

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ**

**ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ**

**ΤΟΜΕΑΣ:** Ηλεκτρονικής και Υπολογιστών (Η.Υ.)

**ΕΡΓΑΣΤΗΡΙΟ:** Διαδραστικών Τεχνολογιών

**«Σχεδίαση εξωτερικού ελεγκτή για τον έλεγχο περιβάλλοντος μικτής πραγματικότητας»**

**«Design of external controller for interaction with mixed reality environments»**

**Διπλωματική Εργασία**

του

Παπαδούλη Γεωργίου

Αριθμός Μητρώου: 1020865

Αριθμός Διπλωματικής Εργασίας: (1020865 / 2022)

Επιβλέπων : Νικόλαος Αβούρης

Πάτρα, Σεπτέμβριος 2022

**ΠΙΣΤΟΠΟΙΗΣΗ**

Πιστοποιείται ότι η Διπλωματική Εργασία με θέμα

**«Σχεδίαση εξωτερικού ελεγκτή για τον έλεγχο περιβάλλοντος μικτής πραγματικότητας»**

**«Design of external controller for interaction with mixed reality environments»**

Του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών

Παπαδούλη Γεωργίου

Αριθμός Μητρώου: 1020865

Παρουσιάστηκε δημόσια και εξετάστηκε στο Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών στις

…….. / …….. / ……..

Ο Επιβλέπων Ο φοιτητής Ο Διευθυντής του Τομέα

Αβούρης Νικόλαος Παπαδούλης Γεώργιος Παλιουράς Βασίλειος

Καθηγητής Καθηγητής

# Ευχαριστίες

# Περίληψη

# Abstract

# Περιεχόμενα

[Ευχαριστίες I](#_Toc115044441)

[Περίληψη II](#_Toc115044442)

[Abstract III](#_Toc115044443)

[Περιεχόμενα IV](#_Toc115044444)

[Κατάλογος Σχημάτων V](#_Toc115044445)

[Πρόλογος - 1 -](#_Toc115044446)

[Κεφάλαιο 1 – Θεωρητικό & Τεχνολογικό Υπόβαθρο - 2 -](#_Toc115044447)

[1.1 Εκτεταμένη Πραγματικότητα - 2 -](#_Toc115044448)

[1.1.1 Εικονική Πραγματικότητα (*VR*) - 3 -](#_Toc115044449)

[1.1.2 Επαυξημένη Πραγματικότητα (AR) - 3 -](#_Toc115044450)

[1.1.3 Μικτή Πραγματικότητα (MR) - 4 -](#_Toc115044451)

[1.2 Γραφικά και αναπαράσταση στο χώρο[20], [21] - 5 -](#_Toc115044452)

[1.2.1 Γραφικά και τρισδιάστατος χώρος [20], [21] - 5 -](#_Toc115044453)

[1.2.2 Μεταφορά (*Translation*) - 5 -](#_Toc115044454)

[1.2.3 Περιστροφή γύρω από άξονα (*Rotation*) - 6 -](#_Toc115044455)

[1.2.4 Γενικευμένη Περιστροφή - 7 -](#_Toc115044456)

[1.2.5 Αλλαγή μεγέθους (*Scale*) - 7 -](#_Toc115044457)

[1.2.6 Συνδυασμός μεταφοράς, περιστροφής και αλλαγής μεγέθους - 8 -](#_Toc115044458)

[1.2.7 Τετραδόνιο (*Quaternion)* [22], [23] - 10 -](#_Toc115044459)

[1.3 Επεξεργασία εικόνας [24] - 12 -](#_Toc115044460)

[1.3.1 Ορισμός-σύντομη περιγραφή - 13 -](#_Toc115044461)

[1.3.2 Ανίχνευση ακμών [25] - 13 -](#_Toc115044462)

[1.3.3 Υπολογιστική Όραση - 14 -](#_Toc115044463)

[1.4 Οπτική Αδρανειακή Οδομετρία – Ταυτόχρονος Εντοπισμός και Χαρτογράφηση - 16 -](#_Toc115044464)

[1.4.1 Οδομετρία (*Odometry*) - 16 -](#_Toc115044465)

[1.4.2 Οπτική Οδομετρία (*Visual Odometry*) [27]–[29] - 16 -](#_Toc115044466)

[1.4.3 Αδρανειακή Οδομετρία (Inertial Odometry) [32], [33] - 17 -](#_Toc115044467)

[1.4.4 Οπτική Αδρανειακή Οδομετρία (*Visual Inertial Odometry*) [27] - 19 -](#_Toc115044468)

[1.4.5 Ταυτόχρονος Εντοπισμός και Χαρτογράφηση (*Simultaneous Localization and Mapping*) – *SLAM* [34], [35] - 20 -](#_Toc115044469)

[1.4.5 Εφαρμογές - 21 -](#_Toc115044470)

[1.5 Microsoft Hololens 2 [37] - 22 -](#_Toc115044471)

[1.5.1 Περιγραφή συσκευής - 22 -](#_Toc115044472)

[1.5.2 Βασικά εξαρτήματα - 23 -](#_Toc115044473)

[1.5.3 Αισθητήρες - 23 -](#_Toc115044474)

[*1.5.4* Τρόποι αλληλεπίδρασης – ελέγχου του *HoloLens 2* [38] - 25 -](#_Toc115044475)

[1.6 Πλατφόρμα ανάπτυξης Unity [46] - 30 -](#_Toc115044476)

[1.6.1 Βασικά στοιχεία - 30 -](#_Toc115044477)

[1.6.2 Κύκλος ζωής μιας εφαρμογής Unity [47] - 31 -](#_Toc115044478)

[- 33 -](#_Toc115044479)

[1.6.3 Βασικό Περιβάλλον Ανάπτυξης (Editor) - 33 -](#_Toc115044480)

[1.6.4 Plugins - 35 -](#_Toc115044481)

[1.6.4 *AR Foundation* [48] - 35 -](#_Toc115044482)

[1.6.5 MRTK (Mixed Reality Toolkit) [49] - 37 -](#_Toc115044483)

[Βιβλιογραφία - 40 -](#_Toc115044484)

# Κατάλογος Σχημάτων

[ΕΙΚΟΝΑ 1 Πραγματικό - εικονικο συνεχες - milgram et al.'s continuum - 2 -](#_Toc115020274)

[ΕΙΚΟΝΑ 2 Περιστροφη κατα τον αξονα Z - 6 -](#_Toc115020275)

[ΕΙΚΟΝΑ 3 Διάγραμμα Ροήσ Ενός Γενικού Συστήματος Επεξεργασίας Εικόνας - 12 -](#_Toc115020276)

[ΕΙΚΟΝΑ 4 Πρωτότυπη Εικόνα - 14 -](file:///C:\Users\salok\Documents\GitHub\Hololens-Development\Hololens\Dev%20Log\Thesis\Thesis_Papadoulis.docx#_Toc115020277)

[ΕΙΚΟΝΑ 5 Αποτέλεσμα Ανίχνευσης - 14 -](file:///C:\Users\salok\Documents\GitHub\Hololens-Development\Hololens\Dev%20Log\Thesis\Thesis_Papadoulis.docx#_Toc115020278)

