα παράγεται από έναν κόμβο TransformNode3D που με την σειρά του είναι υποκλάση του Group3D) και μπορεί να περιέχει άλλους κόμβους.

### *Camera3D*

Το μάτι στον εικονικό κόσμο. Ορίζεται έτσι ώστε να ακολουθεί έναν συγκεκριμένο χρήστη ή οποιοδήποτε άλλο κόμβο, να στέκεται μόνο του ή να μετατρέπεται σε ένα κανονικό γεωμετρικό κόμβο σε μια αλυσίδα κινήσεων. Πολλαπλές κάμερες μπορεί να υπάρχουν αλλά μόνο μία είναι η κύρια που δίνει το αποτέλεσμα.

### *Input3D*

Είναι η βασική κλάση όλων των τύπων των ελεγκτών κόμβων. Ο βασικός κόμβος εισόδου διαχειρίζεται την είσοδο που δέχεται από το GUI μέσω του ποντικιού ή εναλλακτικά του joystick. Πρόκειται για μια αφηρημένη διεπαφή συσκευής που παρέχει κανονικοποιημένες τιμές αξόνων και καταστάσεις κουμπιών για περιήγηση και ενεργοποίηση γεγονότων. Συνοδεύεται από έναν μεγάλο αριθμό προκαθορισμένων γεγονότων, που συνδέονται κυρίως με τις αλλαγές στην κατάσταση των μεταβλητών. Κάθε υποκλάση πρέπει να εκτελεί αυτό το αφηρημένο μοντέλο προκειμένου ένας κόμβος User3D να δέσει μαζί του και να μπορεί να δεχτεί έλεγχο. Κατά την διάρκεια φόρτωσης σκηνής, ο κόμβος Input3D καθορίζει την συσκευή που ζητείται (εάν είναι διαθέσιμη) και δημιουργεί ένα εσωτερικό αντικείμενο που θα ενεργοποιήσει την φυσική συσκευή να λειτουργήσει μέσω της διεπαφής DirectX SDK.

### *Group3D*

Η βασική συλλογή κόμβων στο γράφημα σκηνής. Κάθε κόμβος που χρειάζεται να διατηρεί επαφές με άλλους κόμβους πρέπει να παράγεται ως υποκλάση αυτού του κόμβου ή κάποιας υποκλάσης του.

### *TrasnformNode3D*

Ενεργεί ως ένας κανονικός γεωμετρικός κόμβος. Υποστηρίζονται όλοι οι μετασχηματισμοί, περιστροφή, αλλαγή κλίμακας και μετατόπιση αλλά και μέθοδοι για να αποκτηθεί και να μεταβληθεί ο πίνακας απευθείας. Η σειρά των μετασχηματισμών είναι καθορισμένη, αλλαγή κλίμακας, έπειτα περιστροφή και στην συνέχεια μετατόπιση.

### *Geometry3D*

Ο απλούστερος κόμβος για την αναπαράσταση γεωμετρικών πληροφοριών. Μέχρι στιγμής, υποστηρίζεται η διαχείριση και εμφάνιση στατικής γεωμετρίας που φορτώνεται από ένα αρχείο.

### *ProximityTrigger3D*

Το γράφημα σκηνής, συνήθως, περιέχει πολλούς τύπους κόμβων που δρουν για την ενεργοποίηση γεγονότων. Αυτός είναι ο απλούστερος κόμβος που ενεργοποιείται και απενεργοποιείται κάθε φορά που ο καθορισμένος κόμβος εισέρχεται ή εξέρχεται της περιοχής επίδρασης του. Η περιοχή επίδρασης ενός κόμβου ProximityTrigger3D μπορεί να είναι σφαιρική, ένα κουτί ή ένα επίπεδο.

## **6.4.1 Εισαγωγή του SceneGraph3D σε εφαρμογές**

Για την εισαγωγή του SceneGraph3D στον κώδικα μιας εφαρμογής πρέπει να ενώσουμε το εκτελέσιμο κομμάτι με την στατική βιβλιοθήκη SceneGraph3D.lib και να συμπεριλάβουμε το αρχείο επικεφαλίδας οπουδήποτε γίνονται κλήσεις σε κάποιον κόμβο. Το αρχείο SceneGraph3D.h περιέχει μία δήλωση συμπερίληψης για κάθε αρχείο επικεφαλίδας που παρέχει τον ορισμό όλων των τύπων κόμβων. Ο φάκελος include που υπάρχει με το πρόγραμμα EaZD περιέχει όλα αυτά τα αρχεία επικεφαλίδας, επομένως το μόνο που χρειάζεται είναι να συμπεριληφθεί το κύριο αρχείο επικεφαλίδας SceneGraph3D.h στην εφαρμογή.

## **6.4.2 Αρχικοποίηση και φόρτωση σκηνής**

Η διαδικασία φόρτωσης σκηνής αποτελείται από την ανάγνωση του XML αρχείου σκηνής. Αφού ολοκληρωθεί η διαδικασία, όλοι οι κόμβοι είναι σε ετοιμότητα αλλά όχι

ακόμη έτοιμοι να τρέξουν. Προκειμένου να προετοιμαστούν οι κόμβοι για την διαδικασία rendering πρέπει να κληθεί η συνάρτηση αρχικοποίησης που εκτελεί τις ενέργειες : μετατρέπει όλες τις αναφορές ονομάτων σε πραγματικές αναφορές αντικειμένων, διαβάζει όλα τα μηνύματα γεγονότων και τα στοιβάζει σε έναν πίνακα για γρηγορότερη ενεργοποίηση, ελέγχει το γράφημα σκηνής για σοβαρά λάθη (εάν ένας κόμβος Camera3D ή ένας κόμβος User3D λείπει, διορθώνει το πρόβλημα που σχετίζεται με την κατάσταση), φορτώνει αρχικές τιμές στις διάφορες καταστάσεις των κόμβων. Αυτές οι αρχικές τιμές μπορούν να επαναφερθούν οποιαδήποτε στιγμή εκτελώντας επαναφορά στον κόμβο world (είτε χειρονακτικά, , World3D::reset(), είτε τοποθετώντας ένα μήνυμα επαναφοράς στον ειδικό κόμβο %world).

### 6.4.3 Ο κύκλος ζωής του SceneGraph

Μετά την αρχικοποίηση, η εφαρμογή εισέρχεται σε έναν άπειρο βρόγχο, μέχρι να ενεργοποιηθεί ένα σήμα διαφυγής (exit signal). Σε κάθε κύκλο της εφαρμογής, το γράφημα σκηνής πρέπει να εκτελεί τρεις βασικές λειτουργίες : την προσομοίωση συμπεριφοράς της εφαρμογής (app stage), το “κόψιμο” (culling) της γεωμετρίας που δεν είναι ορατή στο τρέχον οπτικό πεδίο (cull stage) και στην σχεδίαση των αποτελεσμάτων (draw stage). Κάθε ένα από τα τρία στάδια εκτελείται για ολόκληρο το γράφημα σκηνής ιεραρχικά, προτού μεταβεί στο επόμενο στάδιο, για να διατηρούνται σε συγχρονισμό οι κόμβοι.

#### *App stage*

Όλοι οι κόμβοι αρχικά εξετάζονται για εισερχόμενα μηνύματα. Κάθε κόμβος περιέχει μια ουρά εισόδου και έναν buffer εξόδου για την παραλαβή και εκτέλεση μηνυμάτων. Όλα τα εισερχόμενα μηνύματα εξετάζονται ένα προς ένα και για κάθε ένα που αναγνωρίζεται εκτελείται η αντίστοιχη ενέργεια. Αυτό γίνεται στην μέθοδο virtual void processMessage(char \* msg) που υπάρχει σε κάθε κλάση κόμβου. Η υπερκλάση Node3D::processMessages() είναι υπεύθυνη για την επαναλαμβανόμενη κλήση της processMessage() για όλα τα εισερχόμενα μηνύματα. Έπειτα, όλοι οι κόμβοι του γραφήματος σκηνής επεξεργάζονται την κατάσταση τους μέσω της μεθόδου τους app(). Η μέθοδος αυτή καθορίζει την συμπεριφορά κάθε κόμβου. Κατά την διάρκεια αυτή μπορεί να ενεργοποιηθούν νέα μηνύματα που προκαλούν την ενεργοποίηση μιας σημαίας γεγονότων (flag event) στην εσωτερική κατάσταση του κόμβου. Τελικά, όλοι οι κόμβοι εκτελούν μια επιθεώρηση των σημαιών, συγκρίνουν τα αποτελέσματα με τις

καταστάσεις των μηνυμάτων και στέλνουν τα κατάλληλα μηνύματα στους κόμβους αποδέκτες. Αυτή η διαδικασία εκτελείται μέσω της μεθόδου `dispatchMessages()`.

#### *Cull stage*

Αν και ο ρόλος αυτής της λειτουργίας δεν είναι απαραίτητος για την διαδικασία rendering, καθώς εκτελούνται απαιτητικές λειτουργίες με μεγάλη πολυπλοκότητα, η διαδικασία culling μπορεί να γλιτώσει πολύ χρόνο για μεγάλα περιβάλλοντα. Ο μηχανισμός της διαδικασίας είναι διαφορετικός για κάθε τύπο κόμβου και μερικές φορές απαιτούνται διαφορετικές στρατηγικές.

#### *Draw stage*

Η τελευταία λειτουργία στον κύκλο ζωής του SceneGraph. Όλες οι πληροφορίες και οι καταστάσεις κόμβων είναι έτοιμες να σχεδιαστούν. Ο κόμβος world παρέχει την μέθοδο του `draw()` και μέσω του πρωταρχικού του μέλους camera ο εικονικός κόσμος ξετυλίγεται στα μάτια των χρηστών.

### **6.4.4 Η γλώσσα περιγραφής μιας σκηνής του SceneGraph3D**

Η περιγραφή μιας σκηνής παρέχεται σε ένα αρχείο XML, του οποίου η δομή επιτρέπει δηλώσεις για βασικά στοιχεία, όπως γεωμετρία, μετασχηματισμοί και ομάδες, και παρέχει έναν μηχανισμό αποστολής μηνυμάτων όταν πραγματοποιείται κάποιο γεγονός για την ενεργοποίηση λειτουργιών ή αλλαγή καταστάσεων και ιδιοτήτων κόμβων.

Ο θεμελιώδης και υποχρεωτικός "λίθος" του αρχείου σκηνής είναι ο κόμβος world που αντιπροσωπεύει όλη την τρισδιάστατη σκηνή. Μέσα στις ετικέτες του βασικού αυτού κόμβου φωλιάζονται ορισμοί άλλων κόμβων οι οποίοι αναπαριστούν το περιεχόμενο τους στον εικονικό κόσμο. Με την ίδια λογική χτίζεται ολόκληρο το αρχείο σκηνής (με κόμβους φωλιασμένους μέσα σε άλλους, δηλαδή) και μπορούμε να δούμε τον νέο κόσμο.

Σχεδόν κάθε τύπος κόμβου σηματοδοτεί την πραγματοποίηση γεγονότων που αναπαριστούν κάποια αλλαγή στην κατάσταση του. Για παράδειγμα, ένας κόμβος τύπου proximity trigger παράγει ένα γεγονός εισόδου (enter) και εξόδου (exit) όταν ένας τουλάχιστον από τους καθορισμένους κόμβους εισέρχονται ή αφήνουν την περιοχή επίδρασης του.

Για την πρόσβαση ενός κόμβου, ώστε να λάβει ένα μήνυμα ή να χρησιμοποιηθεί ως χαρακτηριστικό σε κάποιο άλλο κόμβο, προσδιορίζουμε ένα όνομα (name). Η ονομασία

των κόμβων στο αρχείο XML είναι προαιρετική αλλά είναι χρήσιμη για τον σωστό εντοπισμό του κόμβου στόχου. Για την ρίζα του scene graph χρησιμοποιείται ένα ειδικό όνομα, το "%world". Τα ονόματα, οι τύποι κόμβων και τα χαρακτηριστικά είναι όλα case sensitive (πρέπει να γράφονται όπως ακριβώς είναι με ιδιαίτερη προσοχή στα κεφαλαία και στα μικρά, διαφορετικά δεν αναγνωρίζονται).

Ένα βασικό χαρακτηριστικό του σχεδιασμού scene graph είναι η "απεξάρτηση" των οντοτήτων χρήστη (user), κάμερας (camera) και εισόδου (input). Ένας αυθαίρετος αριθμός χρηστών, καμερών και εισόδων μπορεί να υπάρχει σε μια σκηνή, με τον καθένα να συνδέεται με κάποιον άλλο με όλους τους πιθανούς συνδυασμούς. Με αυτόν τον τρόπο παρέχεται ένας εύκολος τρόπος για την μοντελοποίηση ενός πολυχρηστικού περιβάλλοντος, όπου οι οντότητες user μπορούν να ελέγχονται από ανθρώπους (μέσω συσκευών εισόδου) ή αυτόματα. Ο ίδιος μηχανισμός ισχύει και για τους κόμβους camera, οι οποίοι μπορούν να ακολουθούν οποιονδήποτε άλλον κόμβο ή να στέκονται μόνοι τους. [40]

## Κεφάλαιο 7

### Η εφαρμογή

#### 7.1 Η βιβλιοθήκη OpenCV

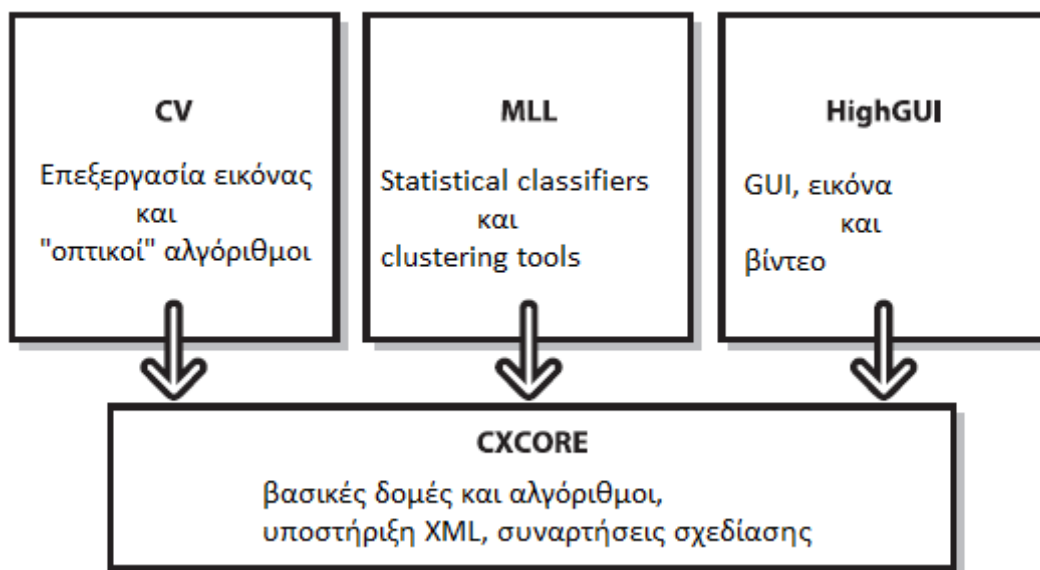
Πρόκειται για μια βιβλιοθήκη που αναπτύχθηκε από την Intel με σκοπό να βοηθήσει στην δημιουργία προγραμμάτων και εφαρμογών στο πεδίο computer vision. Μπορεί κανείς να την χρησιμοποιήσει ελεύθερα και δωρεάν υπό την άδεια της BSD. Λειτουργεί σε όλες τις πλατφόρμες και εστιάζει κυρίως στην επεξεργασία εικόνων πραγματικού χρόνου. Περιέχει, επίσης, περισσότερους από 500 αλγορίθμους εξειδικευμένους σε οπτικά προβλήματα. Είναι κυρίως γραμμένη σε γλώσσα προγραμματισμού C και C++. Ωστόσο, υπάρχουν συναρτήσεις “περιτυλίγματα” (wrapper functions) για άλλες γλώσσες, όπως η C# και η Python, ώστε να προσελκύσει ένα πιο ευρύ κοινό. [26, (όρος OpenCV)]



Εικόνα 72 : Το λογότυπο της OpenCV

Οι αλγόριθμοι της χρησιμοποιούν απλές βασικές μαθηματικές γνώσεις και δεν χρειάζεται κανείς να τα κατανοεί πλήρως προκειμένου να εισάγει συναρτήσεις στα προγράμματα του. Αναφέρονται κυρίως σε απλή άλγεβρα και βασική θεωρία πινάκων ενώ χρησιμοποιούν και μεθόδους βελτιστοποίησης προβλημάτων, κυρίως προσέγγιση ελαχίστων τετραγώνων, αλλά και βασικές γνώσεις γκαουσιανής κατανομής (Gaussian distribution) και παραγώγους απλών συναρτήσεων.

Η βιβλιοθήκη δομείται πάνω σε πέντε κύρια συστατικά. Το τμήμα CV περιέχει πληροφορίες για την βασική διαδικασία επεξεργασίας εικόνας και υψηλού επιπέδου αλγορίθμους του πεδίου computer vision. Το τμήμα ML είναι η βιβλιοθήκη εκμάθησης του υπολογιστή που περιέχει πολλούς στατιστικούς ταξινομητές και εργαλεία clustering. Το τμήμα HighGUI αποτελείται από ρουτίνες εισόδου/εξόδου και συναρτήσεις για την φόρτωση και αποθήκευση βίντεο και εικόνων. Το τμήμα CXCore περιέχει τις βασικές δομές πληροφοριών. Και τέλος το τμήμα CVaux που περιέχει πειραματικούς αλγορίθμους. [2]

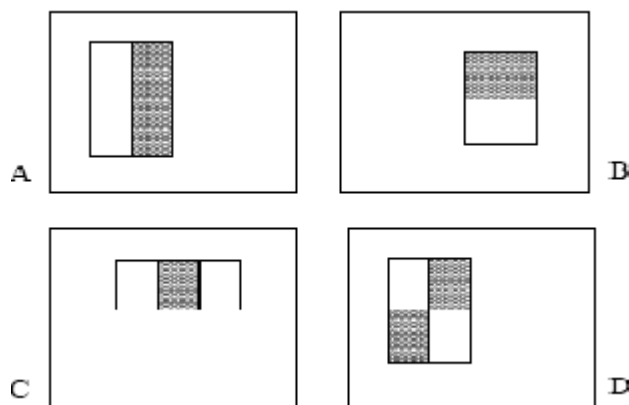


Εικόνα 73 : Η βασική δομή της OpenCV

### 7.1.1 OpenCV και ανίχνευση αντικειμένων

Η ανίχνευση αντικειμένων στην OpenCV στηρίζεται σε ένα σύστημα μηχανικής μάθησης το οποίο είναι ικανό να επεξεργάζεται πολύ γρήγορα μια εικόνα και να πετυχαίνει πολύ υψηλά ποσοστά επιτυχούς ανίχνευσης. Η διαδικασία της ανίχνευσης είναι πολύ πιο αποδοτική όταν βασίζεται στην ανίχνευση χαρακτηριστικών, τα οποία κωδικοποιούν πληροφορία σχετικά με το αντικείμενο που πρόκειται να ανιχνευτεί. Τα χαρακτηριστικά αυτά χρησιμοποιούνται σε σύνολα και κωδικοποιούν πληροφορία που αφορά την αντίθεση φωτεινότητας ανάμεσα σε περιοχές της εικόνας, τον προσανατολισμό της αντίθεσης και την χωρική σχέση αυτών των περιοχών.

Στην διαδικασία της ανίχνευσης υπάρχουν τρία είδη χαρακτηριστικών, τα δι-πολυγωνικά, τα τρι-πολυγωνικά και τα τετρα-πολυγωνικά. Η τιμή των δι-πολυγωνικών χαρακτηριστικών είναι η διαφορά ανάμεσα στο σύνολο των pixels που βρίσκονται μέσα στις δύο πολυγωνικές περιοχές. Η τιμή ενός τρι-πολυγωνικού χαρακτηριστικού υπολογίζεται αφαιρώντας από το άθροισμα των pixels των εξωτερικών πολυγώνων, το άθροισμα των pixels του εσωτερικού πολυγώνου. Τέλος, η τιμή των τετρα-πολυγωνικών χαρακτηριστικών υπολογίζεται από την διαφορά των pixels ανάμεσα στα διαγώνια μέρη του πολυγώνου.



Εικόνα 74 : Παραδείγματα χαρακτηριστικών. Τα A και B είναι δι-πολυγωνικά, το C τρι-πολυγωνικό και το D τετρα-πολυγωνικό

Για τον γρήγορο υπολογισμό των χαρακτηριστικών, σε οποιαδήποτε κλίμακα ή περιοχή, το σύστημα δεν χρησιμοποιεί κατευθείαν τις τιμές φωτεινότητας της εικόνας, αλλά μια ενδιάμεση αναπαράσταση της που ονομάζεται πλήρης εικόνα. Η πλήρης εικόνα μπορεί να υπολογιστεί κατευθείαν από την αρχική εικόνα με χρήση απλών πράξεων στις τιμές των pixels της. Συγκεκριμένα, ο τρόπος υπολογισμού της πλήρους εικόνας γίνεται με την χρήση της σχέσης :

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y'),$$

όπου  $ii(x,y)$  η τιμή της πλήρους εικόνας στο σημείο  $(x,y)$  και  $i(x,y)$  η τιμή της πραγματικής εικόνας στο ίδιο σημείο. Η πλήρης εικόνα μπορεί να υπολογιστεί με ένα μόνο πέρασμα πάνω από την πραγματική εικόνα.

Η διαδικασία ανίχνευσης αντικειμένων που χρησιμοποιεί η OpenCV προτάθηκε αρχικά από τον Paul Viola και υλοποιήθηκε από τον Rainer Lienhart. Αρχικά ένας ταξινομητής εκπαιδεύεται με ένα μεγάλο αριθμό θετικών δειγμάτων (εικόνες που περιέχουν το αντικείμενο προς ανίχνευση και έχουν όλες το ίδιο μέγεθος) και με αρνητικά δείγματα (αυθαίρετες εικόνες που δεν περιέχουν το αντικείμενο και είναι ίδιου μεγέθους με τα θετικά δείγματα).

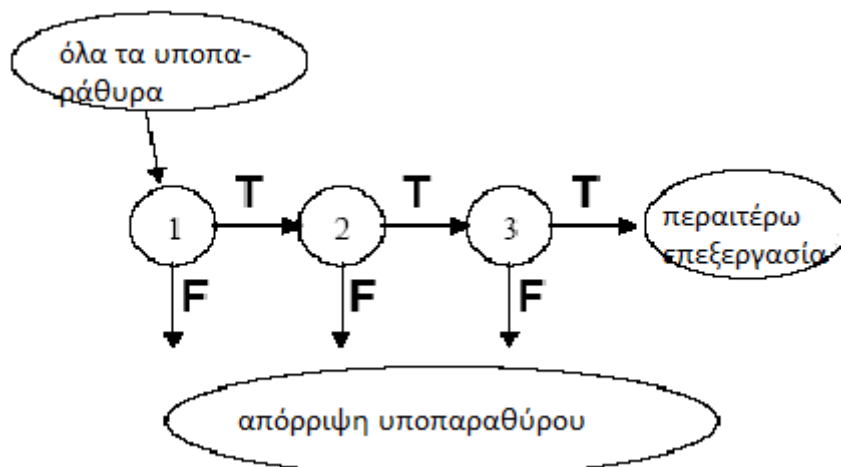
Από κάθε θετικό δείγμα εξάγεται ένας πολύ μεγάλος αριθμός χαρακτηριστικών. Πειραματικά έχει αποδειχθεί ότι ένα πολύ μικρό σύνολο αυτών των χαρακτηριστικών είναι αρκετό για την δημιουργία ενός πολύ αποδοτικού ταξινομητή. Για την εξαγωγή αυτών των χαρακτηριστικών χρησιμοποιείται ένας αλγόριθμος μάθησης ο οποίος έχει σχεδιαστεί να επιλέγει μόνο τα πολυγωνικά χαρακτηριστικά τα οποία ανταποκρίνονται καλύτερα στα θετικά και στα αρνητικά δείγματα.



### 7.1.2 Διαδικασία ανίχνευσης

Από τη στιγμή που έχουν δημιουργηθεί οι ταξινομητές μπορεί να ξεκινήσει η διαδικασία ανίχνευσης αντικειμένων σε μια εικόνα ή ένα βίντεο. Η ανίχνευση αντικειμένων γίνεται με εφαρμογή ταξινομητών στην εικόνα. Η εφαρμογή των ταξινομητών έχει έξοδο 1 αν η περιοχή στην οποία εφαρμόστηκε μοιάζει να περιέχει το αντικείμενο προς αναγνώριση αλλιώς η έξοδος είναι 0. Για την ανίχνευση σε όλη την εικόνα ο ταξινομητής εξετάζει την εικόνα σειριακά με ένα κυλιόμενο παράθυρο. Επίσης, ο ταξινομητής έχει σχεδιαστεί έτσι ώστε εύκολα να μπορεί να μεταβάλλει το μέγεθος του παραθύρου που εξετάζει προκειμένου να μπορεί να ανιχνεύει αντικείμενα διαφορετικών μεγεθών. Το μέγεθος του υποπαραθύρου μεταβάλλεται κάθε φορά που εξετάζεται όλη η εικόνα και με συγκεκριμένο ρυθμό. Ο ρυθμός μεταβολής του μεγέθους του υποπαραθύρου ορίζεται από τον προγραμματιστή ή από τον χρήστη.

Η διαδικασία της ανίχνευσης σε μια εικόνα χωρίζεται σε στάδια. Αρχικά, στην εικόνα εφαρμόζονται οι πιο απλοί ταξινομητές ώστε να απορριφθεί η πλειοψηφία των υποπαραθύρων της εικόνας τα οποία δεν περιέχουν το αντικείμενο προς ανίχνευση και στη συνέχεια εφαρμόζονται ταξινομητές πιο περίπλοκοι που σκοπό έχουν την μείωση του ποσοστού λάθους ανίχνευσης. Αν κάποιο υποπαραθύρο της εικόνας απορριφθεί κατά την διαδικασία ανίχνευσης δεν επανεξετάζεται από περίπλοκους ταξινομητές. [59]



Εικόνα 75 : η διαδικασία ανίχνευσης

### 7.1.3 Χρησιμοποιούμενες δομές OpenCV στην εφαρμογή μας

#### *IplImage*

Πίνακας στον οποίο αποθηκεύονται εικόνες. Μπορεί να έχει περισσότερα του ενός κανάλια η εικόνα.

#### *CvSize*

Δομή για την αποθήκευση του μεγέθους των εικόνων.

#### *CvScalar*

Δομή που έχει 4 μεταβλητές double, μία για κάθε κανάλι μιας εικόνας.

#### *CvCapture*

Δομή απαραίτητη για την "σύλληψη" εικόνας πραγματικού χρόνου από μια webcam.

#### *CvSeq*

Αποτελεί την βάση για όλες τις δυναμικές δομές δεδομένων της OpenCV.

#### *CvRect*

Η βασική δομή για δισδιάστατα ορθογώνια.

#### *CvHaarClassifierCascade*

Δομή που χρησιμοποιείται για την αναπαράσταση των ταξινομητών.

#### *CvMemStorage*

Μία χαμηλού επιπέδου δομή που χρησιμοποιείται για την αποθήκευση δυναμικά αυξανόμενων δομών, όπως ακολουθίες, περιγράμματα (contours) και γραφήματα. Είναι οργανωμένη σαν μια λίστα τμημάτων μνήμης ίδιου μεγέθους.

### 7.1.4 "Σύλληψη" εικόνας από μία webcam

Προτού ξεκινήσουμε να γράφουμε κώδικα σε OpenCV δεν πρέπει να ξεχνάμε την εισαγωγή των κατάλληλων αρχείων επικεφαλίδας της βιβλιοθήκης. Για αυτό στην αρχή του κώδικα βρίσκονται οι εντολές `#include "cv.h"`, `#include "cxcore.h"` και `#include "highgui.h"`.

Στην συνέχεια γίνονται οι απαραίτητες ενέργειες προκειμένου να ανοίξει η κάμερα. Αρχικοποιούμε την δομή στην οποία θα αποθηκευτεί η “σύλληψη” με την εντολή `CvCapture* capture = cvCaptureFromCAM(0);`. Ελέγχουμε εάν η αρχικοποίηση έγινε σωστά με τις εντολές :

```
if ( !capture ) {
    fprintf( stderr, "Cannot open initialize webcam!\n" );
    return;
}
```

Τώρα γίνεται η “σύλληψη” ενός πλαισίου από την κάμερα :

```
CvSize sz = cvGetSize(cvQueryFrame( capture));
```

Η εντολή αυτή επιστρέφει το ύψος και το πλάτος του παραθύρου της κάμερας.

Στη συνέχεια με την εντολή:

```
IplImage* src = cvCreateImage( sz, 8, 3 );
```

γίνεται αρχικοποίηση του παραθύρου στο οποίο θα εμφανιστεί το αποτέλεσμα της κάμερας. Οι παράμετροι που δέχεται η συνάρτηση είναι από αριστερά προς δεξιά : `CvSize size`, το πλάτος και το ύψος του παραθύρου, `int depth`, το βάθος των στοιχείων της εικόνας σε bits, `int channels`, ο αριθμός των καναλιών ανά pixel. Μπορεί να είναι 1, 2, 3 ή 4.

Επειτα γίνεται έλεγχος για το αν έγινε “σύλληψη” κάποιου πλαισίου:

```
if( !cvGrabFrame( capture ) ) {
    fprintf( stderr, "Could not grab a frame!\n" );
    return;
}
```

Η ανάκτηση του πλαισίου γίνεται με την εντολή:

```
src = cvRetrieveFrame( capture );
```

Ενώ η εμφάνιση του παραθύρου της κάμερας και η ονομασία του σε `src` με την :

```
cvShowImage( "src", src);
```

Και τέλος απελευθερώνεται η κάμερα :

```
cvReleaseCapture(&capture); [2]
```

### 7.1.5 Κώδικας για τις δικές μας ανιχνεύσεις

#### *Ανίχνευση ματιών – Ανίχνευση προσώπου*

Η βιβλιοθήκη OpenCV παρέχει έναν ενδιαφέροντα τρόπο για αναγνώριση προσώπου. Πιο συγκεκριμένα, διαθέτει προγράμματα (ή συναρτήσεις) που χρησιμοποιούνται για την “εκπαίδευση” ταξινομητών στο σύστημα αναγνώρισης προσώπου, το οποίο ονομάζεται HaarTraining. Παράγει ένα xml cascade αρχείο, το οποίο χρησιμοποιείται στην ανίχνευση αντικειμένων κατά την διαδικασία μάθησης. Υπάρχουν μερικά βασικά χαρακτηριστικά στο αρχείο αυτό : αριθμός σταδίων και τύπος ταξινομητών σε κάθε στάδιο. Η OpenCV προσφέρει έτοιμα τέτοια αρχεία για ανίχνευση ματιού, ανίχνευση ματιού όταν ο χρήστης φορά γυαλιά οράσεως, προσώπου (είτε προφίλ είτε κανονικά) και σώματος. Βρίσκονται στον φάκελο OpenCV/data/haarcascades και μπορεί κανείς να τα χρησιμοποιήσει στα προγράμματα του. Ωστόσο, μπορεί κανείς να δημιουργήσει τα δικά του με ανάλογη λογική (να συντάξει ένα xml αρχείο αφού έχει συγκεντρώσει τον απαραίτητο αριθμό θετικών και αρνητικών δειγμάτων).

Έτσι, λοιπόν, εμείς χρησιμοποιήσαμε το αρχείο haarcascade\_eye.xml για την ανίχνευση του ματιού του χρήστη και το αρχείο haarcascade\_frontalface\_alt\_tree.xml για την ανίχνευση του προσώπου του. Χρησιμοποιήσαμε μια μεταβλητή συμβολοσειράς για την ανάγνωση του κατάλληλου xml αρχείου. Η δήλωση της έγινε με την εντολή `const char* filename;`. Έτσι, στην περίπτωση ανίχνευσης ματιού η μεταβλητή παίρνει την τιμή `filename = "C:\\Program Files\\OpenCV\\data\\haarcascades\\haarcascade_eye.xml";` και στην περίπτωση ανίχνευσης προσώπου την τιμή `filename = "C:\\ProgramFiles\\OpenCV\\data\\haarcascades\\haarcascade_frontalface_alt_tree.xml";`. Η τιμή είναι το μονοπάτι (path) το οποίο ακολουθούμε για να βρούμε τα αντίστοιχα αποθηκευμένα αρχεία. Μπορούμε, βέβαια, να αλλάξουμε την θέση αποθήκευσης τους και να τα βάλουμε στον ίδιο φάκελο που βρίσκεται το αρχείο κώδικα που γράφουμε. Και στην συνέχεια φορτώνεται το αρχείο με την εντολή `cascade = ( CvHaarClassifierCascade* )cvLoad( filename, 0, 0, 0 );`. Η συνάρτηση `cvLoad()` επιστρέφει αποτέλεσμα τύπου `IntPtr`, για αυτό και χρησιμοποιείται η διαδικασία μετατροπής τύπου (type casting). Οι παράμετροι που δέχεται είναι με τη σειρά : `string fileName`, το όνομα του αρχείου που θα φορτωθεί, `IntPtr memstorage`, μνήμη για δυναμικές δομές, `string name`, εναλλακτικό όνομα του αντικειμένου και `IntPtr realName`, θα περιέχει το όνομα του αντικειμένου που θα φορτωθεί. Αρχικοποιείται ο `buffer` μνήμης, που είναι απαραίτητος για τον ανιχνευτή, με την εντολή `CvMemStorage`

\*storage = cvCreateMemStorage(0); Τώρα μπορεί να γίνει η ανίχνευση, η οποία υλοποιείται σε ξεχωριστή συνάρτηση, και για αυτό καλείται η συνάρτηση αυτή detectFaces( src );. Δέχεται ως παράμετρο την εικόνα ή την λήψη από την webcamera στην οποία θα γίνει η ανίχνευση και είναι τύπου void αφού δεν επιστρέφει κάποιο αποτέλεσμα.

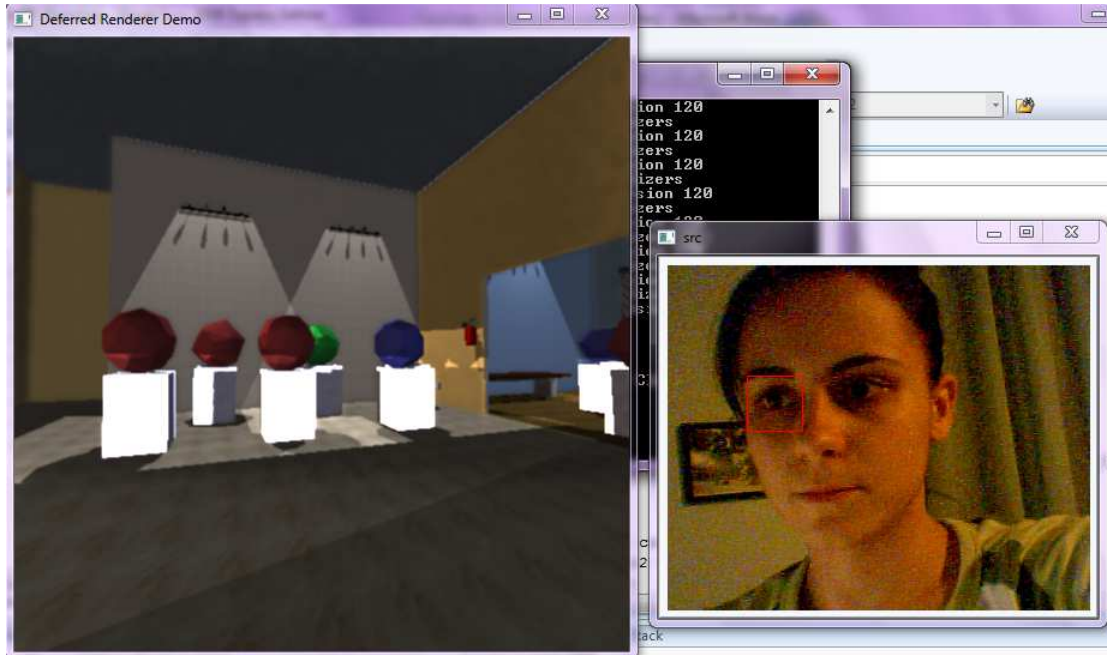
Θα ξεκινήσει η ανίχνευση και ο ανιχνευτής θα τοποθετήσει την λίστα με τα ανιχνευμένα πρόσωπα/μάτια στην μεταβλητή faces. Η εντολή που χρησιμοποιείται είναι η

```
CvSeq *faces = cvHaarDetectObjects(img, cascade, storage, 1.1, 3, 0, cvSize( 40, 40 ) );,
```

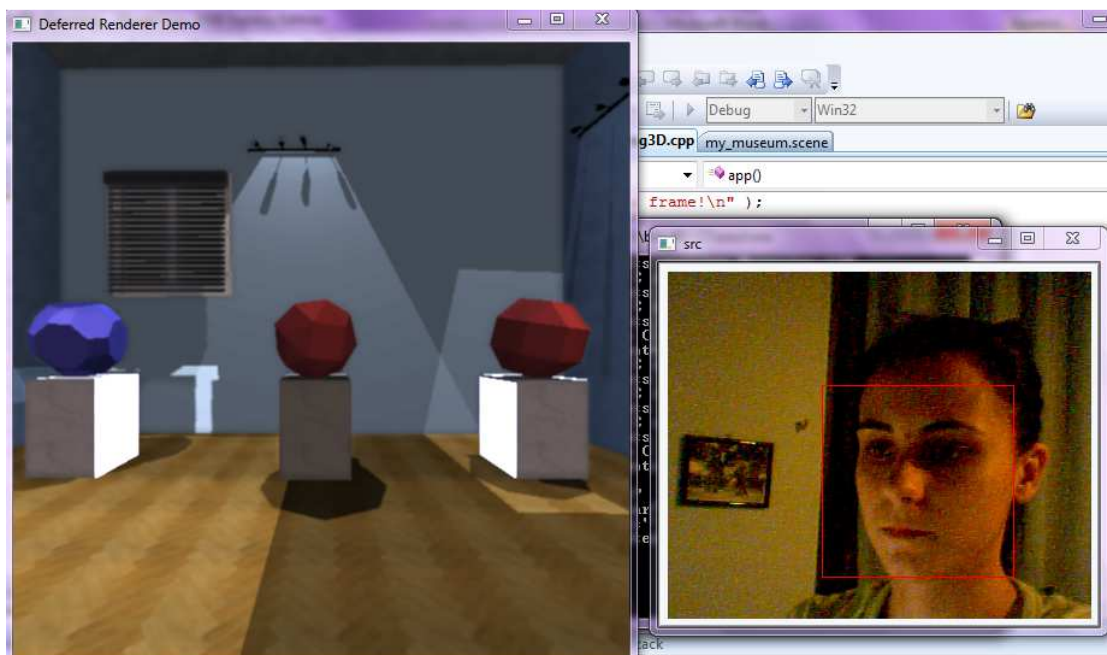
όπου img είναι η λήψη από την webcamera, cascade το αρχείο xml που φορτώσαμε προηγουμένως, storage ο buffer μνήμης, 1.1, ο παράγοντας κλίμακας, δηλαδή, κατά πόσο θα κλιμακωθεί το παράθυρο έρευνας κατά τις στιγμές έρευνας (το 1.1 σημαίνει ότι το παράθυρο θα αυξηθεί κατά 10%), 3, ο ελάχιστος αριθμός γειτονικών ορθογωνίων που δημιουργούν ένα αντικείμενο (όλες οι ομάδες μικρότερου αριθμού ορθογωνίων από τον αριθμό αυτό-1 απορρίπτονται), 0 ο τύπος λειτουργίας (προς το παρόν ο μόνος τύπος είναι ο CV\_HAAR\_DO\_CANNY\_PRUNING, είναι κάτι σαν τιμή κατωφλίου για την απόρριψη περιοχών που δεν περιέχονται στο αντικείμενο που θέλουμε να ανιχνευτεί) και τέλος cvSize( 40, 40 ), το ελάχιστο μέγεθος παραθύρου.

Αφού έγινε η ανίχνευση πρέπει να εμφανιστεί κάτι στην οθόνη για να αντιληφθεί ο χρήστης εάν όντως το αντικείμενο ανιχνεύτηκε. Εμείς σχεδιάζουμε ένα κόκκινο ορθογώνιο που περικλείει την περιοχή που ανιχνεύτηκε. Αυτό γίνεται για κάθε πρόσωπο/μάτι που ανιχνεύτηκε με τις εντολές :

```
for( i = 0 ; i < ( faces ? faces->total : 0 ) ; i++ ) {
    CvRect *r = ( CvRect* )cvGetSeqElem( faces, i );
    cvRectangle( img, cvPoint( r->x, r->y ),
                cvPoint( r->x + r->width, r->y + r->height ),
                CV_RGB( 255, 0, 0 ), 1, 8, 0 );
}
```



Εικόνα 76 : Ανίχνευση του ματιού του χρήστη



Εικόνα 77 : Ανίχνευση του προσώπου του χρήστη

### Ανίχνευση χεριού

Για την ανίχνευση χεριού χρησιμοποιήσαμε διαφορετική μέθοδο. Σε αυτήν την περίπτωση βασιστήκαμε στην χρήση contours (περιγράμματα). Η ανίχνευση με την μέθοδο αυτή συναντάται σε πολλά προβλήματα πραγματικού χρόνου στο πεδίο

computer vision και είναι μία ακόμη μέθοδος που προσφέρει η OpenCV με ικανοποιητικά αποτελέσματα .

Η μέθοδος αυτή βασίζεται στην κατάλληλη εφαρμογή καμπυλών που μπορεί να βοηθήσουν σημαντικά στην κατάτμηση των αντικειμένων, δηλαδή, στην αναγνώριση των σχημάτων και των ακμών (edges) που υπάρχουν στην εικόνα. Ένα αντικείμενο διαχωρίζεται από το υπόλοιπο μέρος της εικόνας εξαιτίας μιας απότομης αλλαγής στη φωτεινότητα ή στο χρώμα. Σε κάθε περίπτωση, υπάρχει κάποια συννοριακή γραμμή ή καμπύλη η οποία αποτελεί το σύνορο του αντικειμένου και σχηματίζει το περίγραμμα του.

Ακμή κάθε pixel μιας εικόνας είναι μια ιδιότητα του συγκεκριμένου pixel της εικόνας και υπολογίζεται με βάση την πληροφορία σχετικά με τη γειτονιά του αντίστοιχου pixel. Ανιχνεύονται οι ακμές στην εικόνα και δείχνουν ασυνέχειες όσον αφορά την διαβάθμιση του γκρι, το χρώμα, την υφή και άλλα. Υπάρχουν αρκετές συναρτήσεις στην βιβλιοθήκη για την ανίχνευση των ακμών. Η πιο γνωστή, όμως, είναι ο τελεστής Canny. Παίρνει ως είσοδο μία εικόνα σε γκριζα μορφή και δίνει ως έξοδο μία εικόνα στην οποία παρουσιάζονται οι ασυνέχειες που ανιχνεύτηκαν στην αρχική εικόνα. Όμως οι εικόνες που παίρνουμε μόνο μέσω ανίχνευσης ακμών δεν μπορούν να χρησιμοποιηθούν για κατάτμηση εικόνας. Πρέπει, δηλαδή, να αναπτυχθούν επιπρόσθετες διαδικασίες οι οποίες να αντιστοιχούν τις κατάλληλες ακμές με τα σύνορα των περιοχών τις οποίες θέλουμε να εξάγουμε.

Στη συνέχεια ακολουθεί η διαδικασία thresholding. Κατά την διαδικασία αυτή εξετάζεται το χρώμα κάθε pixel. Χρησιμοποιούμε μια τιμή κατώφλι και καθορίζεται εάν το pixel έχει το χρώμα του δέρματος και άρα ανήκει στο χέρι ή όχι. Πώς όμως καθορίζεται η τιμή αυτή; Βρίσκουμε τις τιμές RGB για το χρώμα του δέρματος και εξετάζουμε όλους τους πιθανούς σχεδιασμούς που μπορεί να οδηγούν σε χρώμα κοντά σε αυτό που θέλουμε. Για το λόγο αυτό κατασκευάζουμε ένα γράφημα και μέσα από αυτό καταλήγουμε σε μία τιμή κατώφλι για τις μέγιστες τιμές χρώματος και μία τιμή για τις ελάχιστες. Στο τέλος της διαδικασίας παίρνουμε μια εικόνα ασπρόμαυρη με το περίγραμμα του ανιχνευμένου αντικειμένου.

Στον κώδικα μας χρησιμοποιήσαμε για τους σκοπούς της μεθόδου αυτής τις μεταβλητές `IplImage* hsv_image`, `IplImage* hsv_mask`, `IplImage* hsv_edge`, `CvScalar hsv_min` και `CvScalar hsv_max`. Αρχικοποιήσαμε τα αρχικά τους παράθυρα με την εντολή `cvCreateImage( sz, 8, 1);` για τις τρεις πρώτες. Για την ελάχιστη τιμή κατωφλίου χρησιμοποιήσαμε την εντολή `cvScalar(0, 30, 80, 0);` ενώ για την μέγιστη την εντολή `cvScalar(20, 150, 255, 0);`. Στη συνέχεια, μετατρέπουμε την εικόνα από μορφή RGB σε μορφή HSV, καθώς η μορφή αυτή δεν επηρεάζεται τόσο από τον φωτισμό.

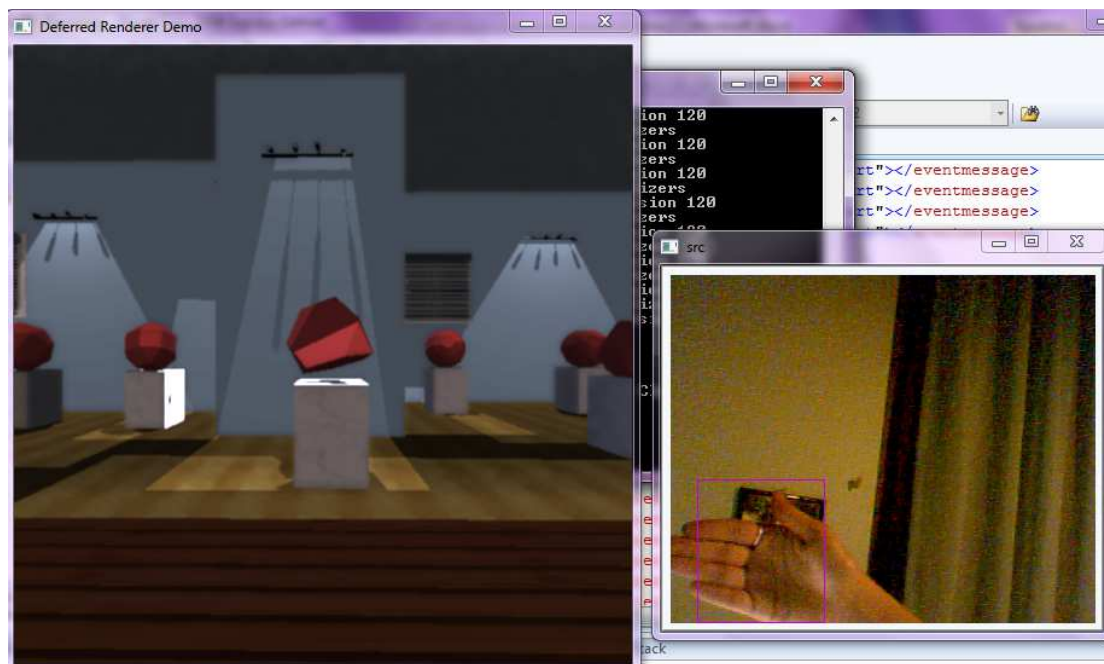
```
cvCvtColor(src, hsv_image, CV_BGR2HSV);
cvInRangeS (hsv_image, hsv_min, hsv_max, hsv_mask);
cvSmooth( hsv_mask, hsv_mask, CV_MEDIAN, 27, 0, 0 );
cvCanny(hsv_mask, hsv_edge, 1, 3, 5);
Και τέλος βρίσκουμε το περίγραμμα :
while(contours)
{
    result = fabs( cvContourArea( contours, CV_WHOLE_SEQ ) );
    if ( result > result2) {result2 = result; contours2 = contours;};
    contours = contours->h_next;
}
```

και σχεδιάζουμε το ορθογώνιο στην περιοχή ενδιαφέροντος :

```
if ( contours2 )
{
    rect = cvBoundingRect( contours2, 0 );
    cvRectangle( src, cvPoint(rect.x, rect.y + rect.height), cvPoint(rect.x + rect.width,
rect.y), CV_RGB(200, 0, 200), 1, 8, 0 );
}
```

Τα περισσότερα προβλήματα, όμως, στους αλγορίθμους αυτούς οφείλονται στην ύπαρξη θορύβου που υπάρχει στην εικόνα. Επίσης σημαντικό πρόβλημα είναι η μη κατάλληλη πληροφορία σε μια εικόνα, δηλαδή η ύπαρξη ακμών όπου δεν υπάρχουν σύνορα αντικειμένων και το αντίστροφο. Ακόμα η αλλαγή φωτισμού έχει καταστροφικές συνέπειες. Σχεδόν όλα τα pixels της εικόνας αλλάζουν κατά πολύ. Επομένως ο φωτισμός πρέπει να παραμένει σταθερός συνεχώς προκειμένου να μην αλλοιώνονται τα αποτελέσματα. Τέλος ένα κακής ποιότητας μέσο λήψης εικόνας έχει ως αποτέλεσμα την εισαγωγή θορύβου και κατά συνέπεια την αλλοίωση των αποτελεσμάτων. [31][32][33]

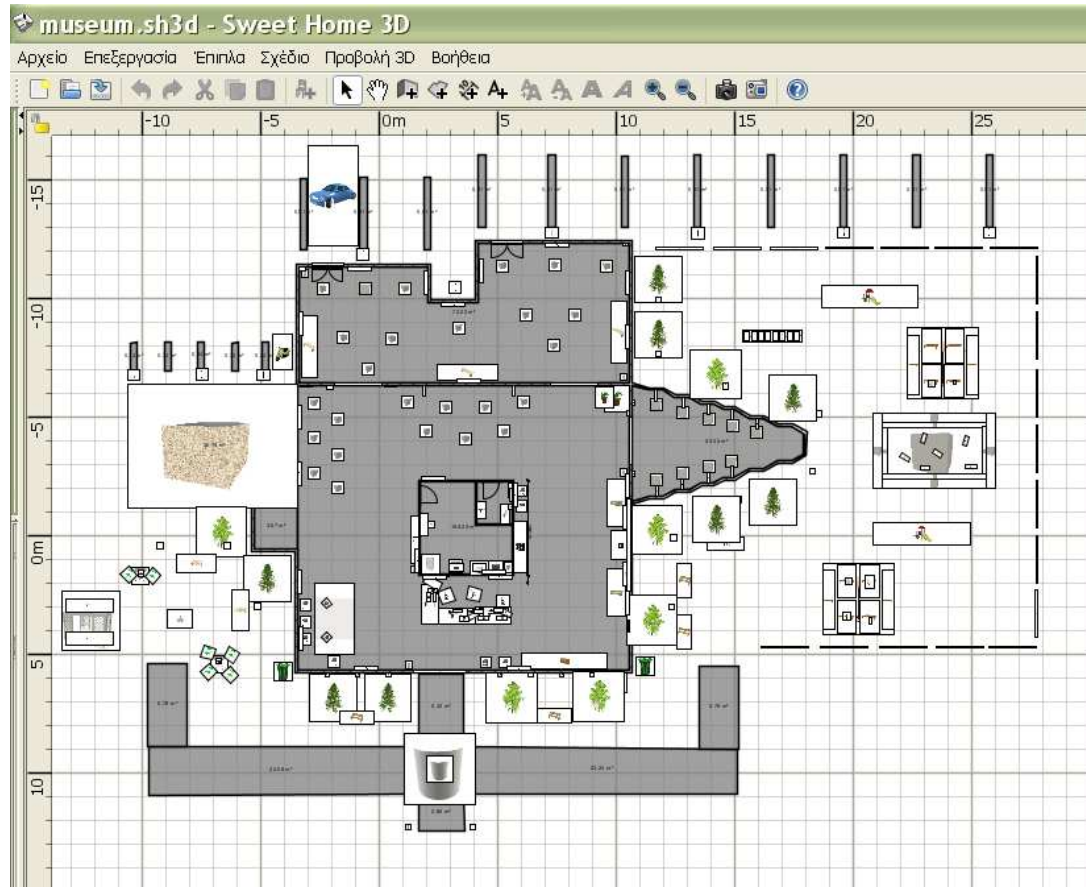




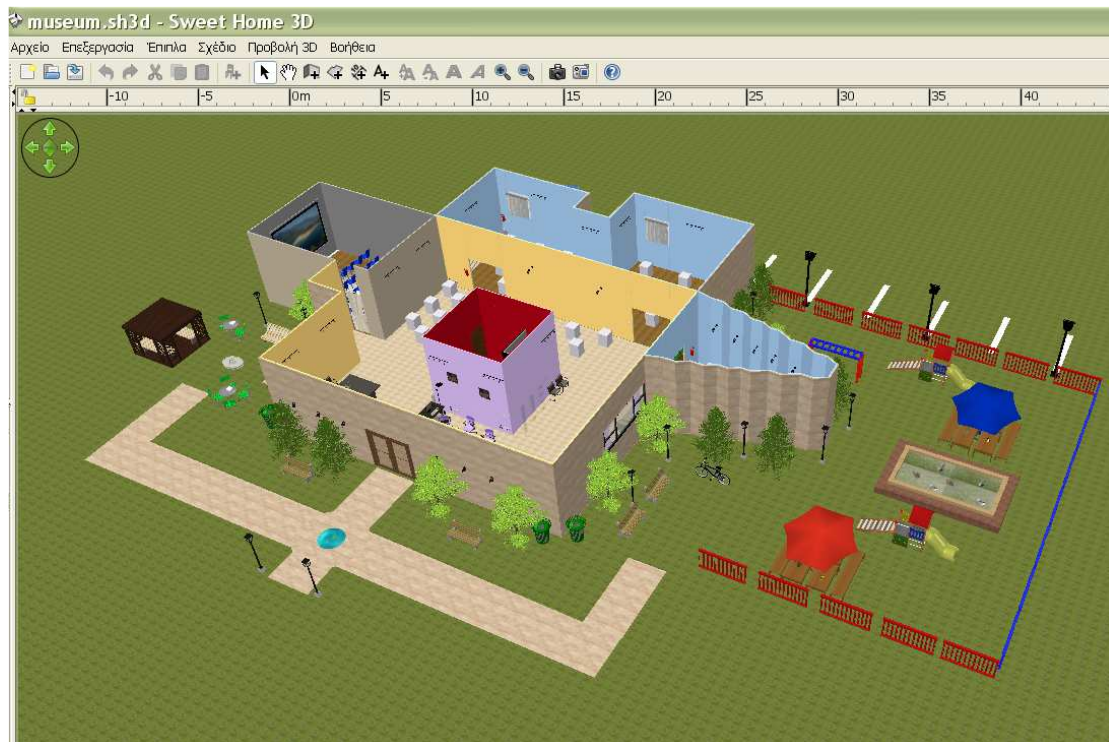
Εικόνα 78 : Ανίχνευση του χεριού του χρήστη