[ΕΙΚΟΝΑ 6 To Πλεγμα που Δημιουργειται μετα τη σαρωση του περιβαλλοντοσ [26] - 15 -](#_Toc115020279)

[ΕΙΚΟΝΑ 7 Stable Platform IMU [31] - 18 -](file:///C:\Users\salok\Documents\GitHub\Hololens-Development\Hololens\Dev%20Log\Thesis\Thesis_Papadoulis.docx#_Toc115020280)

[ΕΙΚΟΝΑ 8 Σύστημα Αδρανειακής Πλοήγησής [33] - 19 -](#_Toc115020281)

[ΕΙΚΟΝΑ 9 Block Διάγραμμα Ενόσ Συστηματοσ vSLAM [35] - 20 -](#_Toc115020282)

[ΕΙΚΟΝΑ 10 Block Διάγραμμα για (α) Vo kai (B) SLAM - 21 -](#_Toc115020283)

[ΕΙΚΟΝΑ 11 Αποδόμηση της Συσκευής - 22 -](#_Toc115020284)

[ΕΙΚΟΝΑ 12 Κύριοι Αισθητήρες του Hololens 2 - 24 -](#_Toc115020285)

[ΕΙΚΟΝΑ 13 Σκελετός και σημεία ενδιαφέροντος του χεριού - 26 -](file:///C:\Users\salok\Documents\GitHub\Hololens-Development\Hololens\Dev%20Log\Thesis\Thesis_Papadoulis.docx#_Toc115020286)

[ΕΙΚΟΝΑ 14 Πλαίσιο Ανίχνευσης Χεριού [42] - 26 -](#_Toc115020287)

[ΕΙΚΟΝΑ 15 Θέσεις ετοιμότητας και θέση tap - 27 -](#_Toc115020288)

[ΕΙΚΟΝΑ 16 Start and One Handed Start gestures - 28 -](file:///C:\Users\salok\Documents\GitHub\Hololens-Development\Hololens\Dev%20Log\Thesis\Thesis_Papadoulis.docx#_Toc115020289)

[ΕΙΚΟΝΑ 17 Χειρισμός από κοντά - 28 -](#_Toc115020290)

[ΕΙΚΟΝΑ 18 Κατάσταση Επικύρωσης (Commit State) - 29 -](file:///C:\Users\salok\Documents\GitHub\Hololens-Development\Hololens\Dev%20Log\Thesis\Thesis_Papadoulis.docx#_Toc115020291)

[ΕΙΚΟΝΑ 19 Ελεύθερη Κατάσταση Δείκτη (Pointing State) - 29 -](file:///C:\Users\salok\Documents\GitHub\Hololens-Development\Hololens\Dev%20Log\Thesis\Thesis_Papadoulis.docx#_Toc115020292)

# Πρόλογος

Η εικονική, επαυξημένη και εκτεταμένη πραγματικότητα είναι πεδία τα οποία τα τελευταία χρόνια αναπτύσσονται με γοργούς ρυθμούς. Έχουν σχεδιασθεί και αναπτυχθεί διάφορες τεχνολογίες που υλοποιούν τα παραπάνω πεδία, από *HMDs* που τοποθετούνται στο κεφάλι στα πιο απλά καθημερινά έξυπνα κινητά και τάμπλετ. Μάλιστα, η βελτίωση είναι εμφανής και από τη μεριά της προσβασιμότητας και της ταχύτητας, καθώς συνεχώς γίνεται φθηνότερη η παραγωγή καλύτερου υλικού και αναπτύσσονται περισσότερο αποδοτικοί αλγόριθμοι και λογισμικό.

Παρατηρούμε λοιπόν τη χρήση αυτών των τεχνολογιών σε διάφορους τομείς. Εφαρμογές εκπαίδευσης, προσομοίωσης, περιήγησης σε περιβάλλοντα ενδιαφέροντος – λόγου χάριν, πολιτισμικά – καθώς και βιντεοπαιχνίδια είναι κάποια παραδείγματα όπου η εικονική, επαυξημένη και εκτεταμένη πραγματικότητα προσφέρουν μια πιο ολοκληρωμένη, πιο κοντά στον άνθρωπο εμπειρία.

Συγκεκριμένα, η παρούσα διπλωματική ασχολείται περισσότερο με το πεδίο της εκτεταμένης πραγματικότητας, όπου ο χρήστης καλείται να περιηγηθεί και αλληλοεπιδράσει με ένα εικονικό περιβάλλον το οποίο όμως συνυπάρχει με τον πραγματικό κόσμο.

Ένα εύλογο ερώτημα που προκύπτει είναι το εξής: «Τι τρόπους / μηχανισμούς αλληλεπίδρασης θα προσφέρω στο χρήστη μίας συσκευής εκτεταμένης πραγματικότητας;» Συνηθέστερα, η αλληλεπίδραση γίνεται είτε με κάποιον ενσωματωμένο ελεγκτή – φωνητική εντολή, αναγνώριση χειρονομιών, εντοπισμός χεριού, εντοπισμός βλέμματος, εντοπισμός κεφαλής – είτε με κάποιο εξωτερικό ελεγκτή / χειριστήριο. Η πρώτη περίπτωση έχει το πλεονέκτημα λιγότερου υλικού, άρα και φθηνότερης υλοποίησης, με περιορισμούς το πεδίο όρασης της συσκευής και την υπολογιστική της ισχύ. Η δεύτερη περίπτωση αυξάνει το κόστος, αλλά προσφέρει έναν επιπλέον μηχανισμό ο οποίος επεκτείνει τις δυνατότητες αλληλεπίδρασης του χρήστη, καθώς και τη συνολική ευρωστία του συστήματος.

Ως μια λύση του υψηλότερου κόστους της χρήσης εξωτερικού ελεγκτή, καθώς και μία συσκευή πειραματισμού για διάφορους τρόπους ελέγχου της εκτεταμένης πραγματικότητας, προτείνεται το κινητό τηλέφωνο που χρησιμοποιούμε καθημερινά. Πλέον, τα κινητά ή έξυπνα τηλέφωνα αποτελούν ισχυρούς υπολογιστές και προσφέρουν πολλές δυνατότητες επέκτασης της αλληλεπίδρασης με ένα εξωτερικό σύστημα.