## 7.2 Σχεδιασμός εικονικού μουσείου

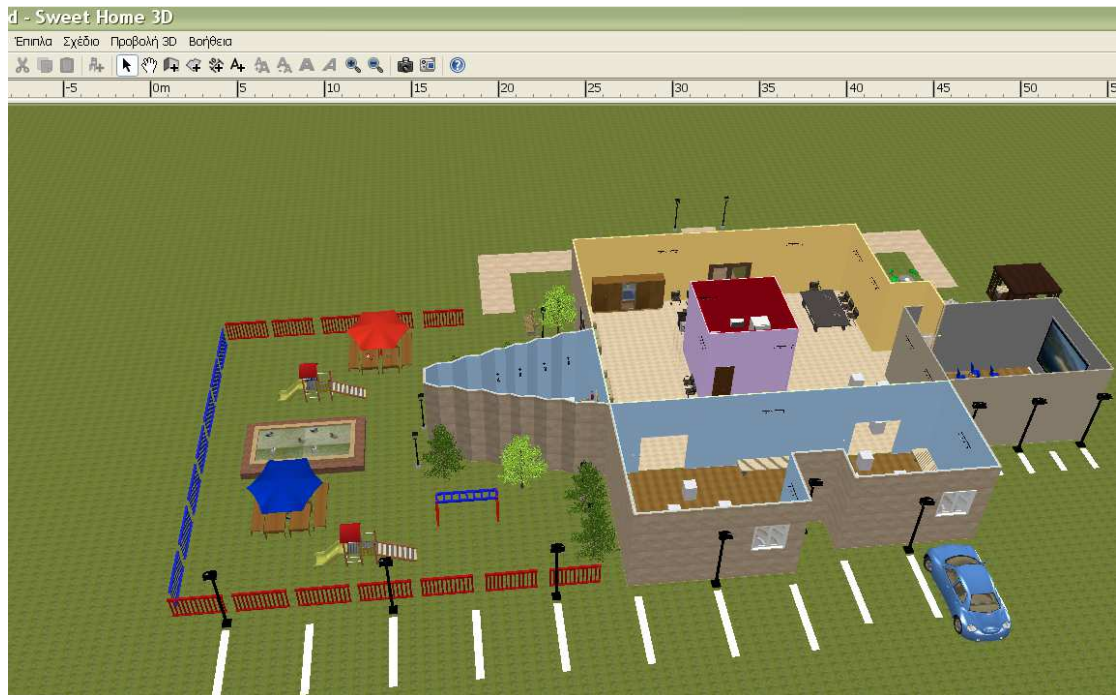
Ο σχεδιασμός του μουσείου έγινε χρησιμοποιώντας το πρόγραμμα SweetHome3D. Αξιοποιώντας τις δυνατότητες που παρέχει το πρόγραμμα αυτό μέσα από τα διαθέσιμα εργαλεία, κατασκευάζουμε το εσωτερικό του μουσείου και τον περιβάλλοντα χώρο. Όσον αφορά το εσωτερικό του μουσείου, αποτελείται από την αίθουσα υποδοχής, με χώρους αναμονής και ξεκούρασης των επισκεπτών, μια αίθουσα εικονικής πραγματικότητας στα αριστερά του μουσείου, έναν ειδικά διαμορφωμένο χώρο για τις ανάγκες και την εξυπηρέτηση των εργαζομένων στο μουσείο και αίθουσες με τα εκθέματα. Τα εκθέματα του μουσείου αποτελούν τρεις κατηγορίες στερεών. Στην κυρίως αίθουσα βρίσκονται τα καταλανικά στερεά, στη δεξιά αίθουσα τα πλατωνικά και τα στερεά Kepler-Poinsot, ενώ στην τελευταία αίθουσα τα αρχιμήδεια στερεά. Κάδρα, διακοσμητικά αντικείμενα, φυτά, προβολείς, παράθυρα και πυροσβεστήρες στοχεύουν να προσομοιώσουν το εσωτερικό ενός πραγματικού μουσείου. Όσον αφορά τον εξωτερικό χώρο, περιλαμβάνει χώρους αναψυχής, μικρό πάρκο για τους νεαρούς επισκέπτες και ειδικούς χώρους παρκαρίσματος. Τα αντικείμενα είτε διατήρησαν το χρώμα με το οποίο παρέχονται από το πρόγραμμα, είτε τους αποδόθηκε κάποιο άλλο επιθυμητό χρώμα, είτε τους αποδόθηκε υφή (μέσω κατάλληλων textures) για πιο ρεαλιστική απεικόνιση των αντικειμένων.



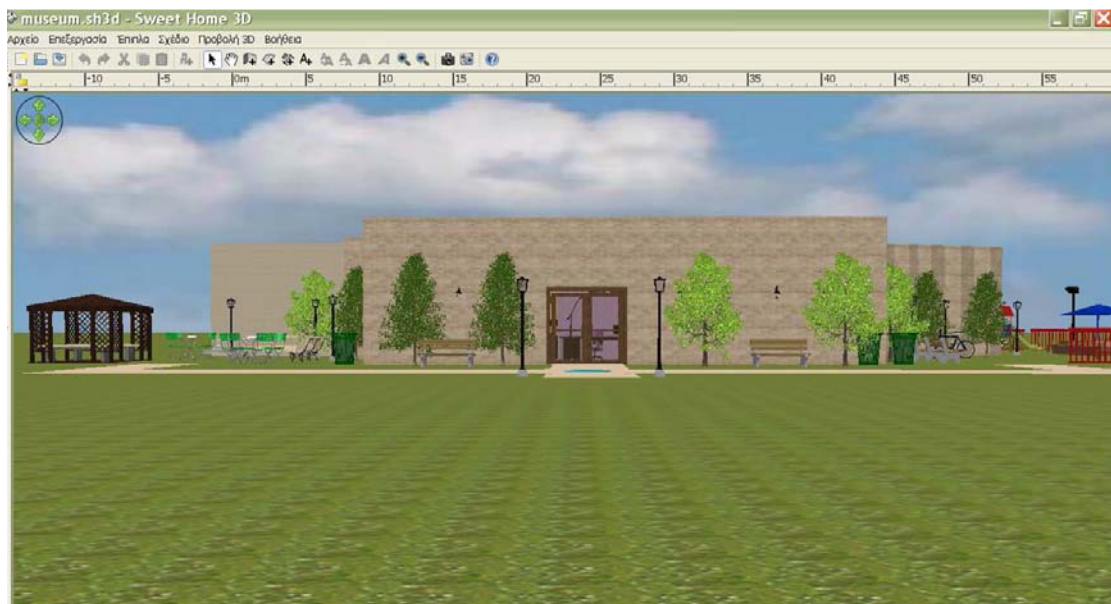
Εικόνα 79 : Η κάτοψη του μουσείου, όπως παρουσιάζεται μέσα από το πρόγραμμα Sweethome3D



Εικόνα 80 : Τρισδιάστατη αναπαράσταση του μουσείου μέσω της δυνατότητας του προγράμματος Sweethome3D για θέαση από ψηλά.



Εικόνα 81 : Τρισδιάστατη αναπαράσταση του μουσείου από την πίσω πλευρά.



Εικόνα 82 : Η είσοδος του μουσείου με την επιλογή “εικονική επίσκεψη” (virtual visit).

Μετά το σχεδιασμό του μουσείου, αφού δηλαδή έχουμε ολοκληρώσει την δημιουργία του αρχείου museum.sh3d, το μετατρέπουμε σε μορφή .obj προκειμένου να μπορούμε να το εισάγουμε στο πρόγραμμα EaZD. Τα στερεά-εκθέματα του μουσείου τα έχουμε πάρει από το διαδίκτυο σε μορφή .obj και δεν τα εισαγάγαμε στο SweetHome3D ως μοντέλα (για αυτό και δεν εμφανίζονται στο πρόγραμμα αυτό). Ο λόγος που έγινε είναι για να έχουμε την δυνατότητα επεξεργασίας τους από το EaZD. Με κατάλληλους

μετασχηματισμούς, που έγιναν μέσω του αρχείου σκηνης, τα στερεά εμφανίζονται στις κατάλληλες θέσεις και με το μέγεθος που εμείς θέλουμε. Το EaZD δημιουργεί το τρισδιάστατο εικονικό περιβάλλον του μουσείου μέσα στο μέγα στο οποίο μπορεί να περιηγηθεί ο χρήστης. Εκεί θα παρατηρήσει τα στερεά του μουσείου τα οποία περιστρέφονται.

Ακολουθούν κάποια στιγμιότυπα από την περιήγηση στο εικονικό μουσείο μέσω του προγράμματος EaZD.



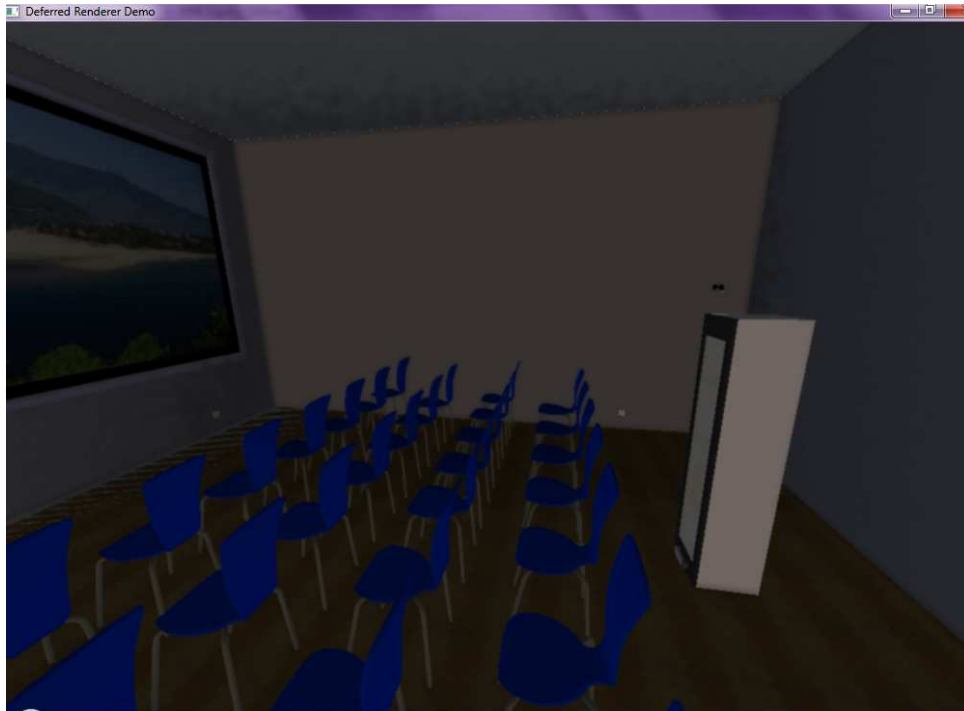
Εικόνα 83 : η αίθουσα με κάποια από τα αρχιμήδεια στερεά



Εικόνα 84 : η αίθουσα με τα πλατωνικά και τα Kepler-Poinsot στερεά



Εικόνα 85 : Η είσοδος του μουσείου και στο βάθος φαίνονται μερικά από τα καταλανικά στερεά



Εικόνα 86 : η αίθουσα εικονικής πραγματικότητας



Εικόνα 87 : χώρος αναμονής στην είσοδο του μουσείου και στο βάθος η αίθουσα με τα αρχιμήδεια στερεά



Εικόνα 88 : μία από τις εξωτερικές πλευρές του μουσείου

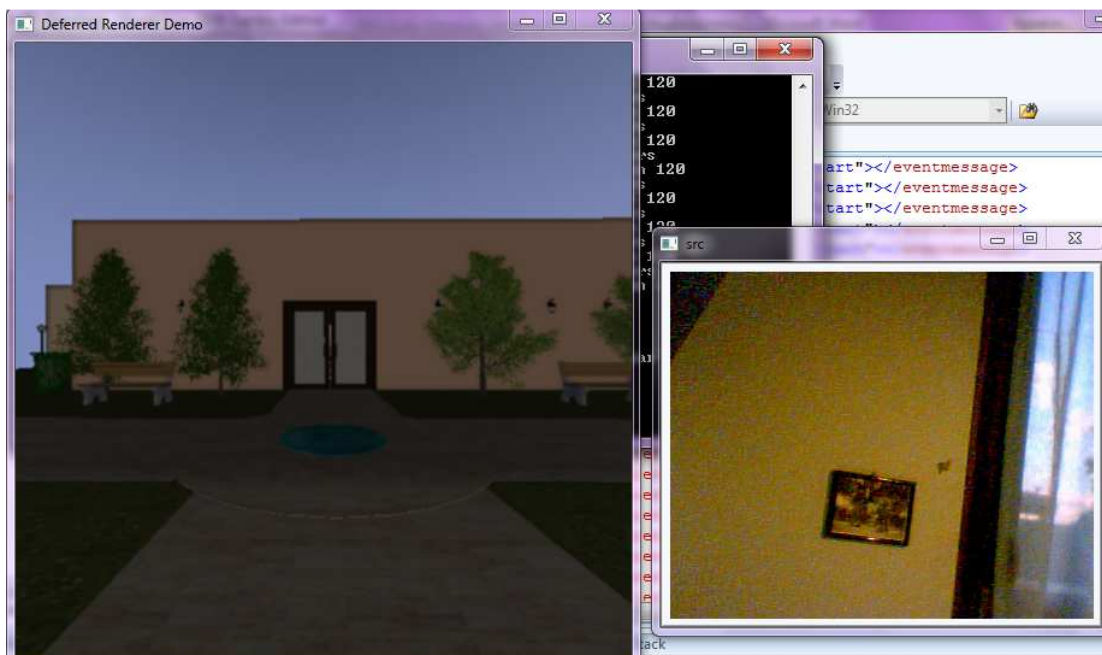
## 7.3 Σύνδεση αρχείου κάμερας με το EaZD

### 7.3.1 Αρχικά βήματα

Προκειμένου να ανοίξει η κάμερα μαζί με το εικονικό περιβάλλον του μουσείου, δηλαδή μέσα από το πρόγραμμα EaZD, πρέπει να προσθέσουμε έναν καινούριο κόμβο scene graph. Η ονομασία της κλάσης του νέου κόμβου ακολουθεί κάποιους κανόνες. Πρέπει το όνομα της να έχει κατάληξη "3D" και το αρχείο επικεφαλίδας αλλά και το αντίστοιχο αρχείο πηγαίου κώδικα (δηλαδή το .cpp) ονομάζονται ακολουθώντας την αρχή SceneGraph\_όνομα κλάσης.h/.cpp. Για αυτό, λοιπόν, ονομάσαμε την κλάση μας CameraTracking3D και τα αντίστοιχα αρχεία SceneGraph\_CameraTracking3D.h και SceneGraph\_CameraTracking3D.cpp. Τα αρχεία αποθηκεύτηκαν στους ήδη υπάρχοντες φακέλους του EaZD. Τα αρχεία επικεφαλίδας στον φάκελο include και τα κυρίως αρχεία στο φάκελο src/SceneGraph. Επίσης συμπεριλάβαμε το καινούριο αρχείο επικεφαλίδας στο κυρίως αρχείο SceneGraph.h με την εντολή `#include "SceneGraph_CameraTracking3D.h"`.

Στη συνέχεια, πρέπει να σκεφτούμε ποιανού κόμβου υποκλάση θα είναι ο καινούριος κόμβος που δημιουργούμε. Η επιλογή του κόμβου γονέα χρειάζεται προσοχή αφού η σωστή επιλογή θα “γλιτώσει” τον προγραμματιστή από υλοποιήσεις συναρτήσεων που δεν χρειάζονται και θα βοηθήσει μελλοντικούς προγραμματιστές στην επέκταση του νέου κώδικα. Για παράδειγμα, εάν κάποιος επιθυμεί να υλοποιήσει μία κίνηση σαν της μπάλας του ring-pong καλό είναι ο νέος κόμβος να είναι υποκλάση του κόμβου TransformNode3D, καθώς όλη η διαδικασία γεωμετρικών μετασχηματισμών είναι ήδη υλοποιημένη. Εμείς θέλουμε ο νέος κόμβος να ανοίγει την κάμερα και στη συνέχεια να γίνεται η ανίχνευση μέσω αυτής. Πρόκειται επομένως για έναν γενικό κόμβο που δεν σχετίζεται με κάποιον από τους υπόλοιπους και επομένως θα είναι υποκλάση του κόμβου Node3D (ο οποίος είναι υπερκλάση για όλους τους κόμβους του scene graph). Αφού βρήκαμε τον κόμβο γονέα τον εισάγουμε στον κόμβο παιδί με την εντολή `#include "SceneGraph_Node3D.h"`.

Έπειτα, πρέπει η νέα κλάση να γίνει γνωστή στον μηχανισμό μεταγλώττισης XML (parsing mechanism) συσχετίζοντας την με ένα όνομα το οποίο θα χρησιμοποιείται στα αρχεία σκηνής XML. Η διαδικασία αυτή γίνεται στο αρχείο `SceneGraph_MessageMap.cpp` στην μέθοδο `Node3D::buildNode()`. Το μόνο που χρειάζεται είναι η πρόσθεση μιας νέας εντολής στις ήδη υπάρχουσες `DECLARE_NODE(όνομα XML, κλάση)`. Εμείς χρησιμοποιήσαμε το όνομα “detect” στα αρχεία XML και για αυτό προσθέσαμε την εντολή `DECLARE_NODE(“detect”,CameraTracking3D)`.



Εικόνα 89 : το άνοιγμα του μουσείου και της κάμερας μέσα από το πρόγραμμα EaZD



### 7.3.2 Δήλωση γεγονότων (event declaration)

Σε περίπτωση που ο νέος κόμβος θέλουμε να πυροδοτεί γεγονότα, events στη γλώσσα του EaZD, πρέπει να ακολουθήσουμε τέσσερα βήματα. Στην περίπτωση μας κάθε φορά που γίνεται η ανίχνευση τα στερεά θα σταματούν την περιστροφή τους. Για να γίνει αυτό πρέπει να στέλνεται κάποιο μήνυμα σε συγκεκριμένους κόμβους. Άρα πρέπει μέσα από τον κώδικα της ανίχνευσης να σηματοδοτείται κάποιο γεγονός. Το γεγονός αυτό το ονομάσαμε “detection” και εκτελέσαμε τα απαραίτητα βήματα.

Αρχικά δηλώσαμε το γεγονός της νέας κλάσης στην συνάρτηση SceneGraphRegisterEvents() του αρχείου SceneGraph\_MessageMap.cpp με την γενική εντολή REGISTER\_EVENT(όνομα νέας κλάσης, “όνομα γεγονότος”). Εμείς, δηλαδή, προσθέσαμε την εντολή REGISTER\_EVENT(CameraTracking3D, “detection”). Στη συνέχεια, στο ίδιο αρχείο αρχικοποιήσαμε την λίστα γεγονότων της κλάσης μας με την εντολή INIT\_MESSAGE\_MAP(όνομα κλάσης). Πιο συγκεκριμένα, προσθέσαμε την εντολή INIT\_MESSAGE\_MAP(CameraTracking3D). Έπειτα προσθέσαμε στην κλάση μας (στο αρχείο SceneGraph\_CameraTracking3D.h) δύο συγκεκριμένα προστατευόμενα μέλη (χαρακτηρισμός protected). Αυτά είναι τα

```
static vector eventmap; και
```

```
virtual int getEventID (char *eventname) GET_EVENT_ID(eventname).
```

Τέλος, στο κυρίως αρχείο και την στιγμή που θέλουμε να ενεργοποιηθεί το γεγονός προσθέσαμε την εντολή EVENT\_OCCURED(“detection”).

### 7.3.3 Συγγραφή αρχείου επικεφαλίδας

Στο αρχείο επικεφαλίδας ορίζεται η νέα κλάση. Περιέχει επίσης, όπως άλλωστε ισχύει σε όλα τα αρχεία επικεφαλίδας, τους ορισμούς των μεθόδων καθώς και τις μεταβλητές που χρησιμοποιεί η κλάση. Όλες οι μέθοδοι της λαμβάνουν το χαρακτηριστικό public ενώ όλες οι μεταβλητές χαρακτηρίζονται protected ή private. Φυσικά δεν ξεχάσαμε την δήλωση του κόμβου γονέα με την εντολή #include "SceneGraph\_Node3D.h".

### 7.3.4 Συγγραφή αρχείου πηγαίου κώδικα

Στο αρχείο αυτό υλοποιούνται οι μέθοδοι που ορίστηκαν στο αρχείο επικεφαλίδας.

Για αυτό δεν πρέπει να ξεχάσουμε να αναφέρουμε το αρχείο επικεφαλίδας με την εντολή `#include "SceneGraph.h"`. Στην μέθοδο με όνομα το όνομα της κλάσης αρχικοποιούνται οι μεταβλητές. Η κύρια λειτουργία του νέου κόμβου υλοποιείται στην μέθοδο `app()`. Τέλος, σε κάθε αρχείο `.cpp` πρέπει να υλοποιείται και η μέθοδος `parse (xmlNodePtr pXMLNode)`, για την λειτουργία του μηχανισμού μεταγλώττισης, κατά την οποία δηλώνονται επιπλέον χαρακτηριστικά τα οποία προσδιορίζει ο χρήστης. Στην δική μας εφαρμογή ο χρήστης επιλέγει τι είδους ανίχνευση θέλει να πραγματοποιηθεί. Έτσι στο αρχείο σκηνής, στο αρχείο XML δηλαδή, στην ετικέτα για το άνοιγμα της κάμερας επιλέγει και το μέρος του σώματος που θέλει να ανιχνευτεί. Οι επιλογές είναι τρεις `hand` (χέρι), `face` (πρόσωπο) ή `eye` (μάτι).

### 7.3.5 Συγγραφή αρχείου σκηνής

Περιγράφονται οι κόμβοι και τα χαρακτηριστικά τους που χρησιμοποιήσαμε για το αρχείο σκηνής.

*camera*

<b>Προέλευση</b>	Node3D		
<b>Όνομα κλάσης</b>	Camera3D		
<b>Τοποθέτηση</b>	Οπουδήποτε		
<b>Περιγραφή</b>	Εισάγει την θέαση στον εικονικό κόσμο. Μπορεί να ρυθμιστεί ώστε να ακολουθεί ένα συγκεκριμένο χρήστη ή οποιοδήποτε άλλο κόμβο, για να μπορεί να στέκεται μόνος του ή να μετασχηματιστεί σε έναν κανονικό γεωμετρικό κόμβο σε μια κινηματική αλυσίδα. Μπορούν να οριστούν πολλαπλές κάμερες, αλλά μόνο μία ενεργεί ως πρωταρχική έξοδος.		
<b>Κληρονομούμενα Γνωρίσματα</b>	name, active, visible		
<b>Νέα Γνωρίσματα</b>	<b>γνώρισμα</b>	<b>τύπος</b>	<b>Περιγραφή</b>
	aperture	FLOAT	Το πλήρες άνοιγμα από άκρη σε άκρη των φακών της κάμερας σε μοίρες, σύμφωνα με την κάθετη κατεύθυνση. Το οριζόντιο "άνοιγμα" καθορίζεται από την αναλογία εικόνας του παραθύρου. Προκαθορισμένη τιμή 90.0 (σε μοίρες).
	near	FLOAT	Η απόσταση κοντινού επιπέδου αποκοπής. Πρέπει να είναι θετικός αριθμός. Η προκαθορισμένη τιμή

			είναι 1.0
	far	FLOAT	Η απόσταση μακρινού επιπέδου αποκοπής. Πρέπει να είναι θετικός αριθμός. Η προκαθορισμένη τιμή είναι 1000.0
	focal_distance	FLOAT	Είναι η απόσταση από το κέντρο της κάμερας στο οποίο η εικόνα είναι σε εστίαση, όταν είναι ενεργοποιημένο το φαινόμενο βάθους πεδίου. Χρησιμοποιείται σε συνδυασμό με το γνώρισμα focal_range. Η προκαθορισμένη τιμή είναι 200.0.
	focal_range	FLOAT	Το εύρος, που μετράται από την απόσταση εστίασης κατά την οποία η εικόνα είναι εντελώς ασαφής-θολή. Τα σημεία στην focal_distance θα είναι απότομα, τα σημεία στα focal_distance-focal_range και focal_distance+focal_range και πέρα θα είναι εντελώς ασαφή και τα σημεία εντός του εύρους αυτού θα είναι εν μέρει μη εστιασμένα.
	primary	BOOLEAN	Δηλώνει ότι αυτή η κάμερα είναι ενεργή για να δώσει την έξοδο.
	follow	STRING	Το όνομα του κόμβου στον οποίο αποδίδεται η κάμερα. Συνήθως είναι ο κόμβος χρήστη, αλλά στην πραγματικότητα μπορεί να είναι οτιδήποτε.