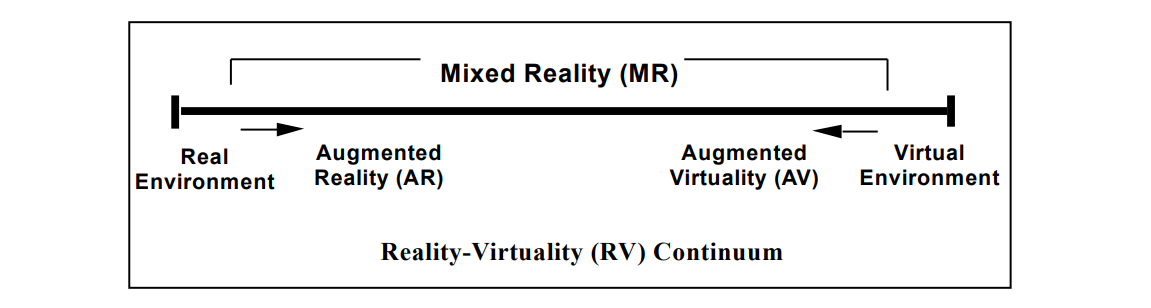
Αυτός είναι και ο σκοπός αυτής της διπλωματικής. Το *HMD* που χρησιμοποιείται στη διπλωματική είναι το *Microsoft Hololens 2* · αν και ο κώδικας επί το πλείστον λειτουργεί και στο *Hololens 1*. Παρουσιάζεται λοιπόν μία προσπάθεια επέκτασης του χειρισμού και της αλληλεπίδρασης που προσφέρει το *Hololens* χρησιμοποιώντας ως εξωτερικό χειριστήριο ένα καθημερινό έξυπνο τηλέφωνο. Επιπλέον, θα εξετάσουμε κατά πόσο είναι δυνατή η μετατροπή του κινητού σε ένα *6DOF Controller*, καθώς και τι άλλες δυνατότητες μπορεί να προσφέρει το κινητό μέσω της οθόνης αφής.

Τέλος, μέσω ενός απλού πειράματος θα συγκρίνουμε μία αλληλεπίδραση που εισαγάγαμε χρησιμοποιώντας την αναπτυχθείσα βιβλιοθήκη (*framework*) με τον εσωτερικό (*native*) τρόπο αλληλεπίδρασης.

# Κεφάλαιο 1 – Θεωρητικό & Τεχνολογικό Υπόβαθρο

Στο κεφάλαιο αυτό παρουσιάζονται τα διάφορα θεωρητικά και τεχνολογικά εργαλεία που απαιτούνται για την κατανόηση και ολοκλήρωση του σκοπού της διπλωματικής εργασίας. Καταρχάς, γίνεται μια αναφορά στα πεδία της εικονικής, επαυξημένης και εκτεταμένης πραγματικότητας, καθώς και συσκευές που χρησιμοποιούνται για την υλοποίηση τους. Ακολουθεί συνοπτική ανάπτυξη χρήσιμων μαθηματικών περιγραφής τρισδιάστατου χώρου, μία εποπτική περιήγηση στην Επεξεργασία Εικόνας καθώς και επεξηγούνται οι όροι *Visual Inertial Odometry* – *Simultaneous Localization and Mapping*, βασικοί όροι για την κατανόηση της εσωτερικής λειτουργίας μιας εκ των αναφερθέντων συσκευών. Στη συνέχεια, παρουσιάζεται το *Hololens 2* και καταγράφονται τα διάφορα είδη ελεγκτών / χειριστηρίων για τη συσκευή αυτή και ο τρόπος αλληλεπίδρασης που προσφέρουν. Τέλος, αναφέρονται τα διάφορα περιβάλλοντα ανάπτυξης της λύσης που προτείνεται στη διπλωματική ·To (Η) Unity και το Visual Studio.

## Εκτεταμένη Πραγματικότητα



ΕΙΚΟΝΑ 1 Πραγματικό - εικονικο συνεχες - milgram et al.'s continuum

Η εκτεταμένη πραγματικότητα αποτελεί μία έννοια που περικλείει την εικονική, την επαυξημένη και τη μικτή πραγματικότητα, καθώς και των ενδιάμεσων αυτών περιοχών. Εντούτοις, αποτελεί ένα υπερσύνολο των παραπάνω, με όρια από το «πλήρως αληθινό» στο «πλήρως εικονικό», όπως προτάθηκε από τον Paul Milgram κατά τα πλαίσια του «πραγματικού-εικονικού συνεχούς» (*Reality-Virtuality Continuum*) [1], [2] (Εικ. 1).

Η ενότητα αυτή εστιάζει στα υποσύνολα της εκτεταμένης πραγματικότητας – εικονική, επαυξημένη και μικτή - καθώς και στην καλύτερη αποσαφήνιση μεταξύ της επαυξημένης και μικτής πραγματικότητας, έννοιες που πολλές φορές χρησιμοποιούνται ως συνώνυμα.

### Εικονική Πραγματικότητα (*VR*)

Η εικονική πραγματικότητα , σε αντίθεση με την επαυξημένη και τη μικτή είναι καλώς ορισμένη. Ως εικονική πραγματικότητα ορίζουμε «μία προσομοιωμένη εμπειρία ή οποία μπορεί να είναι παρόμοια ή τελείως διαφορετική σε σύγκριση με τον πραγματικό κόσμο» [3]. Χαρακτηριστικά της, μεταξύ άλλων, είναι η πλήρης εμβύθιση στον εικονικό κόσμο καθώς και η περιήγηση απομακρυσμένων περιοχών σε σχέση πάντα με την τοποθεσία του χρήστη. Επιπλέον, πολλοί υποστηρίζουν πως η χρήση ενός *HMD* είναι αναγκαία ώστε να θεωρηθεί η εμπειρία εικονική, έναντι ενός απλού τρισδιάστατου βίντεο [4]. Μία συσκευή *VR* είναι το *Oculus Rift*.

Η εικονική πραγματικότητα έχει εφαρμογή σε διάφορους τομείς:

* Στην εκπαίδευση ιατρικού προσωπικού και τη βοήθεια επανένταξης ασθενών, μέσω κατάλληλα σχεδιασμένων προσομοιώσεων. Επιπλέον, στην επέκταση της διαδικασίας της διάγνωσης ενός ασθενούς. [5], [6]
* Στη διασκέδαση και στην ψυχαγωγία, μέσω διάφορων βιντεοπαιχνιδιών. [7], [8]
* Στην εκπαίδευση ως εναλλακτική τεχνική εκμάθησης ή βελτίωση υπαρχόντων τεχνικών. [9]

### Επαυξημένη Πραγματικότητα (AR)

Η επαυξημένη πραγματικότητα διαφέρει αισθητά από την εικονική. Ο χρήστης πλέον δεν είναι πλήρως εμβυθισμένος σε μια εικονική εμπειρία. Στα μάτια του συνεχίζει να αποτυπώνεται ο πραγματικός κόσμος, στον οποίο όμως ταυτόχρονα υπερτίθενται εικονικά αντικείμενα που διατηρούν τη θέση και τον προσανατολισμό τους στο χώρο. [4], [10] Τα αντικείμενα αυτά, όμως, δεν αλληλοεπιδρούν με τον πραγματικό κόσμο, απλά προβάλλονται πάνω σε αυτόν. Βέβαια, όπως προαναφέρθηκε, άλλες πηγές θεωρούν αναγκαία την αλληλεπίδρασή των αντικειμένων με τον πραγματικό κόσμο, γεγονός που προκαλεί σύγχυση μεταξύ των εννοιών της επαυξημένης και μικτής πραγματικότητας. [4] Μια AR συσκευή πλέον είναι οποιοδήποτε σύγχρονο έξυπνο κινητό.

Η επαυξημένη πραγματικότητα με τη σειρά της βρίσκει εφαρμογή σε διάφορους τομείς:

* Στον κόσμο των βιντεοπαιχνιδιών και της ψυχαγωγίας. Μία πολύ γνωστή εφαρμογή επαυξημένης πραγματικότητας είναι το παιχνίδι Pokémon Go. [11]
* Στον τομέα της αρχιτεκτονικής, του εσωτερικού σχεδιασμού και της διακόσμησης, όπου οι χρήστες μπορούν να προβάλλουν εικονικά αντικείμενα στο χώρο τους για να εκτιμήσουν καλύτερα μια πιθανή αγορά. [12], [13]
* Στον τομέα της πολιτισμικής κληρονομιάς, όπου συνεχώς αναζητούνται νέοι τρόποι προσέγγισης και βελτίωσης της εμπειρίας ενός επισκέπτη. [14]

### 1.1.3 Μικτή Πραγματικότητα (MR)

Η μικτή πραγματικότητα δεν είναι καλώς ορισμένη, γεγονός που γίνεται εύκολα αντιληπτό αν παρατηρήσουμε διαφορετικούς ορισμούς που έχουν δώσει ειδικοί όταν ερωτήθηκαν για αυτό. Συγκεκριμένα, παρατηρούνται 4 διαφορετικοί ορισμοί [4]:

* Η μικτή πραγματικότητα βάσει του *RV Continuum* των *Milgram et al.,* όπως αυτό παρουσιάζεται στην Εικ. 1. Σε αυτήν την περίπτωση, η μεικτή πραγματικότητα δεν είναι αναγκαίο να περιλαμβάνει την εικονική.
* Η μικτή πραγματικότητα ως συνδυασμός της εικονικής και της επαυξημένης. Η δυνατότητα δηλαδή συνδυασμού των δύο αυτών τεχνολογιών σε μία εφαρμογή ή συσκευή.
* Η μικτή πραγματικότητα ως επέκταση της επαυξημένης. Εδώ ορίζουμε τη μικτή ως μια ικανότερη εκδοχή της επαυξημένης · Λόγου χάρη, καλύτερη κατανόηση του φυσικού περιβάλλοντος στο οποίο βρίσκεται ο χρήστης.
* Η μικτή και η επαυξημένη πραγματικότητα ως συνώνυμα.

Λόγω αυτής της έλλειψης ενός καθολικά αποδεκτού ορισμού, πολλές εφαρμογές μπορούν να κατηγοριοποιηθούν και ως *AR* και ως *MR*. Στα πλαίσια αυτής της διπλωματικής θα αντιμετωπίσουμε τη μικτή πραγματικότητα ως την καλύτερη δυνατή ανίχνευση του περιβάλλοντος με σκοπό την προβολή ενός εικονικού κόσμου ο οποίος αλληλοεπιδρά με τον πραγματικό με μεγάλη ακρίβεια και σε πραγματικό χρόνο. Επιπλέον, θεωρείται αναγκαία η εμβύθιση του χρήστη στον κόσμο, μέσω παραδείγματος χάριν ενός *HMD*, έναντι μίας συσκευής *AR*, όπως τα έξυπνα κινητά.

Παραδείγματα συσκευών μικτής πραγματικότητας είναι τα Microsoft Hololens (1, 2), Magic Leap (1, 2)

Με βάση τα παραπάνω, έχουμε διάφορες εφαρμογές της μικτής πραγματικότητας. Παρακάτω αναφέρονται μερικές από αυτές, οι οποίες επιλέχθηκαν ώστε να ικανοποιούν τον παραπάνω ορισμό:

* Η χρήση της μικτής πραγματικότητας στον τομέα της ιατρικής (τηλε-ιατρική) για παροχή διάγνωσης και περίθαλψης ασθενών όταν χρειάζεται ομάδα ανθρώπων. Η εφαρμογή αυτή βρήκε χρήση ειδικά κατά την περίοδο του COVID-19, όπου ο επιβαλλόταν η αποφυγή του συνωστισμού ατόμων σε ένα μικρό χώρο αλλά ταυτόχρονα ήταν απαραίτητη η συνεργασία του ιατρικού προσωπικού [15]
* Στην αρχιτεκτονική και στην κατασκευαστική βιομηχανία, ειδικότερα σε περιπτώσεις συνεργασίας. Μια ομάδα ανθρώπων μπορεί μέσω διαφορετικών *HMDs* να συζητήσει / αναπτύξει / σχεδιάσει μία κατασκευή μέσω ενός εικονικού μοντέλου που τοποθετούν στον πραγματικό κόσμο. Επιπλέον, η μικτή πραγματικότητα είναι χρήσιμη στην εκπαίδευση νέου εργατικού δυναμικού. [16], [17]
* Σε πολιτισμικά θέματα. Η μικτή πραγματικότητα αποτελεί ένα πολύ δυνατό εργαλείο ώστε να βελτιωθεί και επαυξηθεί η εμπειρία περιήγησης ενός μουσείο ή γενικότερα πολιτισμικών εκθεμάτων. Είναι ιδιαίτερα σημαντικό να διεγείρεις το ενδιαφέρον τόσο των μικρών όσο και των μεγάλων σε πολιτισμικά θέματα. Μια καλοστημένη εφαρμογή σε περιβάλλον μικτής πραγματικότητας ίσως είναι μια πολύ καλή λύση. [18], [19]

## Γραφικά και αναπαράσταση στο χώρο[20], [21]

Στην ενότητα αυτή παρουσιάζονται συνοπτικά κάποιες σημαντικές έννοιες από την Προβολική Γεωμετρία και πως χρησιμοποιείται στα Γραφικά. Τα μαθηματικά που χρησιμοποιούνται στην Προβολική Γεωμετρία είναι αρκετά σημαντικά για οποιονδήποτε σχεδιαστή μιας τρισδιάστατης εφαρμογής, ανεξαρτήτως αν η εφαρμογή αυτή είναι σχεδιασμένη για εκτεταμένη πραγματικότητα ή απλά για προβολή σε μία τυπική οθόνη. Ειδικά για την υλοποίηση του ελεγκτή στη συγκεκριμένη διπλωματική, η παρούσα ενότητα είναι εξαιρετικά σημαντική.

Αν και κάποιος στις μέρες μας μπορεί να τα αντιμετωπίσει ως ένα «μαύρο κουτί» - δε χρειάζεται και ιδιαίτερη γνώση για να φτιάξεις μια απλή εφαρμογή – η κατανόηση τους εντούτοις μπορεί να κάνει πολύ ευκολότερη την ανάπτυξη της εφαρμογής και να λύσει αρκετά προβλήματα.

### Γραφικά και τρισδιάστατος χώρος [20], [21]

Το κυριότερο πρόβλημα που πρέπει να λυθεί στα γραφικά είναι η ρεαλιστική απεικόνιση ενός τρισδιάστατου αντικειμένου, είτε αυτό αφορά το πεδίο της εκτεταμένης πραγματικότητας ή μία τυπική οθόνη. Ταυτόχρονα απαιτείται η δυνατότητα μεταφοράς (*translation*), περιστροφής (*rotation*) γύρω από έναν άξονα, αλλαγής του μεγέθους (*scale*) ενός αντικειμένου, αλλά και συνδυασμού των 3 παραπάνω πράξεων.