*directory*

<b>Προέλευση</b>	-
<b>Όνομα κλάσης</b>	-
<b>Τοποθέτηση</b>	Μέσα στον κόμβο world
<b>Περιγραφή</b>	Προσδιορίζει ένα πρόσθετο μονοπάτι αναζήτησης για κάθε λειτουργία αναζήτησης αρχείου. Αυτό περιλαμβάνει την αναζήτηση όλων των αρχείων που αναφέρονται άμεσα από το γράφημα σκηνής (π.χ. ένα 3D αρχείο που δηλώνεται ως είσοδος σε ένα κόμβο αντικειμένου), καθώς και μέσω έμμεσων αναφορών. Βοηθά στην οργάνωση των σκηνών σε λογικές ομάδες και στο να τις καταστήσει πιο φορητές. Από προεπιλογή, μόνο ο τρέχων κατάλογος περιλαμβάνεται στο μονοπάτι αναζήτησης. Δεν πρόκειται για ένα κανονικό κόμβο γραφήματος σκηνής. Είναι έγκυρος μόνο ως παιδί του κόμβου "world".
<b>Κληρονομούμενα Γνώρισματα</b>	-

Νέα Γνωρίσματα	γνώρισμα	τύπος	Περιγραφή
	path	STRING	Ενεργοποιεί ή απενεργοποιεί τον κόμβο καθώς και όλα τα παιδιά του (αν υπάρχουν). Όταν ένας κόμβος απενεργοποιείται, η επεξεργασία αναστέλλεται, εκτός από το μηχανισμό ανταλλαγής μηνυμάτων. Προεπιλεγμένη τιμή είναι true.

*eventmessage*

<b>Προέλευση</b>	-		
<b>Όνομα κλάσης</b>	-		
<b>Τοποθέτηση</b>	Μέσα σε οποιοδήποτε κόμβο		
<b>Περιγραφή</b>	Δηλώνει τη σύνδεση μεταξύ ενός γεγονότος που παράγεται από τον κόμβο γονέα και ενός μηνύματος που θα έπρεπε να σταλεί σε έναν κόμβο παραλήπτη. Πολλαπλά eventmessages μπορούν να συμπεριληφθούν στο πεδίο δράσης ενός κόμβου, ακόμα και αν ανταποκρίνονται στο ίδιο γεγονός. Αυτό είναι αναγκαίο εάν πρέπει να ειδοποιηθούν πολλοί κόμβοι ταυτόχρονα, όταν συμβαίνει ένα γεγονός. Τα γνωρίσματα event, message και recipient είναι υποχρεωτικά.		
<b>Κληρονομούμενα Γνωρίσματα</b>	-		
Νέα Γνωρίσματα	γνώρισμα	τύπος	Περιγραφή
	event	STRING	Το γεγονός στο οποίο το eventmessage πρέπει να απαντά. Κάθε φορά που το γεγονός λαμβάνει χώρα, το κείμενο που ορίζεται στο γνώρισμα message αποδίδεται στον κόμβο παραλήπτη, μετά από ένα χρονικό διάστημα που καθορίζεται από το γνώρισμα delay. Δεν υπάρχει προκαθορισμένη τιμή.
	recipient	STRING	Το όνομα του γραφήματος σκηνης του παραλήπτη που θα λάβει το μήνυμα.
	message	STRING	Το μήνυμα που αποδίδεται στον παραλήπτη.

*input*

<b>Προέλευση</b>	Node3D
<b>Όνομα κλάσης</b>	Input3D
<b>Τοποθέτηση</b>	Οπουδήποτε
<b>Περιγραφή</b>	Αναπαριστά την είσοδο ελέγχου από μία διεπαφή χρήστη ή

	<p>συσκευή. Παρέχει απαριθμημένους άξονες με πρωτογενείς και κανονικοποιημένες τιμές, καθώς και μια σειρά από κουμπιά που μπορούν να καταγραφούν ανά πάσα στιγμή. Όταν ένας κόμβος χρήστη συνδέεται στον κόμβο εισόδου, τροποποιεί τον εσωτερικό του μετασχηματισμό σύμφωνα με τις τιμές που καταγράφηκαν σε κάθε κύκλο καρτέ από τον κόμβο εισόδου.</p>		
<b>Κληρονομούμενα Γνωρίσματα</b>	Name		
<b>Νέα Γνωρίσματα</b>	<b>γνώρισμα</b>	<b>τύπος</b>	<b>Περιγραφή</b>
	type	OPTION [INTEGER] mousekeyboard joystick	Καθορίζει το είδος της συσκευής που θα χρησιμοποιηθεί σαν είσοδος. Όταν ένα joystick επιλέγεται ως συσκευή εισόδου, ο χρήστης πρέπει να ορίζει ποιο joystick θα χρησιμοποιήσει. Αυτό δίνει τη δυνατότητα απασχόλησης πολλών χρηστών ταυτόχρονα.

*light*

<b>Προέλευση</b>	Node3D		
<b>Όνομα κλάσης</b>	Light3D		
<b>Τοποθέτηση</b>	Οπουδήποτε		
<b>Περιγραφή</b>	<p>Προσθέτει μια πηγή φωτός στην σκηνή. Ο χειρισμός της πηγής φωτός μπορεί να γίνει όπως κάθε άλλου κόμβου, συμπεριλαμβανομένων ενεργοποίησης, απενεργοποίησης και μετασχηματισμών. Οι πηγές φωτός μπορεί να είναι σημεία φωτός, περιοχές φωτός ή spot φωτός με ή χωρίς εξασθένιση ή σκιές. Τα φώτα παρέχουν μια ποικιλία αποτελεσμάτων και μια RGB τιμή χρώματος που μπορεί να τεθεί πάνω από (1,1,1). Εσωτερικά, ο φωτισμός μπορεί να αντέξει σε 200% υπερέκθεση χωρίς να κορεστεί, αν και το τελικό αποτέλεσμα εξαρτάται από την ακρίβεια χρώματος του buffer εξόδου. Σκιές ενεργοποιούνται χωριστά για κάθε πηγή φωτός.</p>		
<b>Κληρονομούμενα Γνωρίσματα</b>	name,active		
<b>Νέα Γνωρίσματα</b>	<b>γνώρισμα</b>	<b>τύπος</b>	<b>Περιγραφή</b>
	color	VEC3	Το χρώμα της φωτεινής πηγής. Οι τιμές RGB μπορεί να είναι μεγαλύτερες από 1.0.
	shadows	BOOLEAN	Είναι true αν το φως ρίχνει σκιές. Προκαθορισμένη τιμή είναι false.
	attenuation	BOOLEAN	True αν το φως είναι εξασθενημένο. Προκαθορισμένη τιμή είναι

			false. Η εξασθένιση είναι γραμμική και το εύρος φωτός καθορίζεται από το γνώρισμα range. Η εξασθένιση δεν έχει ακόμα υλοποιηθεί.
	range	FLOAT	Ο χάρτης σκιών και το εύρος εξασθένισης. Προεπιλεγμένη τιμή είναι 100.0.
	near_range	FLOAT	Η απόσταση κοντινής αποκοπής του χάρτη σκιών. Προεπιλεγμένη τιμή είναι το 100.0.
	position	VEC3	Η θέση της φωτεινής πηγής αναφορικά με το τοπικό σύστημα συντεταγμένων. Αυτό το σημείο υπόκειται σε μετασχηματισμούς.
	aperture	FLOAT	Το άνοιγμα από πλευρά σε πλευρά του κώνου φωτός σε μοίρες, που χρησιμοποιείται από σποτάκια και χάρτες σκιών. Προκαθορισμένη τιμή οι 90.0 μοίρες.
	size	FLOAT	Η διάμετρος της φωτεινής πηγής σε παγκόσμιες μονάδες. Προεπιλεγμένη τιμή είναι το μηδέν.
	skipframes	INT	Καθορίζει πόσα καρέ πρέπει να παρακαμφθούν πριν από τον εκ νέου υπολογισμό του χάρτη σκιών. Κανονικά, οι χάρτες σκιών υπολογίζονται κάθε καρέ, αλλά αν η θέση της σκηνής/φωτός αλλάζει αργά ίσως είναι επιθυμητή η βελτίωση του ρυθμού καρέ με την αποφυγή της επανασχεδίασης των χαρτών σκιάς σε κάθε καρέ. Προεπιλεγμένη τιμή είναι το 0.
	target	VEC3	Το σημείο στο οποίο κοιτά η φωτεινή πηγή, σε σχέση με το τοπικό σύστημα συντεταγμένων. Χρησιμοποιείται για τον τύπο πηγών φωτός με σποτάκια και για τον υπολογισμό του frustum του χάρτη σκιών. Το σημείο αυτό υπόκειται σε μετασχηματισμούς.

*transformation*

<b>Προέλευση</b>	Group3D		
<b>Όνομα κλάσης</b>	TransformNode3D		
<b>Τοποθέτηση</b>	Οπουδήποτε		
<b>Περιγραφή</b>	Ο κόμβος αυτός μετατρέπει γεωμετρικά όλους τους φωλιασμένους κόμβους, σύμφωνα με τις συγκεκριμένες παραμέτρους μετατόπισης, περιστροφής και κλιμάκωσης. Ως αποτέλεσμα, οι φωλιασμένοι μετασχηματισμοί έχουν ένα συσσωρευτικό αποτέλεσμα. Η σειρά των μετασχηματισμών είναι σταθερή: κλιμάκωση, περιστροφή και μετατόπιση. Αν χρειαστεί να αλλάξει αυτή η σειρά πρέπει να εφαρμοστούν δύο φωλιασμένοι μετασχηματισμοί. Ο εσωτερικός πίνακας επαναπροσδιορίζεται μόνο όταν οι τιμές έχουν αλλάξει και όχι σε κάθε καρέ.		
<b>Κληρονομούμενα Γνωρίσματα</b>	name, active, visible		
<b>Νέα Γνωρίσματα</b>	<b>γνώρισμα</b>	<b>τύπος</b>	<b>Περιγραφή</b>
	translation	VEC3	Το διάνυσμα μετατόπισης σε παγκόσμιες μονάδες.
	scale	VEC3	Ο παράγοντας κλιμάκωσης κατά μήκος καθενός από τους τρεις τοπικούς άξονες.
	rotation	VEC3	Η περιστροφή σύμφωνα με το στυλ της OpenGL. Είναι σε μορφή (γωνία, άξονας x, άξονας y, άξονας z) όπου η γωνία είναι η δεξιόστροφη γωνία περιστροφής σε μοίρες και άξονας, ο άξονας περιστροφής.

*spinner*

<b>Προέλευση</b>	TransformNode3D		
<b>Όνομα κλάσης</b>	Spinner3D		
<b>Τοποθέτηση</b>	Οπουδήποτε		
<b>Περιγραφή</b>	Αυτός ο ειδικός μετασχηματισμός αναπαριστά χρονικά μεταβλητή περιστροφή γύρω από άξονα που καθορίζει ο χρήστης.		
<b>Κληρονομούμενα Γνωρίσματα</b>	name, active, visible		
<b>Νέα Γνωρίσματα</b>	<b>γνώρισμα</b>	<b>τύπος</b>	<b>Περιγραφή</b>
	axis	VEC3	Ο άξονας της κατεύθυνσης περιστροφής. ( η περιστροφή είναι ως προς το (0,0,0) στο τοπικό σύστημα συντεταγμένων)

	period	FLOAT	Η περίοδος περιστροφής
	repeats	FLOAT	Ο αριθμός των ολοκληρωμένων περιστροφών. Είναι μηδέν για άπειρη περιστροφή.

*object*

<b>Προέλευση</b>	Node3D		
<b>Όνομα κλάσης</b>	Geometry3D		
<b>Τοποθέτηση</b>	Οπουδήποτε		
<b>Περιγραφή</b>	Ο κόμβος αντικειμένου γραφήματος σκηνής φορτώνει και εμφανίζει ένα πλέγμα από ένα εξωτερικό αρχείο σε μία από τις υποστηριζόμενες μορφές. Επί του παρόντος, υποστηρίζεται μόνο η μορφή Alias Wavefront Object (Το .obj και το αντίστοιχο. mtl). Όλες οι γνωστές μορφές bitmap υποστηρίζονται για απόδοση υφής.		
<b>Κληρονομούμενα Γνωρίσματα</b>	name, active, visible		
<b>Νέα Γνωρίσματα</b>	<b>γνώρισμα</b>	<b>τύπος</b>	<b>Περιγραφή</b>
	file	STRING	Το αρχείο που θα φορτωθεί. Το μονοπάτι αναζήτησης για το αρχείο παρέχεται από τους κόμβους directory του γραφήματος σκηνής.

*skylight*

<b>Προέλευση</b>	-		
<b>Όνομα κλάσης</b>	-		
<b>Τοποθέτηση</b>	Μέσα στον κόμβο world. Μόνο ένας τέτοιος κόμβος μπορεί να είναι παρών.		
<b>Περιγραφή</b>	Δημιουργεί ένα ουρανό και έναν ήλιο. Ο ουρανός δεν αντιπροσωπεύεται γεωμετρικά. Καθώς οι φυσικές λύσεις τείνουν να είναι πιο "έντονες" όσον αφορά τους υπολογισμούς και το μοντέλο του EaZD είναι εμπειρικό, δεν απαιτεί καμία υφή και παραμετροποιείται μόνο από την πυκνότητα του αέρα (haze). Ο ήλιος συνδέεται με μία από τις υπάρχουσες πηγές φωτός της σκηνής, όπως ορίζονται από το χρήστη. Ο φωτισμός του ουρανού προσαρμόζεται δυναμικά ανάλογα με την θέση του ήλιου και επομένως τη θέση της αντίστοιχης πηγής φωτός. Επιπλέον, ο κόμβος αυτός παρέχει μια επιλογή για την ενεργοποίηση ογκομετρικού φωτός από τον ήλιο για ρεαλιστικές ακτίνες (godrays).		
<b>Κληρονομούμενα Γνωρίσματα</b>			
<b>Νέα Γνωρίσματα</b>	<b>γνώρισμα</b>	<b>τύπος</b>	<b>Περιγραφή</b>
	active	BOOLEAN	Ενεργοποιεί τη χρήση



			skylight.
	air_density	FLOAT	Κανονικοποιημένη πυκνότητα αέρος σε εύρος [0,1] που προκαλεί μία μικρή αύξηση στο haze του ουρανού. Επηρεάζει επίσης την ένταση των ακτίνων godray.
	volumetric_light	BOOLEAN	Ενεργοποιεί τη χρήση ογκομετρικού φωτός ήλιου. Η ποιότητα καθορίζεται από τον αριθμό των δειγμάτων που λαμβάνονται υπόψη.
	samples	INTEGER	Ο αριθμός των δειγμάτων ανά ακτίνα. Δρα μόνο όταν το γνώρισμα volumetric_light τίθεται σε true.
	sun	STRING	Παρέχει το όνομα του ενός κόμβου φωτός για να χρησιμοποιηθεί ως ήλιος. Γνωρίσματα σκιάς και τοποθέτησης της φωτεινής πηγής είναι σεβαστά.

*user*

<b>Προέλευση</b>	TransformNode3D
<b>Όνομα κλάσης</b>	User3D
<b>Τοποθέτηση</b>	Οπουδήποτε
<b>Περιγραφή</b>	Αυτό το είδος κόμβου δημιουργεί μια αυτόνομη οντότητα που μπορεί να περιηγηθεί μέσα στον εικονικό κόσμο, ως αποτέλεσμα ενός ελεγκτή που παρέχει στοιχεία εισόδου και ως εκ τούτου οδηγεί τον κόμβο. Ο κόμβος user δεν είναι μοναδικός. Πολλοί χρήστες μπορεί να κατοικούν στον εικονικό κόσμο, με τον καθένα να αντιπροσωπεύει ένα διαφορετικό πραγματικό χρήστη, χαρακτήρες που ελέγχονται από υπολογιστή ή ζώα κλπ. Οι κόμβοι user δεν εισάγουν σύστημα θέασης, καθώς αυτό είναι μέρος του κόμβου camera, ο οποίος μπορεί να συνδεθεί με οποιοδήποτε κόμβο γραφήματος σκηνής, συμπεριλαμβανομένου του user. Οι κάμερες και οι χρήστες είναι εντελώς διαφορετικές οντότητες στο SceneGraph3D, επιτρέποντας μια μεγάλη ποικιλία από συνθέσεις, πολλαπλές γωνίες και κάθε δυναμικό συνδυασμό αυτών. Ο κόμβος user προέρχεται από τον τύπο κόμβου μετασχηματισμού, γεγονός που σημαίνει ότι μπορεί να περιέχει άλλους κόμβους. Αυτό κάνει την προσθήκη της οπτικής αναπαράστασης ενός avatar πολύ εύκολη, καθώς προστίθεται απλά οποιοσδήποτε κόμβος σαν περιεχόμενο του user αλλά θα μπορεί να κινείται μαζί του.
<b>Κληρονομούμενα</b>	name, active, visible

Γνωρίσματα			
Νέα Γνωρίσματα	γνώρισμα	τύπος	Περιγραφή
	input	STRING	Δηλώνει τον κόμβο εισόδου που θα παρέχει τον έλεγχο για τον χρήστη.
	position	VEC3	Η αρχική θέση του χρήστη. Προκαθορισμένη τιμή το (0,0,0).
	lookat	VEC3	Που κατευθύνεται αρχικά ο χρήστης. Αρχική τιμή το (0,0,-100).
	turn	FLOAT	Ο ρυθμός με τον οποίο στρέφεται ο χρήστης. Προκαθορισμένη τιμή το 1.0, το οποίο είναι αρκετά γρήγορο.
	speed	FLOAT	Η γραμμική ταχύτητα του χρήστη σε παγκόσμιες μονάδες. Η προεπιλεγμένη τιμή είναι 1.0.
	control	OPTION <i>navigate</i> <i>orbit</i> <i>passive</i>	Διαμορφώνει τον τύπο της κίνησης που εκτελεί ο χρήστης σε σχέση με το look-at σημείο και τις παραμέτρους εισόδου. Αν έχει οριστεί η επιλογή <i>navigate</i> (προεπιλογή), ο χρήστης περιστρέφεται χρησιμοποιώντας τη θέση του ως σημείο άξονα. Αν οριστεί η λειτουργία <i>orbit</i> , ο χρήστης στρέφεται γύρω από το σημείο look at. Αν οριστεί η επιλογή <i>passive</i> , ο χρήστης αγνοεί κάθε είσοδο.

*world*

Προέλευση	Group3D		
Όνομα κλάσης	World3D		
Τοποθέτηση	Μόνο ένα μπορεί να υπάρχει σε κάθε αρχείο σκηνής.		
Περιγραφή	Ο ανώτερος κόμβος ενός XML αρχείου σκηνής. Πρέπει να είναι μοναδικός καθώς αναπαριστά τη ρίζα όλου του γραφήματος σκηνής. Παρέχει καθολικά γνωρίσματα και λειτουργίες και περιέχει κάθε άλλο κόμβο. Ο κόμβος world έχει προκαθορισμένο όνομα: %world.		
Κληρονομούμενα Γνωρίσματα	visible, active		
Νέα Γνωρίσματα	γνώρισμα	τύπος	Περιγραφή
	ambient_occlusion	OPTION	Ορισμός της μεθόδου για

		<i>no</i> <i>none</i> <i>ssao</i>	τον υπολογισμό του ατμοσφαιρικού φωτισμού της σκηνής. Επί του παρόντος, δύο επιλογές είναι διαθέσιμες, no και none, οι οποίες χρησιμοποιούν το αδιαμόρφωτο παγκόσμιο ατμοσφαιρικό στοιχείο και η ssao, η οποία υπολογίζει τη σύγκλιση του περιβάλλοντος του χώρου της οθόνης. Μπορεί να προκληθεί σημαντική επιβράδυνση κατά την ενεργοποίηση αυτού του τύπου φωτισμού. Προεπιλεγμένη τιμή είναι none/no.
	ambient	VEC3	Το καθολικό ατμοσφαιρικό χρώμα. Η προεπιλεγμένη τιμή είναι το (0.2,0.2,0.2).
	background	VEC3	Το χρώμα φόντου.
	ao_radius	FLOAT	Το εύρος επίδρασης της σύγκλισης περιβάλλοντος. Ο αλγόριθμος ψάχνει για σύγκλιση μέχρι ao_radius παγκόσμιες μονάδες. Προεπιλεγμένη τιμή είναι το 1.
	bloom	BOOLEAN	Ενεργοποιεί ή απενεργοποιεί το φαινόμενο bloom. Προεπιλεγμένη τιμή είναι off.
	defocus	BOOLEAN	Ενεργοποιεί και απενεργοποιεί το φαινόμενο μη εστίασης στο βάθος πεδίου. Οι παράμετροι που συνδέονται με το φαινόμενο είναι γνωρίσματα του κόμβου κάμερας, καθώς διαφορετικές κάμερες μπορεί να έχουν διαφορετικές προδιαγραφές φακών. Προεπιλογή είναι το off.

	shadows	OPTION <i>simple</i> <i>quad</i> <i>disk</i> <i>slope</i>	Καθορίζει τον τρόπο για τη δειγματοληψία χάρτη σκιών. Τρεις διαφορετικοί τρόποι υποστηρίζονται μέχρι στιγμής, του πλησιέστερου γείτονα, 2X2 πολύ-δειγματοληψία με την τεχνική percentage closer filtering (PCF), 16-tap disk sampling PCF και 16-tap PCF with variable bias. Οι αντίστοιχες επιλογές είναι οι εξής: <i>simple</i> , <i>quad</i> , <i>disk</i> , <i>slope</i> . Ο τρόπος δειγματοληψίας είναι ίδιος για όλους τους χάρτες σκιών. Οι απαλές σκιές παράγονται με όλους τους τρόπους, εκτός από τον πλησιέστερο γείτονα. Το μέγεθος της φωτεινής πηγής καθορίζει την ασάφεια του ημίφωτος. Προκαθορισμένη τιμή η <i>simple</i> . Ο έλεγχος σκιών χρησιμοποιείται επίσης για εσωτερικούς υπολογισμούς όπως στα ογκομετρικά φώτα.
--	---------	---	---

*group*

<b>Προέλευση</b>	Node3D
<b>Όνομα κλάσης</b>	Group3D
<b>Τοποθέτηση</b>	Οπουδήποτε
<b>Περιγραφή</b>	Είναι ο βασικός κόμβος ομαδοποίησης άλλων κόμβων. Οι κόμβοι που περιέχονται τοποθετούνται μεταξύ των ετικετών αρχής/τέλους της ομάδας χρησιμοποιώντας το προκαθορισμένο φωλιασμένο στυλ XML.
<b>Κληρονομούμενα Γνωρίσματα</b>	name, active, visible

*node*

<b>Προέλευση</b>	-
<b>Όνομα κλάσης</b>	Node3D
<b>Τοποθέτηση</b>	Πουθενά
<b>Περιγραφή</b>	Είναι ο γενικός κόμβος, του οποίου κάθε άλλος κόμβος είναι υποκατηγορία. Δεν έχει καμία άλλη λειτουργικότητα στη γλώσσα

	περιγραφής σκηνής, εκτός από τον καθορισμό του ελάχιστου συνόλου των μηνυμάτων που δέχονται οι κόμβοι και κατά συνέπεια δεν συμπεριλαμβάνεται στο αρχείο XML.		
<b>Κληρονομούμενα Γνωρίσματα</b>	-		
<b>Νέα Γνωρίσματα</b>	<b>γνώρισμα</b>	<b>τύπος</b>	<b>Περιγραφή</b>
	active	BOOLEAN	Ενεργοποιεί ή απενεργοποιεί τους κόμβους και τα παιδιά του. Όταν ένας κόμβος απενεργοποιείται αναστέλλεται όλη η επεξεργασία, εκτός από το μηχανισμό μετάδοσης μηνυμάτων. Στην περίπτωση του ειδικού κόμβου world, συνεχίζεται η διατήρηση του χρόνου ακόμα και αν ο κόμβος έχει ανασταλεί. Προεπιλεγμένη τιμή είναι true.
	visible	BOOLEAN	Ορίζει την κατάσταση εμφάνισης του κόμβου και των παιδιών του. Αν ένας κόμβος δεν είναι ορατός, δεν είναι παρών σε κανένα υπολογισμό σχεδίασης. Προεπιλεγμένη τιμή είναι true.
	debug	BOOLEAN	Αν true, ορίζει την κατάσταση σχεδιασμού μιας οντότητας σε μορφή debug, επιτρέποντας σε επιπλέον στοιχεία να σχεδιαστούν.
	name	STRING	Το όνομα του κόμβου. Αυτό το γνώρισμα είναι προαιρετικό αλλά όταν ο κόμβος πρέπει να αναφερθεί από έναν άλλο κόμβο, η αναφορά γίνεται μέσω ονόματος. Επομένως απαιτείται ένα μοναδικό όνομα να αποδοθεί σε κάθε αναφερόμενο κόμβο. Προεπιλεγμένο όνομα για όλους τους κανονικούς κόμβους είναι "none".

[40]

## Κεφάλαιο 8

### Εφαρμογές και συσκευές που σχετίζονται με ανίχνευση

Τα ηλεκτρονικά παιχνίδια διασκεδάζουν το κοινό τους εδώ και σχεδόν 30 χρόνια. Τα γραφικά ηλεκτρονικών υπολογιστών έχουν εξελιχθεί αρκετά από τότε και τα γραφικά ηλεκτρονικών παιχνιδιών αλλάζουν συνεχώς τα όρια μεταξύ εικόνας και πραγματικότητας. Στις μέρες μας, οι ερευνητές και οι μηχανικοί επεκτείνουν τον κόσμο των μηχανών και τον εισάγουν σιγά σιγά στον δικό μας κόσμο. Η νέα τεχνολογία, λοιπόν, θολώνει την διαφορά μεταξύ του πραγματικού και φανταστικού βελτιώνοντας την αίσθηση του τι βλέπουμε, ακούμε, νιώθουμε και μυρίζουμε.