Ορίζουμε λοιπόν έναν τρισδιάστατο χώρο, όπου περιγράφουμε ένα σημείο του με τις συντεταγμένες . Αρχή του χώρο αποτελεί το σημείο

### Μεταφορά (*Translation*)

Η μεταφορά ενός σημείο στο χώρο είναι από τις πιο απλές πράξεις. Ουσιαστικά, αν έχουμε ένα σημείο στο χώρο, μία οποιαδήποτε μεταφορά εκφράζεται ως η αλγεβρική πρόσθεση των συντεταγμένων του σημείου αυτού με το ποσό κατά το οποίο επιθυμούμε να μετακινηθεί. Αυτό επεκτείνεται σε ένα σύνολο από σημεία, ώστε μπορούμε να πούμε πως η πρόσθεση αυτού του ποσού στις συντεταγμένες του κάθε σημείου προκαλεί ομοιόμορφη μεταφορά των σημείων στον τρισδιάστατο χώρο.

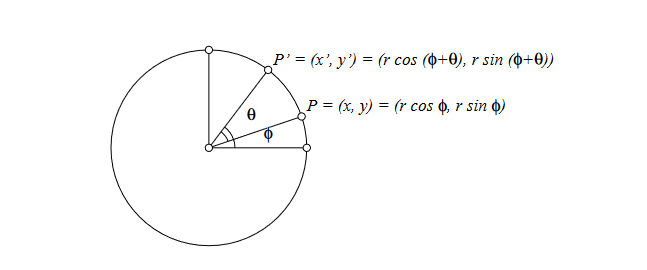
Ορίζουμε ένα γενικό τελεστή μεταφοράς ώστε:

,

όπου το ποσό κατά το οποίο μεταφέρεται ένα σημείο στο χώρο.

### Περιστροφή γύρω από άξονα (*Rotation*)

Θα ξεκινήσουμε περιγράφοντας την περιστροφή ενός σημείου στο επίπεδο -κατά γωνία με βάση το σημείο αναφοράς μας. Το σημείο μπορεί να γραφτεί με πολικές συντεταγμένες ως . Μία μετατόπιση κατά γωνία περιγράφεται καλύτερα στο παρακάτω σχήμα.



ΕΙΚΟΝΑ 2 Περιστροφή κατά τον άξονα Z

Χρησιμοποιώντας την ταυτότητα για την πρόσθεση ημίτονων-συνημίτονων, καταλήγουμε πως

ορίζοντας έτσι τους τελεστές περιστροφής γύρω από τους άξονες ως

Είναι σημαντικό να τονίσουμε εδώ πως τα διάφορα πρόσημα προκύπτουν έτσι επειδή το σύστημα συντεταγμένων μας είναι σύστημα «δεξιού χεριού». Τα συστήματα δεξιού-χεριού και αριστερού-χεριού έχουν να κάνουν με την κατεύθυνση του άξονα που αναφέρεται στο ύψος. Στην προκειμένη περίπτωση, έχουμε ορίσει αυτόν τον άξονα να είναι ο . Αν τα ορίζουν ένα επίπεδο σε ένα τραπέζι, τότε σε ένα σύστημα δεξιού-χεριού η θετική φορά του δείχνει πάνω από το τραπέζι, ενώ σε ένα σύστημα αριστερού-χεριού δείχνει κάτω από το τραπέζι.

### Γενικευμένη Περιστροφή

Στην προηγούμενη υπο-ενότητα ορίσαμε 3 διαφορετικούς τελεστές περιστροφής, έναν για κάθε διαφορετικό άξονα, οι οποίοι περιγράφουν την περιστροφή ενός σημείου με βάση την αναφορά μας (την αρχή των αξόνων). Συνεχίζοντας, θέλουμε να γενικεύσουμε την περιστροφή αυτή ώστε να γίνεται με βάση ένα αυθαίρετο άξονα.

Για να το πετύχουμε αυτό θα ακολουθήσουμε μια απλή τακτική. Επιλέγουμε έναν άξονα, έστω τον . Θα περιστρέψουμε το σημείο μας ώστε ο άξονας περιστροφής που θέλουμε να χρησιμοποιήσουμε - ο οποίος περνάει από την αρχή των αξόνων - να είναι παράλληλος με τον άξονα . Αυτό το πετυχαίνουμε εκτελώντας το πολύ δύο περιστροφές, έστω κατά γωνίες στους άξονες αντίστοιχα. Έπειτα περιστρέφουμε κατά την επιθυμητή γωνία στον άξονα και τέλος αντιστρέφουμε τις αρχικές περιστροφές που εκτελέσαμε. Η διαδικασία περιγράφεται από την παρακάτω εξίσωση

Θεωρώντας πλέον τον τελεστή ως έναν γενικευμένο τελεστή περιστροφής γύρω από οποιονδήποτε άξονα περνάει από το σημείο αναφοράς, μπορούμε με τον τελεστή μεταφοράς να γράψουμε μία τελική εξίσωση η οποία περιγράφει περιστροφή γύρω από ένα αυθαίρετο άξονα:

Η διαδικασία που περιγράφει η εξίσωση είναι απλή. Μεταφέρουμε το κέντρο της περιστροφής στην αρχή των αξόνων, εκτελούμε την περιστροφή και αντιστρέφουμε τη μεταφορά.

### Αλλαγή μεγέθους (*Scale*)

Η αλλαγή του μεγέθους ενός αντικειμένου είναι απλή. Αν έχουμε ένα αντικείμενο που περιγράφεται από ένα σύνολο σημείων στο χώρο, τότε για να μικρύνουμε ή να μεγαλώσουμε το αντικείμενο αυτό, αρκεί να πολλαπλασιάσουμε τις συντεταγμένες του με ένα βαθμωτό αριθμό. Έτσι επιτυγχάνουμε οποιαδήποτε αλλαγή στο μέγεθος, πάντα με βάση τους τρεις άξονες. Παραμένει όμως το πρόβλημα που είχαμε και πριν. Ένας πολλαπλασιασμός των συντεταγμένων θα άλλαζε το μέγεθος του αντικειμένου αλλά βάσει του σημείου αναφοράς, της αρχής δηλαδή των αξόνων. Αν είχαμε για παράδειγμα ένα αμάξι και θέλαμε να το μικρύνουμε κατά ένα συντελεστή 0.1, τότε ναι μεν θα το αμάξι θα ήταν το 1/10 σε μέγεθος, αλλά θα κατέληγε και περίπου κατά το 1/10 από την αρχή των αξόνων, σε σχέση με το σημείο που βρισκόταν αρχικά.

Η λύση σε αυτό το πρόβλημα έχει ήδη προταθεί στην υπο-ενότητα για τη γενική περιστροφή. Αρκεί να μεταφέρουμε το αντικείμενο στην αρχή των αξόνων, να εκτελέσουμε την αλλαγή μεγέθους και να το μεταφέρουμε ξανά πίσω στην αρχική του θέση.

Έτσι, μία γενική εξίσωση για την αλλαγή μεγέθους ενός συνόλου σημείων με βάση την αρχή των αξόνων είναι:

,

όπου ο κάθε συντελεστής αφορά τον κάθε άξονα, ώστε το τελικό αποτέλεσμα μπορεί να είναι και ανομοιόμορφο. Στην περίπτωση μάλιστα που κάποιος συντελεστής είναι αρνητικός αυτό θα έχει ως αποτέλεσμα την «αντανάκλαση» του αντικειμένου κατά επίπεδο το οποίο είναι κάθετο στον εκάστοτε άξονα.

### Συνδυασμός μεταφοράς, περιστροφής και αλλαγής μεγέθους

Συνήθως, αν όχι πάντα, συμφέρει να συνδυάσουμε τους διάφορους αυτούς μετασχηματισμούς ώστε να δημιουργήσουμε ένα πιο σύνθετο μετασχηματισμό ο οποίος περιγράφει ακριβώς αυτό που θέλουμε να κάνουμε. Φαίνεται όμως από την περιγραφή και τον ορισμό των τελεστών πως μια τέτοια διαδικασία θα γινόταν γρήγορα πολύπλοκη, όπως ήδη φαίνεται από τη γενικευμένη περιστροφή.

Η λύση στο παραπάνω πρόβλημα είναι η περιγραφή των 3 αυτών πράξεων με χρήση πινάκων. Η χρήση πινάκων, πέραν του ότι καθιστά πιο ευανάγνωστο και εύχρηστο το μετασχηματισμό, είναι βελτιστοποιημένη, αφού οι υπολογιστές σήμερα έρχονται με λογισμικό και υλικό το οποίο επιταχύνει τις πράξεις σε πίνακες.

Έστω λοιπόν πως τα σημεία αποτελούν διανύσματα στήλες. Μία οποιαδήποτε περιστροφή είναι απλά ένας πολλαπλασιασμός ενός πίνακα περιστροφής με το διάνυσμα στήλη. Έτσι, για τους 3 άξονες έχουμε

,

και

Από αυτήν την περιγραφή προκύπτει κάτι πολύ χρήσιμο και όμορφο. Συνεχόμενες περιστροφές δεν είναι τίποτα άλλο πέραν πολλαπλασιασμοί μεταξύ πινάκων. Αν δηλαδή για οποιοδήποτε λόγο χρειαστεί να περιστρέψουμε το αντικείμενο κατά 5 διαφορετικούς τρόπους, τότε δεν έχουμε παρά να πολλαπλασιάσουμε 5 διαφορετικούς πίνακες περιστροφής και εν τέλη να πολλαπλασιάσουμε το αποτέλεσμα με τα σημεία του αντικειμένου.

Η αλλαγή μεγέθους είναι ακόμα ευκολότερη. Η μετατροπή του τελεστή σε μορφή πινάκων είναι η εξής:

Όπως και πριν, πολλαπλασιασμός διαδοχικών πινάκων ισοδυναμεί με συνεχόμενους μετασχηματισμούς αλλαγής μεγέθους.

Προκύπτει, όμως, ένα πρόβλημα με αυτόν τον τρόπο περιγραφής των ανωτέρω μετασχηματισμών. Η περιγραφή της μεταφοράς με χρήση πινάκων, τουλάχιστον 3 x 3, είναι αδύνατη. Αυτό προκύπτει εύκολα αν εξετάσουμε τον πολλαπλασιασμό του σημείου με οποιονδήποτε 3 x 3 πίνακα. Το αποτέλεσμα πάντα θα είναι . Αυτό αναστατώνει τον τρόπο περιγραφής που έχει αναπτυχθεί, διότι πλέον δε καθίσταται δυνατός ο συνδυασμός και της μεταφοράς με συνεχόμενους πολλαπλασιασμούς πινάκων.

Η λύση προέρχεται από την επέκταση των παραπάνω πινάκων ώστε να έχουν μία παραπάνω στήλη και γραμμή, ή καλύτερα, ένα επιπλέον διάνυσμα στήλη του οποίου ο τελευταίο στοιχείο ισοδυναμεί με 1. Πλέον, ο μετασχηματισμός, παραδείγματος χάρη της περιστροφής γύρω από τον άξονα παίρνει τη μορφή:

,

Πλέον, με αυτήν την σύνθετη συντεταγμένη, ακόμα και ο μετασχηματισμός της μεταφοράς μπορεί να γραφεί με τον εξής τρόπο:

Συνεπώς, οποιοσδήποτε συνδυασμός της μεταφοράς, περιστροφής και αλλαγής μεγέθους δεν είναι τίποτα άλλο παρά συνεχόμενοι πολλαπλασιασμοί πινάκων. Μάλιστα, οι ανωτέρω συντεταγμένες με αυτή την επιπλέον συνιστώσα λέγονται ομοιογενείς. Χρησιμοποιούνται στην Προβολική Γεωμετρία όπως οι καρτεσιανές χρησιμοποιούνται στην Ευκλείδεια.

Η τελευταία συντεταγμένη δε χρειάζεται να είναι 1. Στις περισσότερες περιπτώσεις είναι γνωστή και ως , εφόσον και η κυριότερη χρήση των ομοιογενών συντεταγμένων είναι για μονοδιάστατους, δισδιάστατους και τρισδιάστατους Ευκλείδειους χώρους, Είναι ευνόητο πως η τιμή της δεν επηρεάζει το αποτέλεσμα των μετασχηματισμών.

### Τετραδόνιο (*Quaternion)* [22], [23]

Μία άλλη χρήσιμη έννοια στο πως λειτουργούν τα περιβάλλοντα ανάπτυξης τρισδιάστατων εφαρμογών είναι το τετραδόνιο. Αποτελεί μία σχετικά δυσνόητη έννοια η οποία όμως χρησιμοποιείται αρκετά στα γραφικά. Στο πλαίσιο της διπλωματικής θα τα αντιμετωπίσουμε ως ένα «μαύρο κουτί», περιγράφοντας κάποιες σημαντικές ιδιότητες και τη χρησιμότητά τους.

Το τετραδόνιο στα γραφικά χρησιμοποιείται ως ένας εναλλακτικός τρόπος περιγραφής μιας περιστροφής στον τρισδιάστατο χώρο, έναντι των ευκλείδειων γωνιών τις οποίες περιγράψαμε στις προηγούμενες ενότητες. Οι Euler γωνίες είναι ευάλωτες στο φαινόμενο *gimbal lock*. Το φαινόμενο αυτό συμβαίνει όταν δύο από τους άξονες περιστροφές ευθυγραμμίζονται και έτσι χάνεται ένας βαθμός ελευθερίας. Επιπλέον καθιστούν δυσκολότερη και αυθαίρετη την παρεμβολή (*interpolation*) μεταξύ δύο διαφορετικών περιστροφών / προσανατολισμών. Τα προβλήματα αυτά λύνονται με τη χρήση των τετραδόνιων αριθμών.