Η εικονική πραγματικότητα πλησιάζει ολοένα και περισσότερο στην πραγματικότητα και με την τεχνολογία της επαυξημένης πραγματικότητας δεν ενσωματώνεται κάποιος σε νέο περιβάλλον ούτε εγκαταλείπει το ήδη οικείο περιβάλλον απλά το διανθίζει. Μια live camera, δηλαδή, δίνει την εικόνα του πραγματικού κόσμου μπροστά στη συσκευή, και σε αυτή την εικόνα η συσκευή προσθέτει ανά πάσα στιγμή “στοιχεία” που αλλάζουν τον εξωτερικό κόσμο κατά βούληση. Όλοι μας (από τουρίστες, που ψάχνουν δρόμους και πληροφορίες, μέχρι στρατιώτες, που προσπαθούν να εκπαιδευτούν όσο καλύτερα γίνεται) μπορούμε να εκμεταλλευτούμε την δυνατότητα εισαγωγής χαρακτήρων και στοιχείων, που δημιουργούνται από μηχανές, στο οπτικό μας πεδίο.

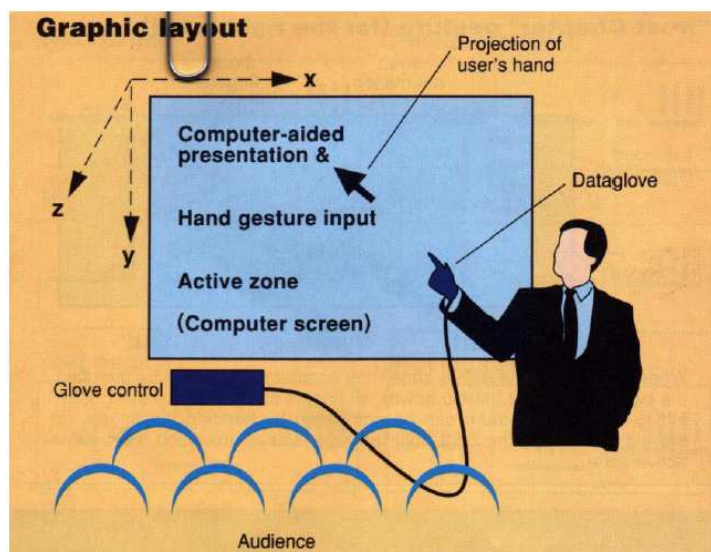
Ο συνδυασμός ανίχνευσης κίνησης και τρισδιάστατης απεικόνισης μάς φέρνει αρκετά βήματα πιο κοντά στο τελικό σύνορο της επαυξημένης πραγματικότητας. Μερικοί πιστεύουν ότι ο στόχος αυτός θα επιτευχθεί σταδιακά, με τη συνεχή ανάπτυξη των περιφερειακών. Άλλοι πιστεύουν ότι κάποτε θα φτάσουμε στο σημείο ώστε το ίδιο μας το σώμα θα παίζει τον ρόλο της εισόδου σε κάποια μηχανή και ότι θα μπορούμε να “ζούμε” μέσα στον εγκέφαλό μας. Βασική προϋπόθεση είναι η μεταβίβαση των αισθήσεων του χρήστη. Μία πρώτη απόπειρα προς αυτή την κατεύθυνση έγινε με τη γνωστή σε όλους μας ανάδραση (force feedback) των χειριστηρίων. Η αλήθεια είναι ότι με την υπάρχουσα τεχνολογία είναι πολύ δύσκολο να δούμε κάτι πιο “εκλεπτυσμένο” σε αυτόν τον τομέα στο άμεσο μέλλον. Υπάρχουν όμως περιπτώσεις που η λογική αυτή εφαρμόζεται για σοβαρούς σκοπούς. Το Γενικό Επιτελείο Στρατού των ΗΠΑ σε συνεργασία με το Πανεπιστήμιο της Νότιας Καλιφόρνια και την εταιρεία Virtually Better Inc. σχεδίασαν το πρόγραμμα εικονικής πραγματικότητας “Virtual Iraq” [41]. Σκοπός του προγράμματος είναι η αντιμετώπιση της διαταραχής μετατραυματικού

στρες από την οποία υποφέρουν εκατοντάδες στρατιώτες που έχουν υπηρετήσει στο Ιράκ και στο Αφγανιστάν. Με βάση τις περιγραφές του ασθενή, ο θεραπευτής αναπαράγει τις συνθήκες που βίωσε ο πρώτος στο πεδίο της μάχης. Όσο πιο ακριβής είναι η αναβίωση του τραυματικού γεγονότος, τόσο καλύτερα είναι τα αποτελέσματα της θεραπείας. Για αυτό το λόγο ο ασθενής φοράει κιάλια εικονικής πραγματικότητας και πατάει σε ειδικό χαλί ανάδρασης. Επιπλέον, ο θεραπευτής μπορεί να αναπαράγει ανά πάσα στιγμή οσμές του πεδίου μάχης, όπως καμένο λάστιχο, καύσιμα, ιδρώτα κλπ, ενισχύοντας την αληθοφάνεια της εμπειρίας.

Στη συνέχεια παρουσιάζονται κάποιες εφαρμογές που κυκλοφορούν στο εμπόριο και κάνουν πράξη πολλά από αυτά που περιγράφηκαν.

## 8.1 Charade

Ένα εξαιρετικό παράδειγμα διεπαφής που βασίζεται αποκλειστικά σε ανίχνευση κινήσεων του χρήστη είναι το Charade (Baudel & Beaudouin-Lafon, 1993). Χρησιμοποιεί τις δυναμικές χειρονομίες και κινήσεις του ανθρώπινου χεριού για τον έλεγχο μιας παρουσίασης μέσω υπολογιστή. Ο χρήστης στέκεται μπροστά από μια οθόνη προβολής φορώντας ένα ειδικό γάντι το οποίο συνδέεται με έναν υπολογιστή. Όταν κινεί το χέρι του προς την οθόνη προβολής μπορεί να θεωρηθεί ότι δίνει εντολές ελέγχου της παρουσίασης που εμφανίζεται στην οθόνη, όπως η μετάβαση στην επόμενη διαφάνεια. Υπάρχουν συνολικά 16 εντολές που μπορούν να αναγνωριστούν. Το Charade αναγνωρίζει σωστά το 70-80% των χειρονομιών ενός νέου χρήστη και έχει ακρίβεια 90-98% για τους εκπαιδευμένους χρήστες.



Εικόνα 90 : Η γενική ιδέα του συστήματος Charade

Για να εξασφαλιστεί η ακριβής αναγνώριση χειρονομιών και μια διαισθητική διεπαφή εφαρμόστηκε μια σειρά από περιορισμούς. Μια περιοχή μπροστά από την οθόνη προβολής ορίζεται ως η ενεργή ζώνη και χειρονομίες αγνοούνται αν πραγματοποιηθούν έξω από αυτή την περιοχή. Αυτό επιτρέπει στους χρήστες να αναμειγνύουν εντολές χειρονομιών με κανονικές ομιλητικές χειρονομίες. Οι χειρονομίες ορίζονται επίσης από ένα σύνολο θέσεων εκκίνησης και τέλους καθώς και δυναμικής κίνησης μεταξύ των θέσεων αρχής και τέλους. [11]

## 8.2 Nintendo Wii

Wii είναι η ονομασία της νέας κονσόλας οικιακής ψυχαγωγίας που δημιούργησε η εταιρία Nintendo και κυκλοφόρησε το 2006. Πρόκειται για την κονσόλα 7ης γενιάς της εταιρίας, της ίδιας δηλαδή γενιάς με το PlayStation 3 (της εταιρίας Sony) και το Xbox 360 (της εταιρίας Microsoft). Αν και υστερεί από άποψη γραφικών και άλλων τεχνικών χαρακτηριστικών σε σχέση με τους ανταγωνιστές της, PlayStation 3 και Xbox 360, ο επαναστατικός χειρισμός του, κάνει το Wii να αντιπροσωπεύει ένα καινοτόμο βήμα προς μια νέα εποχή ψυχαγωγίας. Όπως έχει δηλώσει η εταιρία, η κονσόλα δεν βασίζεται στους πιο γρήγορους επεξεργαστές ή τις υπερδύναμες κάρτες γραφικών, στις οποίες βασίζονται οι δύο άλλες κονσόλες, αλλά στο πραγματικά πρωτοποριακό χειρισμό του, ο οποίος γίνεται μέσω δύο τμημάτων: το Wii Remote, το οποίο μοιάζει με τηλεχειριστήριο και το Nunchuk, το οποίο είναι ένα κλασικό joystick σε μικρότερο μέγεθος. Τα δύο αυτά τμήματα συνδέονται με ένα καλώδιο μεταξύ τους. Το Remote ενσωματώνει αισθητήρες κίνησης, οι οποίοι ανιχνεύουν πλήρως τις κινήσεις του παίκτη στον τρισδιάστατο χώρο στον οποίο κινείται και τις μεταφέρουν στο παιχνίδι. Για παράδειγμα σε ένα παιχνίδι τένις μπορεί το Remote να χρησιμοποιηθεί σαν ρακέτα, σε ένα παιχνίδι ξιφομαχίας σαν σπαθί κτλ. Αντίστοιχα, και το Nunchuk αντιλαμβάνεται κίνηση στο χώρο και είναι απαραίτητο συμπλήρωμα για τα περισσότερα παιχνίδια.





Εικόνα 91 : Η κονσόλα Wii με το χειριστήριο Wii remote

Αρχικά η νέα κονσόλα ονομάστηκε Revolution. Τον Απρίλιο του 2006, η Nintendo έκρινε πως το όνομα της νέας της κονσόλας θα έπρεπε να είναι μια λέξη που να μπορεί να ειπωθεί εύκολα από ανθρώπους οπουδήποτε στον κόσμο. Ταυτόχρονα τα δύο "i" συμβολίζουν τόσο το ιδιαίτερο χειριστήριο της κονσόλας, όσο (σύμφωνα με την Nintendo) και δύο ανθρώπους δίπλα-δίπλα, θέλοντας έτσι να τονίσει τη διασκεδαστική συνύπαρξη πολλών παικτών και να τους προσκαλέσει να παίξουν όλοι μαζί. Σε αυτό παραπέμπει και η ηχητική του ονόματος μιας και όταν διαβάζεται ακούγεται σαν τη λέξη "we", δηλαδή "εμείς".

Στόχος της εταιρίας ήταν να προσελκύσει όλων των ειδών παίκτες, από τους πιο φανατικούς μέχρι τους απλούς περιστασιακούς, ακόμα και τους πιο αδιάφορους. Εστιάζει σε καινούριες εμπειρίες παιχνιδιού με μεγάλες συγκινήσεις και ακόμα μεγαλύτερη εμπλοκή του παίκτη στην εξέλιξη, αφού στο Wii, βασικό ρόλο παίζει η σωστή αλληλεπίδραση του χεριού του χρήστη με το παιχνίδι.

Μέχρι και 4 χειριστήρια μπορούν να συνδεθούν ταυτόχρονα στην κονσόλα Wii χρησιμοποιώντας την ασύρματη τεχνολογία Bluetooth, οπότε μέχρι και τέσσερις παίκτες μπορούν να παίξουν με το δικό τους Wii Remote έχοντας την δυνατότητα ενός Wii-Party.

Το Wii διαθέτει μια πολύ κομψή εμφάνιση με μικρό μέγεθος, περίπου όσο τρεις θήκες DVD τοποθετημένες η μία πάνω στην άλλη. Μπορεί να αποτελέσει φυσική προσθήκη σε οποιαδήποτε τηλεόραση και να τοποθετηθεί είτε σε οριζόντια θέση, είτε σε κάθετη.

Η έναρξη της κυκλοφορίας της κονσόλας έγινε στις 19 Νοεμβρίου 2006 στην Αμερική, στις 2 Δεκεμβρίου στην Ιαπωνία, στις 7 Δεκεμβρίου στην Αυστραλία και στις 8

Δεκεμβρίου στην Ευρώπη. Η κυκλοφορία συνοδεύτηκε από εορταστικά δρώμενα σε όλες τις χώρες. [26 (όρος Wii)]

## 8.3 Eye-Toy

### 8.3.1 Εισαγωγή

Το EyeToy, είναι μια έγχρωμη ψηφιακή συσκευή κάμερας, για το PlayStation 2. Χρησιμοποιεί την τεχνολογία της “τεχνητής όρασης” και αναγνώρισης χειρονομιών για να επεξεργαστεί εικόνες που λαμβάνονται από την κάμερα. Αυτό επιτρέπει στους παίκτες να αλληλεπιδρούν με τα παιχνίδια χρησιμοποιώντας την κίνηση, την αναγνώριση χρώματος και ήχου, μέσω του ενσωματωμένου μικροφώνου. Το EyeToy της Sony αποτελεί ένα μοναδικό περιφερειακό που επιτρέπει στους παίκτες να γίνουν μέρος του παιχνιδιού. Η κάμερα έχει κατασκευαστεί από την Logitech, αν και οι νεότερες EyeToys κατασκευάζονται από την Namtai. Η κάμερα χρησιμοποιείται κυρίως για να παίζονται παιχνίδια EyeToy που αναπτύχθηκαν από την Sony και άλλες εταιρείες. Κύρια λειτουργία της είναι να προβάλλει την εικόνα του παίκτη στην οθόνη. Χρησιμοποιώντας τις κινήσεις του σώματος τους, οι παίκτες μπορούν να αλληλεπιδρούν με τα αντικείμενα στο παιχνίδι. Δεν απαιτείται κανένας ελεγκτής ακόμα και για την επιλογή του μενού. Το γεγονός αυτό καθιστά το EyeToy προσιτό για παίκτες όλων των ικανοτήτων. Δεν προορίζεται για χρήση ως μια κανονική κάμερα υπολογιστή, αν και έχουν αναπτυχθεί ανεπίσημοι οδηγοί για το σκοπό αυτό. Μέχρι την 6η Νοεμβρίου 2008, το EyeToy είχε πουλήσει 10.500.000 μονάδες παγκοσμίως, αποτελώντας μια εκπληκτική συσκευή διασκέδασης. [26, (όρος Eye Toy)]



Εικόνα 92 : Η συσκευή EyeToy

### 8.3.2 Ιστορία

Η ιδέα του EyeToy επινοήθηκε από τον Richard Marks το 1999. Η ιδέα ήταν να καταστεί δυνατή μια φυσική διεπαφή χρήστη και εικονικές εφαρμογές βιντεοπαιχνιδιών χρησιμοποιώντας μια απλή κάμερα και την υπολογιστική δύναμη του PlayStation 2 για την εφαρμογή τεχνολογιών υπολογιστικής όρασης και αναγνώρισης χειρονομιών. Οπότε εντάχθηκε στην Sony Computer Entertainment America (SCEA) δουλεύοντας για το σκοπό αυτό. Η αρχική ιδέα κατέληξε σε μεγάλη επιτυχία και το υλικό κυκλοφόρησε στη Ευρώπη τον Ιούλιο του 2003 μαζί με το πρώτο συμβατό παιχνίδι, EyeToy:Play. [26, (όρος Eye Toy)]

### 8.3.3 Σχεδιασμός

Η ίδια η συσκευή EyeToy είναι μια μικρή κάμερα που ρυθμίζεται και τοποθετείται δίπλα στην τηλεόραση ή την οθόνη. Οι παίκτες βλέπουν τους εαυτούς τους να προβάλλονται στην οθόνη. Αντί για την τυπική διεπαφή ελεγκτή, η αλληλεπίδραση μπορεί να γίνει χρησιμοποιώντας κινήσεις του ανθρώπινου σώματος όπως μετακίνηση ώμου ή χεριών. Είναι ένα κομμάτι της τεχνολογίας που αυξάνει τη διαδραστικότητα στα παιχνίδια. Η ρύθμιση της κάμερας είναι απλή. Το EyeToy συνδέεται με την USB θύρα του PS2 και ύστερα τοποθετείται πάνω ή κάτω από την τηλεόραση. Για τη σωστή λειτουργία της, ο παίκτης πρέπει να βρίσκεται αρκετά μακριά από την κάμερα. Διαφορετικά η εικόνα του παίκτη θα πιάσει πολύ χώρο στην οθόνη. Θα πρέπει να ελεγχθεί ο φωτισμός στο δωμάτιο να είναι αρκετά καλός, διαφορετικά η συσκευή θα δυσκολεύεται στην αναγνώριση της εικόνας του παίκτη. Μόλις έχει ρυθμιστεί η σωστή απόσταση και το ύψος, μπορεί να ρυθμιστεί η εστίαση της κάμερας γυρίζοντας το δακτύλιο μέχρι να προκύψει μια ευκρινής εικόνα. Αφού ολοκληρωθεί η βασική εγκατάσταση, η διεπαφή θα είναι εξαιρετικά εύκολο να χρησιμοποιηθεί. Η περιήγηση στο μενού γίνεται με την κίνηση των χεριών πάνω στα εικονίδια με γρήγορο ρυθμό. Για την έξοδο από το παιχνίδι ή για την επιστροφή στο κυρίως μενού μπορεί να τοποθετηθεί το χέρι του παίκτη σε ολόκληρη την κάμερα. Η διεπαφή είναι εξαιρετικά διαισθητική. Διαθέτει επίσης δύο LEDs στο μπροστινό μέρος. Ένα μπλε φως ανάβει όταν το PS2 είναι ενεργοποιημένο, που δείχνει ότι είναι έτοιμο να χρησιμοποιηθεί, ενώ το κόκκινο φως αναβοσβήνει όταν υπάρχει ανεπαρκές φως στο δωμάτιο. Υπάρχει επίσης ένα

μικρόφωνο ενσωματωμένο, για χρήσεις όπως η καταγραφή ενός σύντομου μηνύματος βίντεο για κάποιον. [34] [26, (όρος Eye Toy)]

### 8.3.4 Τεχνικοί περιορισμοί

Λόγω της ανάγκης της κάμερας να "παρακολουθεί" τον παίκτη καθώς παίζει, η κάμερα μπορεί να είναι πολύ σχολαστική σχετικά με το φως στο δωμάτιο. Διαφορετικά παιχνίδια έχουν διαφορετική ανοχή για διάφορες συνθήκες φωτισμού. [26, (όρος Eye Toy)]

### 8.3.5 Επιπλέον δυνατότητες

Επιπλέον των mini-games, οι παίκτες μπορούν να κάνουν και άλλα πράγματα με το EyeToy. Υπάρχει μια ειδική λειτουργία, όπου ο παίκτης μετακινείται στο χώρο ενώ ειδικά εφέ βελτιώνουν την εικόνα. Το άλλο σημαντικό χαρακτηριστικό είναι η βιντεοσκόπηση μηνυμάτων. Καταγράφεται η φωνή του παίκτη μαζί με το βίντεο. Χρησιμοποιώντας αυτό το χαρακτηριστικό, ένα σύντομο μήνυμα 10 έως 60 δευτερολέπτων, μπορεί να αποθηκευτεί στην κάρτα μνήμης. Μόλις ολοκληρωθεί η καταγραφή του μηνύματος, μπορεί να το μοιραστεί με κάποιον άλλο. Αυτά απαιτούν πολύ χώρο στην κάρτα μνήμης, περιορίζοντας τον αριθμό των μηνυμάτων που μπορεί να αποθηκευτούν, αλλά είναι μια εξαιρετική ιδέα και καλή χρήση της κάμερας. Δυστυχώς, υπάρχουν μερικά μικρά προβλήματα που μειώνουν ελαφρώς τη δύναμη της εντυπωσιακής τεχνολογίας του Eye Toy. Η ποιότητα της εικόνας δεν είναι η μέγιστη δυνατή. Ωστόσο αυτό δεν είναι πραγματικά μεγάλο πρόβλημα, δεδομένου ότι κάποιος επικεντρώνεται στο ίδιο το παιχνίδι και όχι στη ευκρίνεια της εικόνας. Ένα μεγαλύτερο πρόβλημα έγκειται στην αναγνώριση, η οποία μπορεί να είναι τυχαία, μερικές φορές διαβάζοντας κινήσεις, ενώ άλλες φορές με ελάχιστη εγγραφή στην οθόνη. Αυτό είναι ιδιαίτερα ενοχλητικό αφού χαθεί πολύς χρόνος στην προσπάθεια να επιτευχθεί ένα αξιοπρεπές επίπεδο φωτισμού. Μια καλή λύση είναι η χρήση κομματιών άσπρου χαρτιού στα χέρια, το οποίο διευκολύνει τη σύλληψη της κίνησης από το EyeToy. Ωστόσο τα θέματα αυτά με τις εν λόγω δυσλειτουργίες είναι σποραδικά και πραγματικά δεν αναιρούν το γεγονός ότι πρόκειται για μια εντυπωσιακή τεχνολογία. Ενώ πολλές καινοτομίες έχουν περάσει κατά τη διάρκεια των χρόνων, το EyeToy παρέχει τεράστια δύναμη και δεν πρέπει να μείνει στο περιθώριο. Η ίδια η κάμερα είναι εύκολο να

ρυθμιστεί και να λειτουργήσει και φέρνει τον παίκτη άμεσα στη δράση. Το EyeToy της Sony προσφέρει κάτι μοναδικό στη σημερινή αγορά και σίγουρα αξίζει για παίκτες που θέλουν να δοκιμάσουν κάτι διαφορετικό. [34]

## 8.4 CamSpace

“Έχάστε το joystick, το πληκτρολόγιο και το ποντίκι”, διαλαλούν οι δημιουργοί της εφαρμογής CamSpace. Με την εφαρμογή αυτή μπορούμε να χειριστούμε το ηλεκτρονικό παιχνίδι μέσω της webcam μας. Η πλατφόρμα αυτή συνδέει τον εικονικό με τον πραγματικό κόσμο μέσω κινήσεων παιχνιδιού, εμπειριών και δραστηριοτήτων. Ανιχνεύει τις ανθρώπινες χειρονομίες και μετατρέπει καθημερινά αντικείμενα (όπως μπουκάλια, κουτιά, τενεκεδάκια και άλλα) σε συναρπαστικούς ελεγκτές καινούριων ή ήδη υπαρχόντων παιχνιδιών και εφαρμογών. Έχει προσελκύσει ευρύ κοινό και τρέχει μέσα από οποιοδήποτε περιηγητή διαδικτύου και είναι απλή και εύκολη στην χρήση της. Δεν απαιτεί κάποιο συγκεκριμένο υλικό. Το μόνο που χρειάζεται είναι μία σύνδεση στο Internet και μία webcam.



Εικόνα 93 : Το λογότυπο της πλατφόρμας CamSpace

Ο πυρήνας της τεχνολογίας του CamSpace είναι το λογισμικό το οποίο επιτρέπει σχεδόν σε κάθε κανονικό υπολογιστή (95% ποσοστό επιτυχίας) να ανιχνεύσει μέχρι και τέσσερα αντικείμενα – ακόμη και πολύ μικρά, από 5 χιλιοστά – σε πραγματικό χρόνο και με μεγάλη ακρίβεια και αξιοπιστία. Όλη η διαδικασία της ανίχνευσης γίνεται αυτόματα (x, y και z άξονες και γωνίες). Δουλεύει μόνο σε λειτουργικό περιβάλλον Windows. Οι αλγόριθμοι που χρησιμοποιήθηκαν είναι όλοι οι γνωστοί αλγόριθμοι ανίχνευσης οι οποίοι έχουν εξελιχθεί και προσαρμοστεί στις εκάστοτε ανάγκες και φυσικά υπάρχουν αλγόριθμοι που έχουν αναπτυχθεί ολοκληρωτικά από τους κατασκευαστές.

Το λογισμικό CamSpace χρειάζεται μια εφαρμογή την οποία θα διαχειριστεί και θα μιμηθεί την λειτουργία του ποντικιού και του πληκτρολογίου, ακόμη και του joystick. Οι χρήστες επιλέγουν το αντικείμενο που θα χρησιμοποιήσουν για την εφαρμογή και τον τρόπο που θα γίνει η μίμηση των συσκευών εισόδου. Για παράδειγμα, ένας χρήστης θα μπορούσε να κατασκευάσει έναν τροχό (από χαρτόνι για παράδειγμα) για να παίξει ένα

παιχνίδι με αγώνες αυτοκινήτων και να καθορίσει, μέσω του CamSpace, τι δηλώνει η κίνηση στον z άξονα (ίσως την λειτουργία του μοχλού ταχυτήτων). Ένας διαφορετικός χρήστης για το ίδιο παιχνίδι θα μπορούσε να χρησιμοποιήσει δύο αντικείμενα και να καθορίσει το δεύτερο αντικείμενο (για παράδειγμα ένα κουτί αναψυκτικού) ως μοχλό ταχυτήτων.

Η εφαρμογή αυτή αναπτύχθηκε από μια ομάδα τεσσάρων ανθρώπων. Τους Igal Lichtman και John F. Frankel ως επενδυτές και τους Yaron Tanne και Dory Zidon ως προγραμματιστές. Αυτήν την στιγμή κυκλοφορεί η έκδοση 8.8 του CamSpace και σε μορφή beta είναι δωρεάν. [35]

## 8.5 Project Natal

Το Project Natal, γνωστό και ως Kinect είναι ένα σύστημα χειρισμού των παιχνιδιών που επιτρέπει στους παίκτες να χρησιμοποιούν όλο τους το σώμα, ώστε να μπορούν να αλληλεπιδρούν με όσα γίνονται στην οθόνη. Αποτελεί την απάντηση της Microsoft στο Nintendo Wii και το απόκτημα της πρόσφατης εξαγοράς της 3DV Systems. Προσαρμοσμένο στις ανάγκες του Xbox 360 το Project Natal δεν είναι τίποτα περισσότερο ή λιγότερο από μια κάμερα και το απαραίτητο λογισμικό για αναγνώριση χειρονομιών σε πραγματικό χρόνο. Για τον χρήστη παιχνιδιών, αυτό σημαίνει έλεγχο κινήσεων και χαρακτήρων στην οθόνη χρησιμοποιώντας το ανθρώπινο σώμα. Χωρίς χειριστήρια και αισθητήρες κίνησης, τη χρήση επιταχυνσιομέτρου και τον φόβο τραυματισμού από σπασμένο λουράκι, το Xbox 360 είναι σε θέση να εκμεταλλευτεί το ανθρώπινο σώμα ως χειριστήριο σε παιχνίδια ενός ή πολλών παικτών.

Το Xbox 360 κατάφερε με το Project Natal, που αποτελεί έναν εντελώς καινούριο τρόπο παιχνιδιού, καταργώντας το χειριστήριο, να ενθουσιάσει τις κορυφαίες εταιρείες ανάπτυξης βιντεοπαιχνιδιών στον κόσμο. Δημιουργοί από όλο τον κόσμο έχουν ήδη αγκαλιάσει τη νέα τεχνολογία και ξεκίνησαν να εργάζονται για να προσφέρουν μοναδικές εμπειρίες ψυχαγωγίας για όλους.