Το τετραδόνιο ουσιαστικά είναι μία μη-αντιμεταθετική επέκταση της θεωρίας των μιγαδικών αριθμών. Βασικά στοιχεία αποτελούν τα ώστε να ισχύει:

Ο αριθμός λοιπόν αυτός είναι ένας γραμμικός συνδυασμός των παραπάνω στοιχείων, των βασικών δηλαδή τετραδονίων . Έτσι μπορεί να εκφρασθεί με μοναδικό τρόπο ως:

Όπου πραγματικοί αριθμοί.

Χωρίς να επεκτείνουμε παραπάνω λόγω της πολυπλοκότητας των αριθμών αυτών, μπορούμε να ορίσουμε κάποιες σημαντικές ιδιότητες και κάποιους τελεστές που θα φανούν χρήσιμοι. Οι ορισμοί αυτοί δεν είναι μαθηματικά αυστηροί, αποτελούν απλώς μία προσπάθεια επεξήγησης της χρήσης και χρησιμότητας του τετραδόνιου.

1. Εάν έχουμε ένα τετραδόνιο , τότε ορίζουμε ως αντίστροφό του το

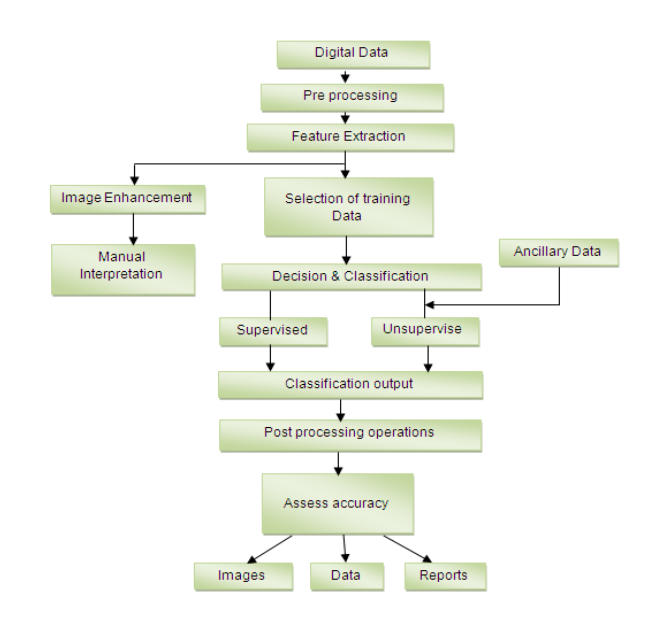
έτσι ώστε . Το δεν είναι τίποτα άλλο παρά ο μιγαδικός συζυγής του .

1. Ορίζουμε τελεστή περιστροφών, όπως ακριβώς εξηγήθηκε για τους διαδοχικούς πολλαπλασιασμούς σε προηγούμενη ενότητα. Για χάριν συντομίας και ευκολίας, ταυτίζουμε τον τελεστή , ώστε να μπορούμε να περιγράψουμε διαδοχικές περιστροφές ως
2. Ορίζουμε τελεστή όπου τετραδόνιο και ένα διάνυσμα τριών στοιχείων, ώστε το αποτέλεσμα να είναι ένα καινούριο διάνυσμα το οποίο έχει περιστραφεί κατά την περιστροφή που περιγράφει το τετραδόνιο. Για χάριν συντομίας και ευκολίας, ταυτίζουμε τον τελεστή , ώστε να είναι ένα διάνυσμα θέσης που προέκυψε από την περιστροφή του αρχικού διανύσματος κατά το τετραδόνιο.
3. Ορίζουμε τελεστή , όπου διάνυσμα στήλη το οποίο περιγράφει μία περιστροφή κατά γωνίες Euler. Επιπλέον, ισχύει και ο αντίθετος του, ώστε

Είναι σημαντικό να αναφερθεί πως ένα τετραδόνιο μπορεί να περιγράψει οποιουδήποτε είδους γενική περιστροφή, κατά οποιονδήποτε δηλαδή αυθαίρετο άξονα.

## Επεξεργασία εικόνας [24]

Στην ενότητα αυτή παρουσιάζεται ο τομέας της Επεξεργασίας Εικόνας. Συγκεκριμένα δίνεται ένας ορισμός, ο σκοπός της ύπαρξης αυτού του επιστημονικού πεδίου και η σημασία του στην υπολογιστική όραση, καθώς και η χρήση της σε περιβάλλοντα εκτεταμένης πραγματικότητας.



ΕΙΚΟΝΑ 3 Διάγραμμα Ροής Ενός Γενικού Συστήματος Επεξεργασίας Εικόνας

### Ορισμός-σύντομη περιγραφή

Η επεξεργασία εικόνας είναι μία μέθοδος κατά την οποία μία εικόνα μετατρέπεται σε ψηφιακή μορφή και έπειτα επεξεργάζεται ώστε να προκύψει μία βελτιωμένη εκδοχή της ή να εξαχθεί χρήσιμη πληροφορία από αυτή. Ως εκ τούτων, ορίζουμε ένα σύστημα επεξεργασίας εικόνας το οποίο έχει ως είσοδο μία εικόνα – όπως λόγου χάρη μία φωτογραφία ή ένα στιγμιότυπο από ένα βίντεο – και ως έξοδο δύναται να έχει πάλι εικόνα ή χαρακτηριστικά που έχουν εξαχθεί από την αρχική εικόνα. Συνηθέστερα ένα τέτοιο σύστημα δέχεται την εικόνα ως ένα δισδιάστατο σήμα και την επεξεργάζεται με διάφορες μεθόδους.

Παρότι υπάρχουν και οι έννοιες της οπτικής και αναλογικής επεξεργασίας εικόνας, εδώ αναφερόμαστε εξ ολοκλήρου στην ψηφιακή.

Τα διάφορα συστήματα του πεδίου αυτού μπορούν να συνοψιστούν ως εξής:

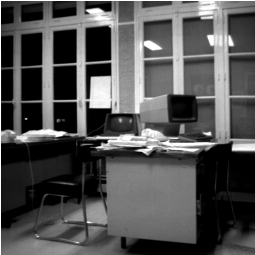
* Επεξεργασία Εικόνας (είσοδος εικόνα => έξοδος εικόνα)
* Ανάλυση Εικόνας (είσοδος εικόνα => εξαγωγή μετρικών)
* Κατανόηση εικόνας (είσοδος εικόνα => σημασιολογική ερμηνεία)

Από την άλλη, ο σκοπός των συστημάτων αυτών μπορεί να κατηγοριοποιηθεί σε 5 διαφορετικές ομάδες

1. Απεικόνιση - Η δυνατότητα παρατήρησης αντικειμένων τα οποία δεν είναι ορατά
2. Βελτίωση και αποκατάσταση εικόνας – η δημιουργία μίας καλύτερης από την αρχική εικόνας
3. Ανάκτηση εικόνας - Η εύρεση μιας εικόνας υψηλού ενδιαφέροντος
4. Αναγνώριση προτύπων
5. Αναγνώριση Εικόνας – Κατηγοριοποίηση και αναγνώριση διάφορων αντικειμένων μέσα σε μία εικόνα.