Το Project Natal, είναι η κωδική ονομασία ενός εντελώς νέου τρόπου παιχνιδιού. Συνδυάζοντας μία κάμερα RGB, έναν αισθητήρα βάθους και ένα πολυδιάστατο μικρόφωνο μαζί με ειδικό λογισμικό, παρέχει στους παίκτες τη δυνατότητα να παίζουν με ένα νέο ρεαλιστικό και φιλικό τρόπο, καταρρίπτοντας τα εμπόδια που συνεπάγεται η χρήση χειριστηρίων και αξιοποιώντας τη κίνηση, ομιλία και όραση του παίκτη.



Εικόνα 94 : Η συσκευή Kinect

Παράλληλα, η εταιρεία αποκάλυψε ότι η τρισδιάστατη κάμερα “αξιολογεί” τρισεκατομμύρια διατάξεις του σώματος σε κάθε καρέ και αυτό το κάνει 30 φορές κάθε δευτερόλεπτο. [38][36]

### 8.5.1 Ιστορία

Το Kinect ανακοινώθηκε για πρώτη φορά την 1<sup>η</sup> Ιουνίου 2009 με το κωδικό όνομα "Project Natal". Ακολουθώντας την παράδοση της Microsoft που χρησιμοποιεί πόλεις σαν κωδικά ονόματα, το "Project Natal" προήλθε από την βραζιλιάνικη πόλη Νατάλ, προς τιμήν της χώρας του διευθυντή της Microsoft Alex Kipman, ο οποίος επωάζεται το έργο και ο οποίος είναι από Βραζιλία. Η ονομασία Natal επιλέχθηκε επίσης επειδή η λέξη "natal" σημαίνει “σχετικός με γέννηση”, αντικατοπτρίζοντας την άποψη της Microsoft για το έργο σαν την γέννηση της επόμενης γενιάς οικιακής ψυχαγωγίας. [26, (όρος Project Natal)]

### 8.5.2 Τεχνολογία

Ο αισθητήρας Kinect είναι μια οριζόντια μπάρα που συνδέεται με μια μικρή βάση και έχει σχεδιαστεί για να τοποθετείται κάτω από την οθόνη. Η συσκευή διαθέτει μια κάμερα RGB, έναν αισθητήρα βάθους και ένα μικρόφωνο, τα οποία παρέχουν δυνατότητες τρισδιάστατης σύλληψης κίνησης του σώματος, αναγνώρισης φωνής και προσώπου.

Ο αισθητήρας βάθους αποτελείται από ένα υπέρυθρο προβολέα σε συνδυασμό με ένα μονόχρωμο αισθητήρα CMOS και επιτρέπει στον αισθητήρα Kinect να δει τρισδιάστατα κάτω από οποιεσδήποτε συνθήκες φωτισμού. Το εύρος ανίχνευσης του αισθητήρα βάθους είναι ρυθμιζόμενο, με το λογισμικό Kinect να είναι ικανό για την αυτόματη διαρρύθμιση του αισθητήρα με βάση το παιχνίδι και το φυσικό περιβάλλον του παίκτη, όπως πχ την παρουσία των επίπλων.

Το Kinect βασίζεται στην τεχνολογία λογισμικού που αναπτύχθηκε από τη Microsoft και στην τεχνολογία range camera που αναπτύχθηκε από τον Ισραηλινή εταιρία PrimeSense, η οποία ερμηνεύει την τρισδιάστατη πληροφορία σκηνής από ένα συνεχώς προβαλλόμενο υπέρυθρο μοτίβο.

Πρωταρχική καινοτομία του Kinect σύμφωνα με την Microsoft είναι η τεχνολογία λογισμικού που επιτρέπει την προηγμένη αναγνώριση χειρονομιών, την αναγνώριση προσώπου, καθώς και την αναγνώριση φωνής. Το Kinect είναι σε θέση να παρακολουθεί ταυτόχρονα μέχρι έξι άτομα, μεταξύ των οποίων δύο ενεργούς παίκτες για ανάλυση κίνησης με εξαγωγή χαρακτηριστικών από 20 αρθρώσεις ανά παίκτη.

Ο αισθητήρας Kinect δίνει video εξόδου με ρυθμό καρτέ των 30 Hz, με τη ροή βίντεο RGB σε ανάλυση VGA χρώματος 32-bit (640 × 480 pixels) και τη μονόχρωμη ροή βίντεο που χρησιμοποιείται για την ανίχνευση βάθους σε ανάλυση QVGA 16-bit (320 × 240 pixel με 65.536 επίπεδα ευαισθησίας). Έχει οπτικό πεδίο των 57° οριζοντίως και 43° κατακόρυφα. [26, (όρος Project Natal)]

## 8.6 EyePoint

Όλοι έχουμε δει συστήματα ανίχνευσης ματιού υψηλής τεχνολογίας, που επιτρέπουν σε ανθρώπους με ειδικές ικανότητες να ελέγχουν τον υπολογιστή τους προσαρμόζοντας το βλέμμα τους, ή φθηνότερα μοντέλα που υπόσχονται να ελευθερώσουν τα χέρια παικτών ηλεκτρονικών παιχνιδιών. Ωστόσο, το πρόβλημα των συστημάτων αυτών είναι ότι αντιμετωπίζουν δυσκολίες στην αντιμετώπιση των γρήγορων και αυθόρμητων καταστάσεων των ματιών (ακούσιο ανοιγοκλείσιμο των ματιών για παράδειγμα) καθιστώντας τα έτσι επιρρεπή σε λάθη και χαρακτηρίζοντας τα με μια πενιχρή λειτουργικότητα. Αυτό, όμως, θα διορθωθεί, με τη βοήθεια του λογισμικού EyePoint που αναπτύχθηκε από τον ερευνητή του πανεπιστημίου του Stanford Manu Kumar και την ομάδα του. Χρησιμοποιεί το ήδη υπάρχον υλικό και μία webcamera υψηλής



ανάλυσης αλλά βελτιώνει την ακρίβεια του μέσω ειδικών αλγορίθμων που ασχολούνται με την περιοχή της κόρης του ματιού. Η τεχνολογία ανίχνευσης ματιού αναπτύχθηκε κυρίως για τους ανθρώπους με ειδικές ικανότητες. Αυτό, όμως, που προσπαθεί να καταφέρει ο Kumar είναι να την καταστήσει πιο χρήσιμη για όλους τους χρήστες. Οι περισσότεροι χρήστες αναζητούν απλές και εύκολες στη χρήση εφαρμογές και όλες αυτές που αναπτύχθηκαν για τους ανθρώπους με ειδικές ανάγκες δεν τους προσέλκυαν.



*Εικόνα 95 : Το λογότυπο του EyePoint*

Με το λογισμικό αυτό στοχεύουν να επαυξήσουν τις δυνατότητες και όχι να αντικαταστήσουν τις ήδη υπάρχουσες αλληλεπιδραστικές τεχνικές. Η πρόσθεση πληροφοριών βλέμματος προσφέρει εναλλακτικές τεχνικές αλληλεπίδρασης με τον υπολογιστή που οι χρήστες μπορούν να επιλέξουν με βάση τις ικανότητες τους, τις δραστηριότητες τους και τις προτιμήσεις τους. Η έρευνα αυτή μελετά την χρήση του βλέμματος σαν βοηθητική είσοδο για την εκτέλεση καθημερινών υπολογιστικών εργασιών. Πιο συγκεκριμένα, εξερευνά την ιδέα χρήσης του βλέμματος του χρήστη για να δείξει και να επιλέξει κάτι (pointing και selection), να εναλλάσσεται μεταξύ εφαρμογών (application switching), να εισάγει κάποιο κωδικό (password entry), να κινηθεί πάνω-κάτω ή δεξιά-αριστερά (scrolling), να μεγεθύνει ή να ελαχιστοποιήσει μια περιοχή (zooming) και να περιηγηθεί σε ένα έγγραφο (document navigation). Η χρήση του λογισμικού από διάφορους ανθρώπους έδειξε πως είναι δυνατόν να αξιοποιηθούν πρωτοποριακές αλληλεπιδραστικές τεχνικές που χρησιμοποιούν το βλέμμα ως είσοδο χωρίς να υπερφορτώνουν το οπτικό κανάλι και να ελαχιστοποιούν τις λανθασμένες ενεργοποιήσεις (αυθόρμητες κινήσεις ματιού).

Η περιήγηση στο διαδίκτυο, για παράδειγμα, με το λογισμικό αυτό γίνεται ως εξής : ο χρήστης κοιτάζει σε έναν σύνδεσμο και πατά ένα πλήκτρο του πληκτρολογίου με επιφορτισμένη λειτουργία (hot key, συνήθως βρίσκεται στο δεξί μέρος του πληκτρολογίου με την αριθμομηχανή). Η περιοχή στην οποία κοιτάζει μεγεθύνεται και στη συνέχεια ο χρήστης, αφού εστιάσει το βλέμμα του στη μεγεθυμένη περιοχή, απελευθερώνει το hot key και ανοίγει η ιστοσελίδα του συνδέσμου. [39]



Εικόνα 96 : Τα hot keys του λογισμικού

## Κεφάλαιο 9

# Προτάσεις για μελλοντική βελτίωση της εφαρμογής

Αν και η εφαρμογή που παρουσιάστηκε αποτελεί μια εισαγωγή στην ανίχνευση του ανθρώπινου σώματος σε εικονικά περιβάλλοντα, μπορεί να αποτελέσει τη βάση για πιο πολύπλοκες εφαρμογές με περισσότερες δυνατότητες και σκοπούς. Σαν συσκευή εισόδου περιοριστήκαμε στην χρήση μιας κάμερας, ωστόσο πιο πολύπλοκες συσκευές όπως το Wii Remote μπορούν να χρησιμοποιηθούν για κάθε είδους ανίχνευση. Η μεγάλη και ταχύτατη διείσδυση του νέου αυτού ελεγκτή στο κοινό και ο απλός σχεδιασμός του, μπορεί να συμβάλλει στη δημιουργία πιο εξελιγμένων συστημάτων ανίχνευσης με καλύτερη πιθανότητα ανάλυση και απουσία θορύβου.

Παράλληλα, πολλοί είναι αυτοί που υποστηρίζουν πως η τεχνολογία εικονικής πραγματικότητας μπορεί να προσφέρει ισχυρά οφέλη στην εκπαιδευτική διαδικασία επηρεάζοντας θετικά την μαθησιακή διαδικασία. Αυτό οφείλεται στο ότι παρέχει ισχυρή αλληλεπίδραση, άμεση ανταπόκριση του συστήματος στις ενέργειες του χρήστη και ελευθερία κινήσεων στην εικονική κατάσταση. Επιπλέον, η εικονική πραγματικότητα, συνδέεται με την φυσική συμπεριφορά. Ο προγραμματισμός, το πληκτρολόγιο, το ποντίκι μπορούν να αντικατασταθούν από φυσικότερες λειτουργίες του μαθητή όπως χειρονομίες, κίνηση και ομιλία. Στα πλαίσια της εισαγωγής της εικονικής πραγματικότητας στην εκπαίδευση, αρκετοί εκπαιδευτικοί χρησιμοποιούν παιχνίδια εικονικών κόσμων στις αίθουσες διδασκαλίας. Τα συμπεράσματα είναι εντυπωσιακά καθώς παρατηρείται αύξηση της κοινωνικό-αισθηματικής, γνωστικής και φυσικής ανάπτυξης του παιδιού, του συντονισμού χεριού-ματιού, της δημιουργικότητας, της συγκέντρωσης και των μηχανισμών επίλυσης προβλημάτων. Οι μαθητές μπορούν για παράδειγμα να περιηγηθούν στο εικονικό μουσείο, να εξετάζουν τα αντικείμενα μέσω της δυνατότητας ανίχνευσης χεριού ή να γράφουν ιδεατά το όνομα του στερεού που συναντούν και αν το αναγνωρίσουν σωστά, να εμφανίζεται το συγκεκριμένο στερεό σε μεγέθυνση σε ξεχωριστό παράθυρο με το όνομα του. Ειδικό παραδέχονται πως εκ πρώτης όψεως, η ιδέα της χρήσης παιχνιδιών εικονικών κόσμων ως εκπαιδευτικό εργαλείο φαίνεται αστεία, ωστόσο μελέτες αποδεικνύουν ότι τα παιχνίδια μπορούν να επιφέρουν από 7 έως 40% βελτίωση σε μαθησιακές πρακτικές. Σε αντίθεση με τις διαλέξεις, τα παιχνίδια μπορούν να προσαρμοστούν στο ρυθμό κάθε χρήστη.

## Βιβλιογραφία

### Βιβλία

[1] Θεοχάρης Θ., Μπεμ Α., Γραφικά, Αρχές & Αλγόριθμοι, Συμμετρία, Αθήνα, 1999

[2] Gary Bradsky, Adrian Kaehler, Learning OpenCV, Computer Vision with the OpenCV Library, O'Reilly, 2008

### Ηλεκτρονικά περιοδικά

[3] Vincent Lepetit, Pascal Fua, Monocular model-based 3D Tracking of Rigid Objects : A survey, Foundations and Trends in Computer Graphics and Vision, vol. 1, No. 1, 2005

### Ηλεκτρονικές δημοσιεύσεις

[4] Θεοφάνους Ν., Μαρούλης Δ., Παπαγέωργας Π., Σαγκριώτης Μ., Σγούρος Ν., Χαϊκάλης Δ., Τεχνολογία Τρισδιάστατης Εικόνας, Μεταβίβαση και Προκλήσεις

[5] Tobias Nowack, Peter Kurtz, Stefan Lutherdt, Torsten Gramsch, Martin Gerlach, Gunnar Sievers, Beate Schlütter, Design of an Evaluation Study for 3D Input Devices

[6] Diego Nehab, Pedro V. Sander, John Isidoro, The Real-Time Reprojection Cache

[7] Frédo Durand, A Short Introduction to Computer Graphics

[8] Amit K. Roy-Chowdhury and Yilei Xu, Face Tracking

[9] Daniel Mohr, Gabriel Zachmann, Real-Time Hand Tracking for Natural and Direct Interaction

[10] Samuel Gateau, The In and Out: Making Games Play Right with Stereoscopic 3D Technologies

[11] Buxton, Haptic Input, Chapter14: GESTURE BASED INTERACTION

[12] Helmut Jorke, Markus Fritz, Stereo Projection Using Interference Filters

[13] Martin Tall, NEOVISUS : GAZE INTERACTION INTERFACE COMPONENTS

## Ηλεκτρονικές σελίδες στο διαδίκτυο

- [14] <http://www.bobpowell.net/transformations.htm>
- [15] [http://www.ccrs.nrcan.gc.ca/resource/tutor/stereo/chap4/chapter4\\_1\\_e.php](http://www.ccrs.nrcan.gc.ca/resource/tutor/stereo/chap4/chapter4_1_e.php)
- [16] <http://www.cgl.ucsf.edu/chimera/docs/UsersGuide/stereo.html>
- [17] <http://www.codeproject.com/KB/usability/eye-tracking.aspx>
- [18] [http://experiencedynamics.blogs.com/site\\_search\\_usability/2004/12/eyetracking\\_stu.html](http://experiencedynamics.blogs.com/site_search_usability/2004/12/eyetracking_stu.html)
- [19] <http://www.gamedev.net/reference/programming/features/scenegraph/>
- [20] [http://www.haydenplanetarium.org/universe/partiviewguide/using\\_global\\_clip.php](http://www.haydenplanetarium.org/universe/partiviewguide/using_global_clip.php)
- [21] <http://www.rhino3d.com/nurbs.htm>
- [22] <http://www.sketchpad.net/basics1.htm>
- [23] <http://www.sweethome3d.com/index.jsp>
- [24] [http://www.tobii.com/corporate/eye\\_tracking/what\\_is\\_eye\\_tracking.aspx](http://www.tobii.com/corporate/eye_tracking/what_is_eye_tracking.aspx)
- [25] <http://www.vision3d.com/stereo.html>
- [26] [http://en.wikipedia.org/wiki/Main\\_Page](http://en.wikipedia.org/wiki/Main_Page)
- [27] [http://en.wikiversity.org/wiki/Introduction\\_to\\_Computers/Input\\_Devices](http://en.wikiversity.org/wiki/Introduction_to_Computers/Input_Devices)
- [28] <http://www.brighthub.com/computing/hardware/articles/35224.aspx>
- [29] <http://www.tech-faq.com/how-an-optical-mouse-works.html>
- [30] <http://www.cs.iit.edu/~agam/cs512/lect-notes/opencv-intro/>
- [31] <http://cgi.cs.indiana.edu/~oleykin/website/OpenCVHelp/>
- [32] <http://opencv.willowgarage.com/documentation/cpp/>
- [33] [http://opencv.jp/opencv-1.0.0\\_org/docs/ref/opencvref\\_highgui.htm#highgui\\_gui](http://opencv.jp/opencv-1.0.0_org/docs/ref/opencvref_highgui.htm#highgui_gui)
- [34] <http://www.the-laser.com/eyetoy.htm>
- [35] <http://camspace.com/>
- [36] [http://news.cnet.com/8301-10797\\_3-10253892-235.html](http://news.cnet.com/8301-10797_3-10253892-235.html)
- [37] <http://www.wisegeek.com/what-is-a-laser-mouse.htm>
- [38] <http://www.networkworld.com/community/blog/xbox-project-natal-now-called-kinect-and-look>
- [39] <http://hci.stanford.edu/research/GUIDe/index.html>
- [40] <http://graphics.cs.aueb.gr/graphics/downloads/EaZD/doc/EaZD.html>
- [41] <http://defense-update.com/products/v/VR-PTSD.htm>
- [42] <http://local.wasp.uwa.edu.au/~pbourke/exhibition/passive/>
- [43] <http://resumbrae.com/archive/papers/spie02/node4.html>
- [44] <http://www.howstuffworks.com/joystick.htm>

- [45] <http://www.epanorama.net/documents/joystick/intro.html>
- [46] [http://whatis.techtarget.com/definition/0,,sid9\\_gci798540,00.html](http://whatis.techtarget.com/definition/0,,sid9_gci798540,00.html)
- [47] <http://techstremist.com/archive3/>
- [48] <http://local.wasp.uwa.edu.au/~pbourke/miscellaneous/interpolation/>
- [49] <http://msdn.microsoft.com/en-us/library/system.inptr%28VS.80%29.aspx>
- [50] <http://hezhaonet.com/writings/ic.pdf>
- [51] [http://http.developer.nvidia.com/GPUGems/gpugems\\_ch11.html](http://http.developer.nvidia.com/GPUGems/gpugems_ch11.html)
- [52] [http://www.tml.tkk.fi/Opinnot/Tik-111.500/2002/paperit/timo\\_ahokas.pdf](http://www.tml.tkk.fi/Opinnot/Tik-111.500/2002/paperit/timo_ahokas.pdf)
- [53] <http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc42.htm>
- [54] <http://www.hcsnet.edu.au/node/1126>
- [55] <http://people.pwf.cam.ac.uk/jw566/research/stereo/system/matching.htm>

### Διατριβές και εισηγήσεις σε συνέδρια

[56] Θεοφάνους Ν., Σγουρός Ν., Αθηναίος Σπ., Παπαγέωργας Π., Σαγκριώτης Εμ., Μαρούλης Δ., Αντωνίου Ι., Δημάκης Α., Στερεοσκοπικές Διατάξεις-Ολοκληρωτική Φωτογραφία, 10<sup>ο</sup> Πανελλήνιο Συνέδριο Φυσικής, 2004

[57] Tom McReynolds, David Blythe, Advanced Graphics Programming Techniques Using OpenGL, April 26, 1998

[58] M.R. Fremouw, Real-time hand tracking using standard computer hardware, A thesis submitted for the degree of Master of Science in Computing Science, November 2009.

[59] Paul Viola, Michael Jones, Rapid Object Detection using a Boosted Cascade of Simple Features, Conference on Computer Vision and Pattern Recognition, 2001

## Παράρτημα

### A

**accelerometer** : ηλεκτρομηχανική συσκευή που έχει την ικανότητα να μετρά δυνάμεις επιτάχυνσης. Αυτές μπορεί να είναι στατικές, όπως η επιτάχυνση της βαρύτητας ή δυναμικές όταν προκαλούνται – προέρχονται από αλλαγές στην ταχύτητα ή στην διεύθυνση της κίνησης (επιταχύνσεις, επιβραδύνσεις, στροφές). Υπάρχουν μοντέλα για υπολογισμό της επιτάχυνσης σε έναν άξονα, ή σε πολλούς, ανιχνεύοντας το μέγεθος και την κατεύθυνση της επιτάχυνσης σαν μια διανυσματική ποσότητα και μπορεί να χρησιμοποιηθούν για να ανιχνεύουν τη θέση και την κατεύθυνση της συσκευής, την αίσθηση δόνησης και κίνησης. Επιταχυνσιόμετρα, χρησιμοποιούνται διαρκώς σε φορητές ηλεκτρικές συσκευές και ελεγκτές βιντεοπαιχνιδιών για να ανιχνεύσουν τον προσανατολισμό της συσκευής. [26, (όρος accelerometer)]

**affine transformation** : κάθε μετασχηματισμός που διατηρεί την συγγραμμικότητα (δηλαδή όλα τα σημεία που βρίσκονται αρχικά σε μια ευθεία εξακολουθούν να βρίσκονται σε αυτή μετά τον μετασχηματισμό) και την αναλογία των αποστάσεων (για παράδειγμα το κέντρο ενός ευθύγραμμου τμήματος εξακολουθεί να είναι κέντρο και μετά τον μετασχηματισμό). [1]

**algorithms for shading and illumination effects** : αλγόριθμοι που συνδυάζουν εξελιγμένες και πρωτοποριακές τεχνικές για την παραγωγή εικόνων ή βίντεο με μεγαλύτερη χρησιμότητα και την δυνατότητα να αποδίδουν την πραγματικότητα αρκετά ρεαλιστικά. [26, (όρος rendering και όρος rasterization)]

**animation** : (στα ελληνικά αποδίδεται με τους όρους Εμφύχωση, Κινούμενη Εικόνα είτε περιφραστικά Απόδοση Κίνησης στην Εικόνα) η ταχεία προβολή μιας σειράς από εικόνες (δισδιάστατης ή τρισδιάστατης μακέτας) ή θέσεων ενός μοντέλου, έτσι ώστε να δημιουργείται η ψευδαίσθηση της κίνησης. [26, (όρος animation)]

**Augmented Reality (AR)** : όρος που χρησιμοποιείται όταν βλέπουμε σε πραγματικό χρόνο ένα αληθοφανές περιβάλλον, του οποίου τα στοιχεία έχουν συγχωνευθεί με στοιχεία που προήλθαν από τον υπολογιστή, δημιουργώντας έτσι μία “ανακατεμένη” πραγματικότητα. Βρίσκεται συμβατικά σε σημασιολογικά και πραγματικού χρόνου

κείμενα με στοιχεία περιβάλλοντος, όπως, για παράδειγμα, αθλητικά σκορ στην τηλεόραση κατά την διάρκεια ενός αγώνα. [26, (όρος Augmented Reality)]

**avatar:** Είναι η αναπαράσταση του χρήστη ενός υπολογιστή είτε με τη μορφή ενός τρισδιάστατου μοντέλου που χρησιμοποιείται σε ηλεκτρονικά παιχνίδια, είτε μιας δισδιάστατης εικόνας. Πρόκειται για ένα αντικείμενο που αντιπροσωπεύει τον χρήστη. Ο όρος "avatar" μπορεί να αναφερθεί δηλαδή στην προσωπικότητα που συνδέεται με την οθόνη, το χρήστη δηλαδή του υπολογιστή, ή το χρήστη του Διαδικτύου. [26, (όρος Avatar (computing))]

## B

**beta :** είναι ένα στάδιο στη φάση ανάπτυξης ενός προϊόντος λογισμικού. Προέρχεται από την ελληνική λέξη "βήτα". Στη φάση αυτή το λογισμικό δεν είναι στην τελική του μορφή αλλά περιέχει όλες τις βασικές λειτουργίες της τελικής του έκδοσης. Είναι ουσιαστικά ένα στάδιο "προεπισκόπησης τεχνολογίας". [26, (όρος Software release life cycle)]

**Bloom effect :** είναι φαινόμενο που εμφανίζεται στα γραφικά υπολογιστών και χρησιμοποιείται κυρίως σε παιχνίδια υπολογιστών. Το φαινόμενο αυτό δημιουργεί "πλαίσια" φωτός γύρω από πολύ φωτεινά αντικείμενα σε μια εικόνα. [26, (όρος Bloom (shader effect))]



Παράδειγμα του φαινομένου bloom. Το φως από το φωτεινό φόντο τροφοδοτεί τις σκοτεινές περιοχές, όπως τοίχους και χαρακτήρες.

**buffer :** μία προσωρινή αποθηκευτική περιοχή, συνήθως στην RAM. Ο σκοπός των περισσότερων buffers είναι να λειτουργήσουν ως αποθήκη επιτρέποντας στην CPU να χειριστεί τα δεδομένα, όπως αυτά εισάγονται από κάποια συσκευή εισόδου, προτού τα μεταφέρει σε κάποια συσκευή εξόδου. [26, (όρος data buffer)]



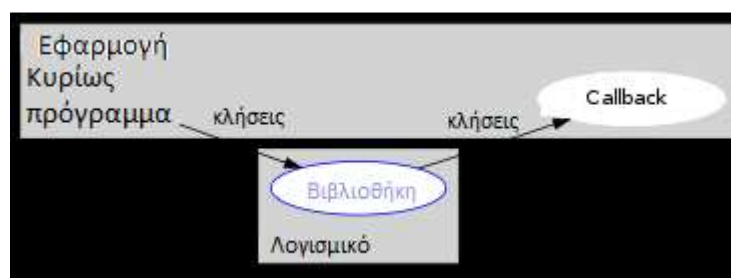
**BSD licence** : ανήκει στην οικογένεια παραχώρησης αδειών για ελεύθερα λογισμικά και χρησιμοποιήθηκε αρχικά από την Berkley Software Distribution, από την οποία και το όνομα. [26, (όρος BSD licence)]

## C

**CAD/ CAM** : (computer-aided design/ computer-aided manufacturing), υπολογιστικά συστήματα που χρησιμοποιούνται για τον σχεδιασμό και την παραγωγή προϊόντων. Ένας μηχανικός, δηλαδή, μπορεί να χρησιμοποιήσει το σύστημα τόσο για τον σχεδιασμό ενός προϊόντος όσο και για τον έλεγχο της διαδικασίας παραγωγής του. Για παράδειγμα μόλις δημιουργηθεί ένα σχέδιο με την βοήθεια του CAD συστατικού, το ίδιο το σχέδιο μπορεί να ελέγξει τα μηχανήματα που θα κατασκευάσουν το τμήμα αυτό. [26, (όρος Computer-aided technologies)]

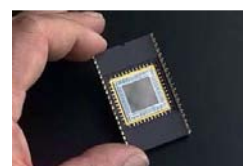
**calibration** : το σύνολο των λειτουργιών που αναπτύσσουν, κάτω από συγκεκριμένες συνθήκες, την σχέση μεταξύ της αξίας ποσοτήτων που υποδεικνύονται από ένα εργαλείο μέτρησης και της αξίας που τους αντιστοιχεί και παράγεται από πρότυπα. Το αποτέλεσμα της διαδικασίας αυτής επιτρέπει είτε την ανάθεση των αξιών των μετρήσεων στις ενδείξεις, είτε τον καθορισμό διορθώσεων σημείων με σεβασμό πάντοτε στις ενδείξεις. [46]

**callback** : αναφορά σε κάποιο κώδικα, ή κάποιο κομμάτι ενός κώδικα, που περνά ως όρισμα σε αυτόν. Αυτό επιτρέπει σε ένα στρώμα λογισμικού χαμηλού επιπέδου να καλεί συναρτήσεις που ορίζονται σε ένα στρώμα υψηλού επιπέδου. [26, (όρος callback)]



Μία αναφορά callback συνήθως βρίσκεται κάτω από το επίπεδο του καλούντος

**CCD camera** : μια συσκευή που είναι σχεδιασμένη έτσι ώστε να μετατρέπει οπτική φωτεινότητα σε ηλεκτρικά σήματα υψηλής συχνότητας, χρησιμοποιώντας πολλά CCDs (συσκευές για την



Ένα από τα πολλά CCDs μιας κάμερας CCD

κίνηση του ηλεκτρικού φορτίου), και στην συνέχεια αναπαράγει την εικόνα ενός αντικειμένου χωρίς χρονικό περιορισμό. [26, (όρος CCD camera)]

**Charles Wheatstone** : (6 Φεβρουαρίου 1802 – 19 Οκτωβρίου 1875), βρετανός επιστήμονας και εφευρέτης πολλών σημαντικών επιστημονικών ανακαλύψεων της βικτοριανής εποχής. Συμπεριλαμβάνεται το στερεοσκόπιο (συσκευή για την απεικόνιση τρισδιάστατων εικόνων), της αγγλικής κονσερτίνας (concertina, μουσικό όργανο όπως το ακορντεόν) και το Playfair cipher (τεχνική κρυπτογραφίας). [26, (όρος Charles Wheatstone)]

**clustering tools** : ο όρος clustering χρησιμοποιείται για να περιγράψει την διαδικασία οργάνωσης αντικειμένων σε ομάδες που τα μέλη τους είναι σχεδόν όμοια. Είναι ένα από τα σημαντικότερα προβλήματα μάθησης. Τα εργαλεία αυτά, λοιπόν, βοηθούν στην διεξαγωγή της διαδικασίας αυτής. [26, (όρος Cluster analysis)]

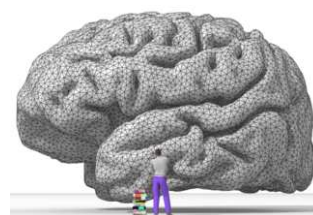
**CMake** : είναι ένα πρόγραμμα ανοιχτού κώδικα που δουλεύει σε όλες τις πλατφόρμες. Χρησιμοποιείται για να ελέγχει και να βοηθά στην ολοκλήρωση της διαδικασίας μεταγλώττισης ενός πακέτου λογισμικού με τον μεταγλωττιστή της επιλογής μας. [26, (όρος CMake)]



Το λογότυπο του CMake

**cognitive science** : η μελέτη του μυαλού ή της σκέψης. Περιλαμβάνει αρκετούς

επιστημονικούς κλάδους ερευνών, όπως η ψυχολογία, η τεχνητή νοημοσύνη, η φιλοσοφία, η ανθρωπολογία, η βιολογία και η γλωσσολογία. Βασίζεται σε μια ποικιλία



Απόδοση του ανθρώπινου εγκεφάλου επιστημονικών μεθοδολογιών (για παράδειγμα πειράματα συμπεριφοράς, προσομοιώσεις συλλογισμών) και καλύπτει πολλά επίπεδα ανάλυσης του μυαλού (από χαμηλού επιπέδου μηχανισμούς μάθησης και απόφασης σε υψηλού επιπέδου λογική και σχεδιασμό). [26, (όρος cognitive science)]

**CMOS** : (complementary metal-oxide-semiconductor), τεχνολογία κατασκευής ολοκληρωμένων κυκλωμάτων. Η τεχνολογία αυτή χρησιμοποιείται σε μικροεπεξεργαστές, μικροελεγκτές, στατικές μνήμες RAM και άλλα ψηφιακά λογικά

κυκλώματα. Χρησιμοποιείται επίσης για αρκετά αναλογικά κυκλώματα, όπως αισθητήρες εικόνας, μετατροπείς δεδομένων, καθώς και πομποδέκτες με υψηλό βαθμό ολοκλήρωσης για πολλά είδη επικοινωνίας. Τα CMOS αποτελούν ευρεσιτεχνία του Frank Wanlass. Τα CMOS μερικές φορές αναφέρονται ως ημιαγωγοί συμπληρωματικής-συμμετρίας. Η “συμπληρωματική συμμετρία” αναφέρεται στο γεγονός ότι η τυπική ψηφιακή σχεδίαση με CMOS χρησιμοποιεί συμπληρωματικά και συμμετρικά ζεύγη p-τύπου και n-τύπου τρανζίστορ για τις λογικές λειτουργίες. Βασικό χαρακτηριστικό τους είναι η χαμηλή στατική κατανάλωση ισχύος. [26, (όρος cmos)]

**computer vision** : η επιστήμη και η τεχνολογία των μηχανημάτων που μπορούν και “βλέπουν”. Αναφέρεται στη θεωρία ανάπτυξης τεχνητών συστημάτων που αποκτούν πληροφορίες από images. Οι πληροφορίες αυτές μπορεί να είναι σε διάφορες μορφές, όπως βίντεο, λήψεις από διαφορετικές κάμερες ή πολυδιάστατες πληροφορίες από κάποιο σαρωτή (scanner). [26, (όρος computer vision)]

**CRT projectors** : Όταν πρωτοεμφανίστηκαν οι προβολείς η τεχνολογία της τηλεόρασης βασιζόταν στον “καθοδικό σωλήνα ακτίνας” (Cathode Ray Tube, CRT) στον οποίο ο θεατής βλέπει την εικόνα. Τρεις μικροί καθοδικοί σωλήνες (ένας για κάθε βασικό χρώμα) συνοδευόμενοι, ο καθένας, με έναν μεγεθυντικό φακό φωτός μπορούν να προβάλλουν μία εικόνα χρώματος σε μία τεράστια οθόνη μέσα σε ένα σκοτεινό δωμάτιο. Αν και οι προβολείς αυτοί δεν είναι διαθέσιμοι στο καταναλωτικό κοινό για γενική χρήση, χρησιμοποιούνται ακόμη για διαφημιστικούς και βιομηχανικούς σκοπούς. [26, (όρος CRT projector)]

## D

**dataglove** : Γάντι εφοδιασμένο με αισθητήρες, που ανιχνεύει τις κινήσεις του χεριού

και μεσολαβεί στην αλληλεπίδραση αυτών των κινήσεων με τον υπολογιστή. Χρησιμοποιείται συχνά σε περιβάλλοντα εικονικής πραγματικότητας, όπου ο χρήστης βλέπει και αντιδρά με τους χαρακτήρες του περιβάλλοντος με τη βοήθεια του. [26, (όρος dataglove)]



Ένα dataglove

**deferred rendering** : τεχνική που εφαρμόζεται σε τρισδιάστατα περιβάλλοντα, κατά την οποία ο αλγόριθμος εξελίσσεται διαιρώντας τον σε μικρότερα τμήματα που

αποθηκεύονται σε έναν ενδιάμεσο αποθηκευτικό χώρο (buffer) για να συνδυαστούν αργότερα όταν συγκεντρώσει πληροφορίες για όλα τα σημεία. Σε άλλες τεχνικές το αποτέλεσμα του αλγορίθμου αποθηκεύεται απευθείας στον buffer. Βρίσκει ιδιαίτερη εφαρμογή σε τρισδιάστατα παιχνίδια όπου βελτιώνεται η ποιότητα των γραφικών καθώς αντιδρά αρκετά καλά με τις δυσκολίες φωτισμού αφού τα pixels που δεν είναι ορατά δεν φωτίζονται. [26, (όρος deferred rendering)]

**depth perception** : η οπτική ικανότητα να παρατηρούμε τον κόσμο και στις τρεις διαστάσεις. [26, (όρος depth perception)]

**DSP** : (digital signal processor) είναι ένας εξειδικευμένος μικροεπεξεργαστής με βελτιστοποιημένη αρχιτεκτονική για τις ανάγκες σε ταχύτητα της ψηφιακής επεξεργασίας σήματος. Οι αλγόριθμοι ψηφιακής επεξεργασίας σήματος συνήθως απαιτούν ένα μεγάλο αριθμό μαθηματικών πράξεων που πρέπει να εκτελεστούν γρήγορα και επανειλημμένα σε ένα σύνολο δεδομένων. Σήματα (ίσως από αισθητήρες ήχου ή βίντεο) συνεχώς μετατρέπονται από αναλογικά σε ψηφιακά και στη συνέχεια μετατρέπονται πάλι σε αναλογική μορφή. Πολλές εφαρμογές DSP έχουν περιορισμούς όσον αφορά την καθυστέρηση. Δηλαδή, για να λειτουργήσει το σύστημα, η λειτουργία DSP πρέπει να ολοκληρωθεί εντός ενός καθορισμένου χρόνου. Οι περισσότεροι μικροεπεξεργαστές γενικού σκοπού και τα λειτουργικά συστήματα μπορούν να εκτελέσουν αλγορίθμους DSP ικανοποιητικά, αλλά δεν είναι κατάλληλοι για χρήση σε φορητές συσκευές όπως τα κινητά τηλέφωνα, λόγω περιορισμών τροφοδοσίας και χώρου. Ένας επεξεργαστής DSP ωστόσο, τείνει να παρέχει μια λύση με χαμηλότερο κόστος, με καλύτερες επιδόσεις, μικρότερη καθυστέρηση χωρίς απαιτήσεις για εξειδικευμένη ψύξη ή μεγάλες μπαταρίες. [26, (όρος Digital signal processor)]

**DirectInput** : στοιχείο του DirectX. Μια διεπαφή για χειρισμό εισόδου από τον χρήστη, μέσω συσκευών εισόδου, όπως το ποντίκι, το πληκτρολόγιο, το joystick και άλλους ελεγκτές παιχνιδιών. Παρέχει, επίσης, ένα σύστημα που επιτρέπει στον χρήστη να αναθέτει συγκεκριμένες ενέργειες, στην διάρκεια ενός παιχνιδιού, στα κουμπιά και τους άξονες των συσκευών εισόδου. Ασχολείται και με συσκευές ανάδρασης εισόδου/εξόδου. [26, (όρος DirectInput)]

**DirectX** : Το DirectX βοηθά τον υπολογιστή να βελτιώσει τις επιδόσεις απαιτητικών προγραμμάτων πολυμέσων όπως είναι τα παιχνίδια και οι ταινίες. Το DirectX είναι μια

τεχνολογία που βοηθά στη δημιουργία των ειδικών οπτικών και ηχητικών εφέ που συνοδεύουν τα παιχνίδια. [26, (όρος DirectX)]

**DLP projector** : (Digital Light Processing), ο πιο ευρέως χρησιμοποιούμενος προβολέας τα τελευταία χρόνια. Η εικόνα απεικονίζεται σε ένα chip όπως και στους προβολείς LCD. Αυτό που αλλάζει είναι το chip που χρησιμοποιείται και στους προβολείς DLP ονομάζεται DMD. Κάθε pixel σε αυτό το chip είναι στην ουσία ένας ανακλώμενος καθρέφτης. Καθώς η εικόνα αλλάζει οι “καθρέφτες” του chip γέρνουν ελαφρά και πολύ γρήγορα. Η διαδικασία αυτή δημιουργεί την εικόνα σε ασπρόμαυρη μορφή. Στη συνέχεια προστίθεται χρώμα καθώς περνά φως μέσα από έναν τροχό γρήγορης ταχύτητας και αντανακλάται στους “καθρέφτες” του chip. [26, (όρος Digital Light Processing)]

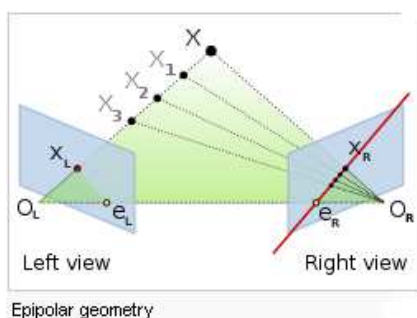


Ένας οικιακός προβολέας DLP

**DOF** : (degrees-of-freedom), ο αριθμός των στοιχείων που απαιτούνται για τον καθορισμό της “θέσης” ή και της “προέλευσης” ενός συστήματος ή ενός αντικειμένου. [26, (όρος Degrees of freedom (mechanics))]

## Ε

**epipolar geometry** : Αναφέρεται στην γεωμετρία της στερεοσκοπικής όρασης. Όταν δύο κάμερες βλέπουν μια τρισδιάστατη σκηνή από δύο ξεχωριστές τοποθεσίες, υπάρχουν αρκετές γεωμετρικές σχέσεις μεταξύ των τρισδιάστατων σημείων και των προβολών τους στις δισδιάστατες εικόνες, που οδηγούν σε περιορισμούς. Αυτές οι σχέσεις βασίζονται στην υπόθεση ότι οι κάμερες μπορούν να αναπαρασταθούν με το μοντέλο pinhole camera model



Η εικόνα απεικονίζει δύο pinhole κάμερες που βλέπουν το σημείο X. Στις πραγματικές κάμερες, το επίπεδο της εικόνας είναι πίσω από το εστιακό σημείο και παράγει μια περιστρεφόμενη εικόνα. Εδώ, το πρόβλημα απλοποιείται τοποθετώντας ένα εικονικό επίπεδο εικόνας μπροστά από το εστιακό σημείο για την παραγωγή μιας μη περιστρεφόμενης

εικόνας. Τα σημεία  $O_L$  και  $O_R$  αντιπροσωπεύουν τα εστιακά σημεία των δύο καμερών. Τα σημεία  $x_L$  και  $x_R$  είναι οι προβολές του σημείου  $X$  στα επίπεδα της εικόνας. Κάθε κάμερα "συλλαμβάνει" μια δισδιάστατη εικόνα του τρισδιάστατου κόσμου. Αυτήν η μετατροπή, από 3D σε 2D, αναφέρεται ως προοπτική προβολή και περιγράφεται από το μοντέλο pinhole camera. [26, (όρος Epipolar geometry)]

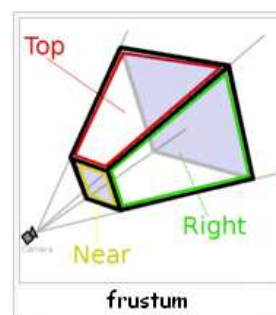
**Ethernet** : Το Ethernet είναι τύπος δικτύωσης υπολογιστών. Πλέον είναι ο μοναδικός τρόπος διασύνδεσης τοπικών δικτύων που έχει επικρατήσει. Για την σύνδεση δύο υπολογιστών μεταξύ τους σε δίκτυο Ethernet απαιτούνται:

- 1) Μία κάρτα δικτύου τύπου Ethernet σε κάθε υπολογιστή και καλώδιο αντεστραμμένων ζευγών που να συνδέει τις δύο κάρτες δικτύου.
- 2) Για τη σύνδεση δύο ή περισσότερων Η/Υ, αντί του αντεστραμμένου καλωδίου (που συνδέει μόνο δύο Η/Υ) μπορεί να χρησιμοποιηθεί ένα Hub/Switch στο οποίο θα συνδέονται όλες οι κάρτες δικτύου με απλό δικτυακό καλώδιο (UTP). Το Hub ή Switch όπως λέγονται τα νέας τεχνολογίας hub, λειτουργούν σαν ένα "πολύπριζο" για το δίκτυο. [26, (όρος Ethernet)]

## F

**fiducial** : είναι ένα αντικείμενο που χρησιμοποιείται σε συστήματα απεικόνισης και εμφανίζεται στην εικόνα που παράγεται για χρήση σαν δείκτης αναφοράς. [26, (όρος Fiduciary marker)]

**frustum** : στερεό σχήμα που παράγεται, όταν δύο παράλληλα επίπεδα κόψουν ένα ευρύτερο στερεό. Όρος που χρησιμοποιείται συχνά στα γραφικά υπολογιστών για να περιγράψει την περιοχή που είναι ορατή στην οθόνη. Είναι το οπτικό πεδίο της ιδεατής κάμερας. Το σχήμα αυτής της περιοχής εξαρτάται από το είδος των φακών της κάμερας που χρησιμοποιούνται.



Σχηματική απεικόνιση του frustum

Συνήθως είναι ένα κομμάτι μιας πυραμίδας. Τα επίπεδα που κόβουν την πυραμίδα δίνοντας το ορθογώνιο σχήμα στην κατεύθυνση θέασης, είναι το near και far. Αντικείμενα πιο κοντά στην κάμερα από το επίπεδο near ή πίσω από το επίπεδο far δεν φαίνονται. [26, (όρος Viewing frustum)]

**G**

**gamepad** : (joypad ή control pad) τύπος ελεγκτή παιχνιδιών που κρατιέται και στα δύο χέρια και ελέγχεται με τους αντίχειρες. Τα gamepads έχουν ένα σύνολο από κουμπιά δράσης που τα χειρίζεται ο δεξιός αντίχειρας και έναν ελεγκτή κατεύθυνσης που τον χειρίζεται ο αριστερός. [26, (όρος gamepad)]

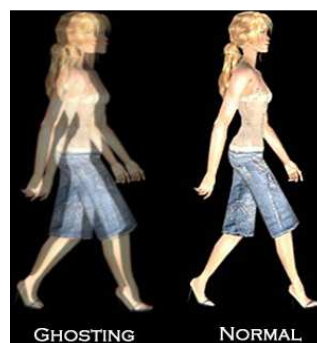


Ένα gamepad

**Gaussian distribution** : η πιο γνωστή αλλά και η πιο χρήσιμη από όλες τις κατανομές. Χρησιμοποιεί και το όνομα κανονική κατανομή. Στην κατανομή αυτή η καμπύλη συχνοτήτων των δεδομένων έχει την μορφή καμπάνας. [26, (όρος Normal distribution)]

**graphics processing unit (GPU)** : εξειδικευμένος μικροεπεξεργαστής που επιτρέπει την επιτάχυνση των τρισδιάστατων ή δισδιάστατων διαδικασιών rendering. Χρησιμοποιείται σε ενσωματωμένα συστήματα, κινητά τηλέφωνα, προσωπικούς ηλεκτρονικούς υπολογιστές και σε κονσόλες ηλεκτρικών παιχνιδιών. Είναι πολύ αποδοτικοί όσον αφορά τα γραφικά υπολογιστών και η υψηλή παράλληλη δομή τους καθιστά πιο αποτελεσματικούς από τους γενικού σκοπούς επεξεργαστές σε μια ποικιλία πολύπλοκων αλγορίθμων. Σε έναν προσωπικό υπολογιστή μπορεί να βρίσκεται σε μια κάρτα γραφικών ή στην μητρική. [26, (όρος Graphics processing unit)]

**ghosting**: η εμφάνιση μιας δευτερεύουσας εικόνας, η οποία επικαλύπτει την πρωτότυπη εικόνα. Πιο συγκεκριμένα, είναι ένα αντίγραφο της επιθυμητής εικόνας που εμφανίζεται πιο αχνό και πιο πίσω σε σχέση με το αρχικό. [47]



Φαινόμενο ghosting

**gradient descent** : αλγόριθμος βελτιστοποίησης. Για την εύρεση ενός τοπικού ελαχίστου μιας συνάρτησης προχωρά με αναλογικά βήματα προς το αρνητικό τμήμα της συνάρτησης. [26, (όρος gradient descent)]

**QVGA** : (Quarter Video Graphics Array) ανάλυση σε 320x240 pixel. Μπορεί να αναφέρεται στην ανάλυση ενός φυσικού στοιχείου, όπως μιας οθόνης ή του ελεγκτή μιας κάμερας, ή στην ανάλυση μιας φωτογραφίας ή ενός αρχείου βίντεο. Χρησιμοποιείται σε ηλεκτρονικές συσκευές, όπως κινητά τηλέφωνα, PDAs και σε κάποιες κονσόλες παιχνιδιών. Υψηλότερες αναλύσεις προσφέρουν περισσότερες λεπτομέρειες. Παρέχει μικρότερη ανάλυση από το VGA το οποίο δίνει ανάλυση 640x480. Ονομάζεται έτσι επειδή η προσφερόμενη ανάλυση είναι το ένα τέταρτο (quarter) της πραγματικής ανάλυσης οθόνης (640x480). [26, (όρος Graphic display resolutions)]

## H

**header αρχεία** : χρησιμοποιούνται από μερικές γλώσσες προγραμματισμού (όπως η C και η C++). Επιτρέπουν στους προγραμματιστές να ξεχωρίσουν ορισμένα στοιχεία του πηγαίου κώδικα και να τα τοποθετήσουν σε αρχεία τα οποία μπορούν να χρησιμοποιηθούν και αλλού. Συνήθως περιέχουν ορισμούς κλάσεων, συναρτήσεων, μεταβλητών και άλλων αναγνωριστικών. [26, (όρος Header file)]

## I

**interpolation techniques** : η ιδέα που χρησιμοποιούν αυτές οι τεχνικές είναι ότι τα σημεία βρίσκονται σε σωστή σειρά κατά μία έννοια, το πρόβλημα είναι η εκτίμηση των τιμών της καμπύλης σε οποιαδήποτε θέση με βάση γνωστά σημεία. Χρησιμοποιούμε, δηλαδή, γνωστές πληροφορίες για την εκτίμηση άγνωστων. [48]

**IntPtr** : ένας ειδικός τύπος ακεραίου που ερμηνεύεται κάθε φορά από την εκάστοτε πλατφόρμα και χρησιμοποιείται συνήθως για να χαρακτηρίσει έναν δείκτη. [49]

**inverse compositional approach** : χρησιμοποιείται κυρίως για την ευθυγράμμιση εικόνων και υποστηρίζει γεωμετρικούς μετασχηματισμούς. Προσπαθεί να επιλύσει το πρόβλημα εκτελώντας αρκετές επαναλήψεις. [50]

## J

**Jacob matrix** : (Ιακωβιανός πίνακας), ένας πίνακας με στοιχεία του τις μερικές παραγώγους συναρτήσεων. [26, (όρος Jacobian matrix and determinant)]

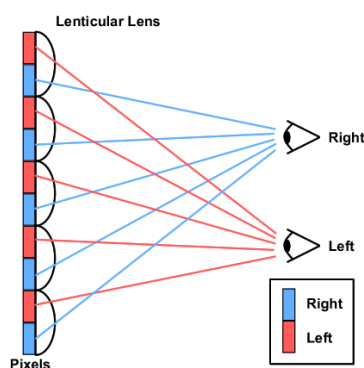


**L**

**LCD projector** : (Liquid Crystal Display), δεν βασίζεται στον παραδοσιακό σωλήνα προβολής, όπως οι προβολείς CRT. Λειτουργεί περνώντας φως από μία ισχυρή πηγή φωτός μέσω ενός διαφανούς chip που αποτελείται από ξεχωριστά pixels (τα οποία απεικονίζουν την εικόνα) και προβάλλει την εικόνα μέσω ενός φακού σε μια μεγάλη οθόνη. [26, (όρος LCD projector)]



**lenticular lens** : (κυλινδροειδής συστοιχία φακών), συστοιχία μεγεθυντικών φακών, οι οποίοι σχεδιάζονται έτσι ώστε όταν κοιτάμε από διαφορετικές γωνίες διαφορετικές εικόνες να μεγεθύνονται. Χρησιμοποιείται σε διάφορες εφαρμογές ώστε να δώσει ψευδαίσθηση βάθους ή οι εικόνες που φαίνονται να αλλάζουν ή να κινούνται καθώς τις επεξεργαζόμαστε από διαφορετικές γωνίες. [26, (όρος Lenticular lens)]



Λειτουργία συστοιχίας

Μία μπλε και μία κόκκινη εικόνα χωρίζονται σε λωρίδες και συνδυάζονται μεταξύ τους. Σε κάθε κόκκινο-μπλε ζευγάρι τοποθετείται ένας φακός. Καθώς αλλάζει η γωνία που κοιτάμε θα βλέπουμε είτε την μπλε είτε την κόκκινη εικόνα.

**local optimization methods** : προβλήματα βελτιστοποίησης κατά τα οποία οι πληροφορίες που χρησιμοποιούνται για την ανανέωση της προσέγγισης κάθε φορά είναι από την κοντινή “γειτονιά” της. [26, (όρος Local search (optimization))]

**look up matrix** : είναι μια δομή δεδομένων, που χρησιμοποιείται συχνά για να αντικαταστήσει έναν υπολογισμό με μια απλούστερη διαδικασία ευρετηρίου σε μορφή πίνακα. Το κέρδος από την άποψη χρόνου επεξεργασίας μπορεί να είναι σημαντικό, δεδομένου ότι η ανάκτηση τιμής από τη μνήμη είναι συχνά πιο γρήγορη από έναν “ακριβό” υπολογισμό ή λειτουργία εισόδου / εξόδου. Οι πίνακες αυτοί μπορεί να έχουν υπολογιστεί εκ των προτέρων και να έχουν αποθηκευτεί σε ένα στατικό πρόγραμμα

αποθήκευσης ή να υπολογίζονται ως μέρος της φάσης αρχικοποίησης των προγραμμάτων. [26, (όρος look up table)]

**Lucas - Kanade algorithm** : ο αλγόριθμος αυτός υποθέτει ότι η λύση βρίσκεται στα γειτονικά pixels και για αυτό οδηγείται στο αποτέλεσμα συνδυάζοντας πληροφορίες από γειτονικά pixels. [26, (όρος Lucas-Kanade Optical Flow Method)]

## M

**mousepad** : επιφάνεια που διευκολύνει τη χρήση του ποντικιού ενός ηλεκτρονικού υπολογιστή. Τα τρία πιο σημαντικά οφέλη από την χρήση του mousepad είναι η υψηλότερη ταχύτητα, η μεγαλύτερη ακρίβεια και η άνεση για τον χρήστη. Ένα δευτερεύον όφελος είναι η προστασία του γραφείου από γδαρσίματα λόγω της συνεχής κίνησης του ποντικιού, αλλά και η διατήρηση του κάτω μέρους του ποντικιού καθαρού. Όταν το οπτικό ποντίκι εμφανίστηκε στην αγορά απαιτούσε ειδικά κατασκευασμένα mousepads με οπτικά μοτίβα αποτυπωμένα σε αυτά. Τα σύγχρονα οπτικά ποντίκια μπορούν να λειτουργήσουν με ακρίβεια πάνω από χαρτί ή άλλες επιφάνειες. Τα σύγχρονα mousepads είναι συνήθως κατασκευασμένα από καουτσούκ (ανοικτή με ύφασμα προσκολλημένο στην επάνω επιφάνεια. [26, (όρος Mousepad)]

**multimodal systems** : (Πολυτροπικά συστήματα), τα πολυτροπικά συστήματα υποστηρίζουν τη φυσικότερη και πιο αποτελεσματική αλληλεπίδραση ανθρώπου-υπολογιστή με την παροχή πολλαπλών τρόπων εισαγωγής και παραγωγής δεδομένων όπως την ομιλία, τις χειρονομίες των χεριών, το βλέμμα και τις οθόνες δυναμικών γραφικών. Το κρίσιμο πλεονέκτημά τους είναι ότι επιτρέπουν την εισαγωγή δεδομένων και την παραγωγή αποτελεσμάτων με τον τρόπο ή τους τρόπους τους οποίους είναι καταλληλότεροι για τον κάθε χρήστη, λαμβάνοντας υπόψη τον επιθυμητό στόχο, τις προτιμήσεις του χρήστη και το φυσικό και κοινωνικό περιβάλλον της αλληλεπίδρασης. [54]

**multiple targets** : (multiple render target, MRT), μία τεχνική όπου ο φωτισμός της σκηνής μπορεί να γίνει αφού τελειώσει η διαδικασία rendering για όλα τα υπόλοιπα γεωμετρικά στοιχεία, μειώνοντας έτσι τα πολλαπλά περάσματα από την σκηνή. Το αποτέλεσμα του φωτισμού μπορεί να γίνει ρεαλιστικό και παράλληλα να αποφύγουμε τον ανούσιο χρόνο επεξεργασίας των pixels που δεν συνεισφέρουν στο ορατό κομμάτι της σκηνής. [26, (όρος Multiple Render Targets)]

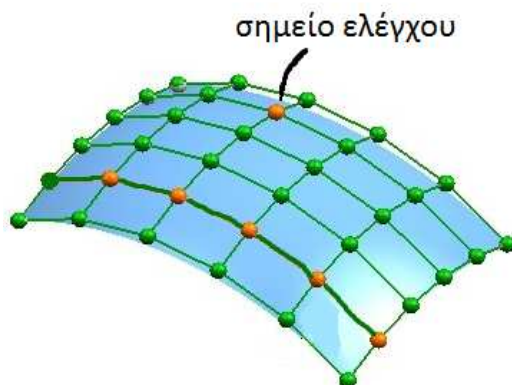
**N**

**near-infrared** : κατηγορία υπέρυθρης ακτινοβολίας με μήκος κύματος 0.75-1.4  $\mu\text{m}$ , που χαρακτηρίζεται από την απορρόφηση της από το νερό και χρησιμοποιείται κυρίως στις τηλεπικοινωνίες με οπτική ίνα. [26, (όρος infrared)]

**norm** : μία συνάρτηση που αναθέτει αυστηρά μια θετική τιμή σε όλα τα διανύσματα ενός διανυσματικού χώρου. Ένα πολύ απλό παράδειγμα είναι η ευκλείδεια νόρμα στον δισδιάστατο ευκλείδειο διανυσματικό χώρο  $\mathbb{R}^2$ . [26, (όρος Norm (mathematics))]

**Nunchuck** : Ο ελεγκτής Nunchuk είναι ένα περιφερειακό που συνδέεται με το Wii remote μέσω καλωδίου μήκους 1 - 1.2 μέτρων. Η εμφάνιση του όταν συνδέεται με αυτό μοιάζει με το κινέζικο όπλο nunchaku, από όπου προέκυψε και το όνομα. Διαθέτει έναν αναλογικό μοχλό και δύο κουμπιά στο μπροστινό του άκρο, C και Z που λειτουργούν σαν σκανδάλη. Λειτουργεί μαζί με τον κύριο ελεγκτή σε πολλά παιχνίδια. Όπως το Wii Remote έτσι και το Nunchuk διαθέτει επιταχυνσιόμετρο τριών αξόνων για την ανίχνευση επιτάχυνσης με 3 βαθμούς ελευθερίας και την αίσθηση κίνησης. Τα δεδομένα στέλνονται και λαμβάνονται από το Nunchuk γράφοντας σε και διαβάζοντας από κάποιο εύρος διευθύνσεων στη μνήμη flash του Wiimote. Το Wiimote με τη σειρά του αναμεταδίδει τα δεδομένα στον κεντρικό ελεγκτή μέσω Bluetooth. Δεν διαθέτει μικρόφωνο, λειτουργία δόνησης ή επίδειξης της κατεύθυνσης. [26, (όρος Wii Remote)]

**NURBS** : (Non-Uniform Rational B-Splines) μαθηματικές αναπαραστάσεις τρισδιάστατης γεωμετρίας που μπορούν να περιγράψουν με ακρίβεια οποιοδήποτε σχήμα, από μια απλή τρισδιάστατη γραμμή, έναν κύκλο, ένα τόξο ή μια καμπύλη, μέχρι την πιο πολύπλοκη τρισδιάστατη χωρίς ιδιαίτερη μορφή επιφάνεια ή στερεό. Είναι αρκετά ευέλικτα και ακριβή. [21]



Μία επιφάνεια σχεδιασμένη με την βοήθεια των NURBS

**Ο**

**OpenGL** : (OPEN Graphics Language), μία διεπαφή που τρέχει σε όλα τα λειτουργικά και αναπτύχθηκε για την συγγραφή εφαρμογών παραγωγής δισδιάστατων και τρισδιάστατων γραφικών. Αποτελείται από περισσότερες από 250 διαφορετικές κλήσεις συναρτήσεων, οι οποίες χρησιμοποιούνται για την σχεδίαση πολύπλοκων τρισδιάστατων σκηνών από απλά πρωταρχικά σχέδια. [26, (όρος OpenGL)]



Το λογότυπο της OpenGL

**OpenGL extensions** : προσφέρουν επιπλέον λειτουργικότητα, καθώς νέα τεχνολογία δημιουργείται συνεχώς. Μπορεί να εισάγουν νέες συναρτήσεις και νέες σταθερές ή να προσθέτουν ή να αφαιρούν περιορισμούς στις ήδη υπάρχουσες συναρτήσεις της OpenGL. [26, (όρος OpenGL)]

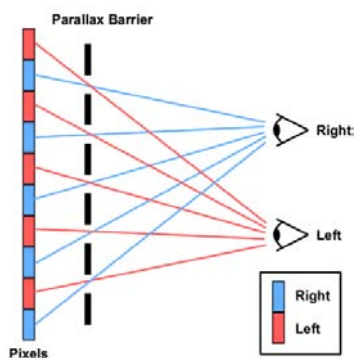
**OpenGL Extensions Wrangler** : (GLEW), μια βιβλιοθήκη της C/C++, για όλες τις πλατφόρμες, που βοηθά στην αναζήτηση και χρήση των πρόσθετων της OpenGL (OpenGL extensions). Παρέχει αποδοτικούς μηχανισμούς για τον καθορισμό των πρόσθετων που υποστηρίζονται στην πλατφόρμα που βρίσκεται κάθε φορά. [26, (όρος OpenGL Extension Wrangler Library)]

**optimization problem** : ο όρος χρησιμοποιείται στα μαθηματικά, στην επιστήμη των υπολογιστών και στα οικονομικά και αναφέρεται στην εύρεση του καλύτερου στοιχείου από ένα σύνολο διαθέσιμων στοιχείων. Στην απλούστερη περίπτωση, αυτό σημαίνει επίλυση προβλημάτων στα οποία αναζητούνται τα μέγιστα ή τα ελάχιστα μιας συνάρτησης. [26, (όρος optimization problem)]

**orthonormal matrix** : ένας πίνακας είναι ορθοκανονικός εάν τα στοιχεία του είναι διανύσματα ορθοκανονικά. Δύο διανύσματα είναι ορθοκανονικά εάν είναι ορθογώνια και μοναδιαίου μήκους. [26, (όρος orthonormality)]

**P**

**parallax barrier** : συσκευή που επιτρέπει σε μια οθόνη υγρών κρυστάλλων να απεικονίζει μια τρισδιάστατη εικόνα χωρίς να χρειάζεται ο θεατής να φορά γυαλιά. Τοποθετείται μπροστά από μία κανονική LCD και αποτελείται από ένα στρώμα υλικού με σχισμές, που επιτρέπει σε κάθε μάτι να βλέπει ένα διαφορετικό σύνολο από pixels ώστε να δίνει την αίσθηση του βάθους μέσω parallax. Δυστυχώς, ο θεατής θα πρέπει να τοποθετηθεί κατάλληλα ώστε να αποκτήσει τρισδιάστατη εμπειρία. [26, (όρος parallax barrier)]



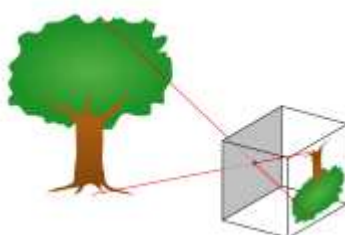
Λειτουργία barrier

Ένα κάλυμμα τοποθετείται μπροστά από την οθόνη LCD και κατευθύνει το φως προς κάθε μάτι μέσα από τις σχισμές (το ίδιο ζευγάρι κόκκινο-μπλε να περνά από την ίδια σχισμή, ενώ το επόμενο ζευγάρι περνά από την επόμενη σχισμή).

**percentage closer filtering (PCF)** : Παρόλο που είναι εξαιρετικά απλή και γενική, η χαρτογράφηση σκιών αντιμετωπίζει προβλήματα ταύτισης (aliasing), διότι οι πυκνότητες δειγματοληψίας στην οθόνη και στον χάρτη σκιών μπορεί να είναι πολύ διαφορετικές. Μια λύση είναι να αυξηθεί η αποτελεσματική ανάλυση του χάρτη σκιών. Μια απλούστερη εναλλακτική λύση είναι η χρήση της μεθόδου Percentage Closer Filtering (του Reeves. Υπολογίζει το ποσοστό της επιφάνειας που είναι πιο κοντά στο φως και όχι στην σκιά. [51])

**perspective division** : με την προοπτική προβολή, συνήθως αυτόματα, διαιρούνται οι συντεταγμένες του επιπέδου αποκοπής (οι αξίες  $x$ ,  $y$  και  $z$ ) με μια συντεταγμένη του επιπέδου  $w$ , προκειμένου να δημιουργηθούν κανονικοποιημένες συντεταγμένες. Με τον διαχωρισμό αυτό συρρικνώνεται η απομακρυσμένη γεωμετρία και επεκτείνεται η κοντινή. [26, (όρος Transformation matrix)]

**pinhole camera model** : είναι η κάμερα χωρίς φωτογραφικό φακό. Την θέση του φακού παίρνει μια μικρή τρύπα και χάρη στην ευθύγραμμη μετάδοση του φωτός, παίρνουμε στην πλάτη της κάμερας ένα αντεστραμμένο είδωλο. Αυτή η κάμερα δεν χρειάζεται να είναι τίποτε περισσότερο από ένα σκοτεινό κουτί, που να μπορεί να ανοίγει για να μπορούμε να τοποθετούμε μέσα του το φωτοευαίσθητο υλικό. [26, (όρος pinhole camera model)]



Μία pinhole κάμερα

Ακτίνες φωτός από ένα αντικείμενο περνούν από μια μικρή τρύπα για τον σχεδιασμό μιας εικόνας

## R

**range camera** : στην διαδικασία range imaging, ένα σύνολο από τεχνικές χρησιμοποιούνται για την παραγωγή μιας δισδιάστατης εικόνας (*range image*) που δείχνει την απόσταση στα σημεία μιας σκηνής από ένα συγκεκριμένο σημείο. Η συσκευή που χρησιμοποιείται για την παραγωγή της εικόνας λέγεται range camera. [26 (όρος range imaging)]

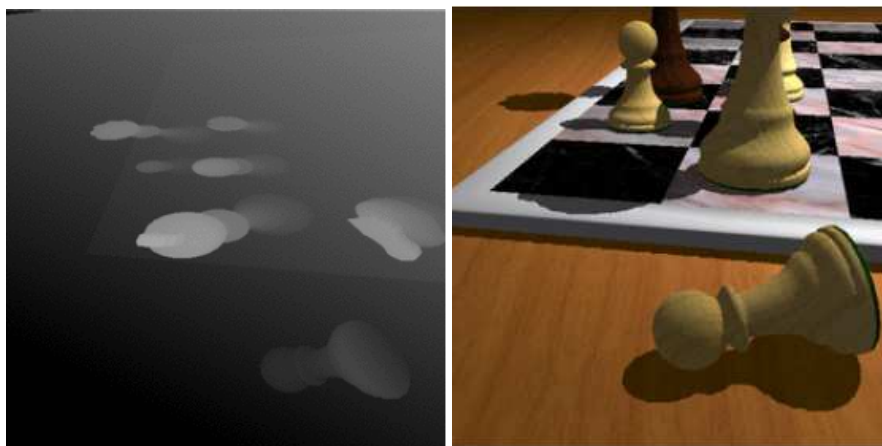
**red-eye effect** : η εμφάνιση κόκκινης κόρης του ματιού σε έγχρωμες φωτογραφίες. [26 (όρος red-eye effect)]

**retroreflector** : μία συσκευή ή επιφάνεια που ανακλά το φως πίσω στην πηγή του με μία ελάχιστη διάχυση του στον χώρο. Ένα ηλεκτρομαγνητικό κύμα ανακλάται πίσω κατά μήκος ενός διανύσματος που είναι παράλληλο αλλά αντίρροπο από το κύμα της πηγής. Η γωνία που δημιουργείται με το retroreflector είναι μεγαλύτερη του 0. [26 (όρος retroreflector)]

**S**

**scene graph** : είναι ένας τρόπος οργάνωσης των πληροφοριών με ιεραρχική δομή όπου οι κόμβοι γονιοί επηρεάζουν τους κόμβους παιδιά. Είναι, δηλαδή, μια δομή n-δέντρου, άλλα πιο εξελιγμένη. Μπορεί να έχει όσα παιδιά θέλει και κάθε κόμβος αντιπροσωπεύει ενέργειες που πρέπει να γίνουν. [19]

**shadow map** : ο χάρτης σκιάς είναι ένας αποδοτικός τρόπος παραγωγής ρεαλιστικών σκιών σε σκηνές που απαιτεί λογικές τιμές χώρου και χρόνου. Με αυτόν τον τρόπο προσθέτονται οι σκιές στην σκηνή εξετάζοντας κάθε pixel και εάν αυτό βρίσκεται σε θέση που είναι ορατή από την πηγή φωτός. [52]



Ένας τυπικός χάρτης σκιών και η αντίστοιχη σκηνή

**skype** : Το Skype είναι μια εξαιρετικά δημοφιλής εφαρμογή VoIP με εκατομμύρια χρήστες από όλον τον κόσμο. Αρχικά ήταν για επικοινωνία από Η/Υ σε Η/Υ. Πλέον προσφέρει κλήσεις σε οποιοδήποτε μέρος του κόσμου, σε οποιοδήποτε δίκτυο τηλεφωνίας, σταθερής και κινητής, με χαμηλές χρεώσεις, δυνατότητα αποστολής SMS, αποστολής αρχείων και δυνατότητα συνδιάσκεψης. Διατίθεται σε εκδόσεις για Linux, Mac, Pocket PC και Windows. Η σύνδεση που παρέχει μεταξύ των χρηστών γίνεται μέσω Peer-to-Peer άμεσων συνδέσεων, χωρίς την διαμεσολάβηση διακομιστών. [26 (όρος Skype)]

**smoothing techniques** : η συλλογή πληροφοριών πολλές φορές δίνει τελικά αποτελέσματα τυχαίων διακυμάνσεων. Η μείωση ή η εξάλειψη αυτών των περιορισμών πετυχαίνεται με τις τεχνικές smoothing. Όταν οι τεχνικές αυτές εφαρμοστούν σωστά δίνουν αποτελέσματα πιο κοντά στην πραγματικότητα. [53]

**spline καμπύλες** : οι καμπύλες που προκύπτουν από την ένωση επιμέρους καμπυλών με κατάλληλες συνθήκες συνέχειας. Υπάρχουν οι καμπύλες  $\nu$ -,  $\beta$ - και  $\gamma$ - splines. [26, (όρος Spline (mathematics))]

**stereo matching** : το πιο δύσκολο πρόβλημα στην στερεοσκοπική ανάλυση είναι να καθορίσουμε ποια pixels της αριστερής και δεξιάς εικόνας αντιστοιχούν στο ίδιο σημείο στον τρισδιάστατο χώρο. Γνωστό και σαν πρόβλημα αντιστοίχισης. [55]

**SVD ανάλυση** : (Singular Value Decomposition), τεχνική μείωσης διαστάσεων. Στην γραμμική άλγεβρα χρησιμοποιείται για την παραγοντοποίηση ενός ορθογωνίου πίνακα και βρίσκει πολλές εφαρμογές στην επεξεργασία σημάτων και στην στατιστική. Μερικές από τις εφαρμογές της SVD ανάλυσης είναι ο υπολογισμός του ψευδός αντίστροφου και ο καθορισμός της τάξης ενός πίνακα. [26, (όρος Singular value decomposition)]

## T

**Tokyo Game Show** : είναι γνωστό και σαν TGS. Πρόκειται για μία έκθεση βιντεοπαιχνιδιών που διεξάγεται ετησίως στην Chiba της Ιαπωνίας. Συμμετέχουν πολλοί υπεύθυνοι πάνω στην ανάπτυξη βιντεοπαιχνιδιών από όλο τον κόσμο για να επιδείξουν τα επερχόμενα παιχνίδια και το υλικό πίσω από αυτά. Κυρίως εστιάζουν στην ιαπωνική αγορά παρά στο εξωτερικό. Σε αντίθεση με την Ηλεκτρονική έκθεση ψυχαγωγίας, τις τελευταίες δύο μέρες επιτρέπεται η πρόσβαση στο ευρύ κοινό. Για αυτό και έχει μεγαλύτερη συμμετοχή από την E3. [26, (όρος Tokyo Game Show)]

**type casting** : μετατροπή ενός χαρακτηριστικού συγκεκριμένου τύπου σε άλλον τύπο που εμείς επιθυμούμε. [26, (όρος Type conversion)]

## V

**volumetric light** : είδος φωτός που χρησιμοποιείται στα τρισδιάστατα γραφικά. Επιτρέπει στον χρήστη να βλέπει ακτίνες φωτός να διαχέονται στο περιβάλλον. Το φως παράγεται από μία πηγή φωτός, μοντελοποιείται ως αδιαφανές αντικείμενο και θεωρείται ως κάτι που έχει μεγάλο όγκο. Περνώντας μέσα του το φως διακυμαίνεται



ανάλογα με τον όγκο κάθε φορά (όπως η ομίχλη, ο καπνός ή η σκόνη), όπως και στον πραγματικό κόσμο. [26, (όρος Volumetric lighting)]

## W

**WiFi** : Ο όρος WiFi (Wireless Fidelity, κατά την ορολογία High Fidelity η οποία αφορά την εγγραφή ήχου) χρησιμοποιείται για να προσδιορίσει τις συσκευές που βασίζονται στην προδιαγραφή IEEE 802.11 b/g και εκπέμπουν σε συχνότητες 2.4GHz. Το IEEE 802.11 είναι μια οικογένεια προτύπων της IEEE για ασύρματα τοπικά δίκτυα (WLAN) που είχαν ως σκοπό να επεκτείνουν το 802.3 (Ethernet, το συνηθέστερο πρωτόκολλο ενσύρματης δικτύωσης υπολογιστών) στην ασύρματη περιοχή. Τα πρότυπα 802.11 είναι ευρύτερα γνωστά ως "WiFi" επειδή η WiFi Alliance, ένας οργανισμός ανεξάρτητος της IEEE, παρέχει την πιστοποίηση για τα προϊόντα που υπακούν στις προδιαγραφές του 802.11. Αυτή η οικογένεια πρωτοκόλλων αποτελεί το καθιερωμένο πρότυπο της βιομηχανίας στο χώρο των ασύρματων τοπικών δικτύων. Ωστόσο το WiFi ("ασύρματη πιστότητα" στα ελληνικά) έχει επικρατήσει και ως όρος αναφερόμενος συνολικά στα ασύρματα τοπικά δίκτυα. Συνήθεις εφαρμογές του είναι η παροχή ασύρματων δυνατοτήτων πρόσβασης στο Internet, τηλεφωνίας μέσω διαδικτύου (VoIP) και διασύνδεσης μεταξύ ηλεκτρονικών συσκευών όπως τηλεοράσεις, ψηφιακές κάμερες, DVD Player και ηλεκτρονικοί υπολογιστές. Σε φορητές ηλεκτρονικές συσκευές το 802.11 βρίσκει εφαρμογές ασύρματης μετάδοσης, όπως π.χ. στη μεταφορά φωτογραφιών από ψηφιακές κάμερες σε υπολογιστές για περαιτέρω επεξεργασία και εκτύπωση, αν και σε αυτόν τον τομέα έχει υποσκελιστεί από το πρωτόκολλο Bluetooth για τα πολύ μικρότερης εμβέλειας ασύρματα προσωπικά δίκτυα. [26, (όρος Wi-Fi)]

**Windows Live Messenger** : Το Windows Live Messenger (πρώην MSN Messenger) είναι λογισμικό instant messaging από την Microsoft. Λειτουργεί στα Windows XP (μόνο 32-bit), Windows Vista, Windows 7 και Windows Mobile. Επιτρέπει την επικοινωνία μέσω ανταλλαγής μηνυμάτων, φωτογραφιών, αρχείων δεδομένων και τη συνομιλία μέσω βίντεο. Αποτελεί κομμάτι των υπηρεσιών Windows Live. [26, (όρος Windows Live Messenger)]

**wire frame model** : Είναι η οπτική απεικόνιση ενός τρισδιάστατου ή φυσικού αντικειμένου με την βοήθεια γραφικών στον υπολογιστή. [26, (όρος Wire-frame model)]



Μορφή πλέγματος για έναν κύβο, ένα πολύγωνο και μια σφαίρα

**wrapper function** : συνάρτηση σε ένα πρόγραμμα που έχει ως κύριο στόχο την κλήση μιας άλλης συνάρτησης χωρίς κάποιο ιδιαίτερο υπολογισμό. [26, (όρος wrapper function)]

## V

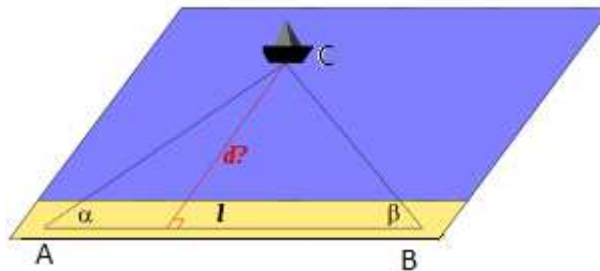
**VGA** : (Video Graphics Array), αποτελεί ένα πολύ διαδεδομένο αναλογικό πρότυπο απεικόνισης σε οθόνες ηλεκτρονικών υπολογιστών, το οποίο εμφανίστηκε για πρώτη φορά το 1987, από την IBM. Συνήθως, ο όρος VGA αναφέρεται στην ανάλυση 640×480 εικονοστοιχείων (pixels). Το VGA διαδέχτηκε το XGA (1024x768 pixels) της IBM, και από πολλές ακόμη επεκτάσεις. [26, (όρος Video Graphics Array)]

## A - Ω

**ποτενσιόμετρο** : όργανο μέτρησης του δυναμικού (ή της τάσης) σε ένα κύκλωμα. Το όργανο χρησιμοποιεί ένα φράγμα (μια γνωστή τάση μιας αντίστασης) και το συγκρίνει με την άγνωστη τάση, με την βοήθεια ενός γαλβανόμετρου (όργανο για την ανίχνευση και την μέτρηση ηλεκτρικού φορτίου).

**τεχνική πλησιέστερου γείτονα** : σύμφωνα με την τεχνική αυτή κατά την δειγματοληψία από όλα τα δείγματα (συμπεριλαμβανομένου και του αρχικού) επιλέγεται ως επόμενο αυτό που είναι πλησιέστερο στο προηγούμενο. Η διαδικασία ολοκληρώνεται όταν έχουν εξεταστεί όλα τα δείγματα και επιστρέφουμε στο αρχικό σημείο.

**τριγωνοποίηση** : η διαδικασία του προσδιορισμού της θέσης ενός σημείου με τη μέτρηση των γωνιών που σχηματίζονται ανάμεσα σε αυτό και γνωστά σημεία σε κάθε άκρο μιας σταθερής γραμμής αναφοράς, αντί για τη μέτρηση των αποστάσεων από το σημείο άμεσα. Το σημείο μπορεί στη συνέχεια να θεωρηθεί ως το τρίτο σημείο ενός τριγώνου με μια γνωστή πλευρά και δύο γνωστές γωνίες. [26 (όρος triangulation)]



Η τριγωνοποίηση μπορεί να χρησιμοποιηθεί για τον υπολογισμό των συντεταγμένων και της απόστασης από την ακτή στο πλοίο. Ο παρατηρητής A μετρά τη γωνία  $\alpha$  ανάμεσα στο πλοίο και την ακτή ενώ ο παρατηρητής B κάνει το ίδιο για τη γωνία  $\beta$ . Με γνωστό το μήκος  $l$  ή τις συντεταγμένες των A και B μπορεί να εφαρμοστεί ο νόμος των ημιτόνων για να βρεθούν οι συντεταγμένες του πλοίου στο C και η απόσταση  $d$ .

### 1 - 9

**16-tap disk sampling PCF** : Παραλλαγή της τεχνικής PRCF. Η ιδέα εδώ είναι να ενσωματωθούν τα αποτελέσματα των ελέγχων σκιάς σε μια γειτονιά του χάρτη σκιών. Η διαδικασία διενεργείται στοχαστικά, με ένα πρότυπο δειγματοληψίας δίσκου Poisson, το οποίο μετατρέπει την ταύτιση σε θόρυβο υψηλής συχνότητας. Ο θόρυβος γίνεται ελάχιστα ορατός όταν βρεθεί ο μέσος όρος από 16 taps στο χάρτη σκιών. [6]