### Ανίχνευση ακμών [25]

Ένα από τα πολύ αναπτυγμένα πεδία της Επεξεργασίας Εικόνας είναι η ανίχνευση ακμών. Με τον όρο ανίχνευση ακμών εννοούμε τον εντοπισμό σημαντικών παραλλαγών σε μια ασπρόμαυρη εικόνα και η κατηγοριοποίηση των φυσικών φαινομένων που προκάλεσαν τις παραλλαγές αυτές. Η πληροφορία αυτή είναι πολλή σημαντική σε ποικίλα συστήματα, όπως τρισδιάστατη ανακατασκευή, κίνηση, αναγνώριση, βελτίωση, αποκατάσταση κ.ο.κ

Συγκεκριμένα, στα πλαίσια της διπλωματικής θα δούμε πως οι αλγόριθμοι που χρησιμοποιούν διάφορα συστήματα επαυξημένης – μικτής πραγματικότητας βασίζονται εσωτερικά σε κάποιου είδους ανίχνευσης σημείων ενδιαφέροντος για να προσδιορίσουν τη θέση της συσκευής στο χώρο.



ΕΙΚΟΝΑ 4 Πρωτότυπη Εικόνα

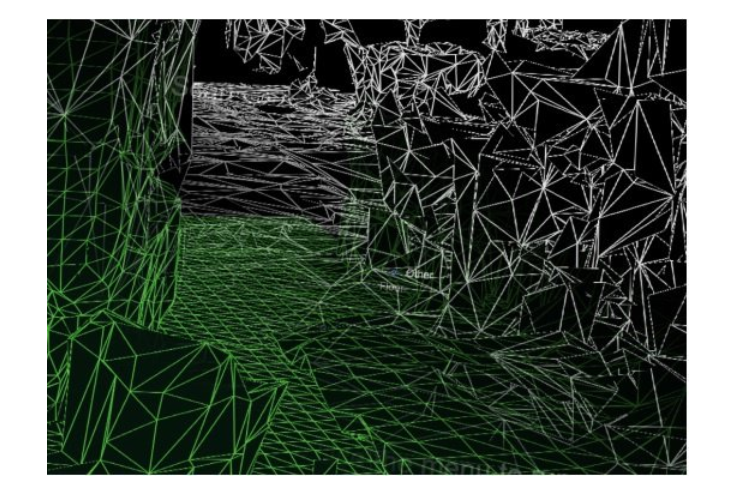
ΕΙΚΟΝΑ 5 Αποτέλεσμα Ανίχνευσης

### Υπολογιστική Όραση

Αποτελεί υψηλού επιπέδου επεξεργασία εικόνας κατά την οποία ένα μηχάνημα / υπολογιστής / λογισμικού προσπαθεί να αποκρυπτογραφήσει τα φυσικά περιεχόμενα μιας εικόνα. (παραδείγματος χάρη βίντεο ή τρισδιάστατες μαγνητικές τομογραφίες)

Είναι πλέον σημαντικός όρος για τα περιβάλλοντα μικτής πραγματικότητας, όπου η συσκευή πρέπει να αναγνωρίσει φυσικά στοιχεία του περιβάλλοντος ώστε ο εικονικός κόσμος να μπορεί να αλληλοεπιδρά με το φυσικό. Ουσιαστικά, μία συσκευή μικτής πραγματικότητας πρέπει:

* Να σαρώνει συνεχώς το περιβάλλον με σκοπό τη δημιουργία ενός σύννεφου σημείων (*point cloud*) ή αλλιώς πλέγματος (*mesh*) το οποίο είναι η εικονική αναπαράσταση του πραγματικού κόσμου στο περιβάλλον της συσκευής.
* Εκτός από τη εικονική αναπαράσταση του περιβάλλοντος, πρέπει να γίνεται σάρωση και αναγνώριση κινήσεων των χεριών του χρήστη, καθώς και χειρονομιών, ώστε να είναι ικανός να αλληλοεπιδράσει με το περιβάλλον.



ΕΙΚΟΝΑ 6 To Πλεγμα που Δημιουργειται μετα τη σαρωση του περιβαλλοντοσ [26]

## 1.4 Οπτική Αδρανειακή Οδομετρία – Ταυτόχρονος Εντοπισμός και Χαρτογράφηση

(*Visual Inertial Odometry [VIO] – Simultaneous Localization and Mapping [SLAM]*)

Η ενότητα αυτή παρουσιάζει τις έννοιεςΟπτική Αδρανειακή Οδομετρία (*VIO*) και Ταυτόχρονος Εντοπισμός και Χαρτογράφηση (*SLAM).* Είναι ιδιαίτερα σημαντικές έννοιες όσον αφορά την επαυξημένη και μικτή πραγματικότητα, διότι χωρίς αυτές, η συσκευή δε θα είχε τρόπο να αναγνωρίσει τη θέση της και τον προσανατολισμό της σε σχέση με τον εικονικό – μικτό κόσμο που δημιουργείται. Επιπλέον, χρησιμοποιήθηκαν στη διπλωματική εκτός της ίδιας της συσκευής, στον εξωτερικό ελεγκτή για την ανάπτυξη ενός συστήματος που παρέχει 6 βαθμούς ελευθερίας στο χρήστη.

Λόγω του ότι είναι ταχέα αναπτυσσόμενα πεδία και υπάρχουν πολλές και διαφορετικές τεχνικές, θα εστιάσουμε περισσότερο σε μια γενική εποπτεία των όρων αυτών, ώστε να γίνει αντιληπτή περισσότερο η φιλοσοφία των τεχνικών αυτών, παρά η υλοποίηση καθαυτή.

### Οδομετρία (*Odometry*)

Ως οδομετρία ορίζεται η χρήση δεδομένων τα οποία έχουν συλλεχθεί από αισθητήρες κίνησης ώστε να γίνει μία πρόβλεψη στη μεταβολή της θέσης ενός αντικειμένου. Ένα απλό οδόμετρο υπάρχει σε κάθε αμάξι, το οποίο μετρά τα συνολικά χιλιόμετρα που έχει διανύσει. Επιπλέον, βρίσκει χρήση στη ρομποτική σε δίποδα ή τροχοφόρα ρομπότ ώστε να γίνεται μία εκτίμηση της θέσης του ρομπότ σε σχέση με ένα αρχικό σημείο αναφοράς.

Μάλιστα, όσον φορά τη ρομποτική, η μέθοδος είναι ευαίσθητη στα σφάλματα λόγω της ολοκλήρωσης που απαιτείται από μετρήσεις τις επιτάχυνσης ή της ταχύτητας, που λόγω θορύβου δε δίνει σωστά αποτελέσματα. Μία καλή υλοποίηση της οδομετρίας απαιτεί τη συλλογή των μετρήσεων των αισθητήρων με ακρίβεια, τη σωστή βαθμονόμηση των οργάνων και την καλή επεξεργασία των δεδομένων.

Η εκτίμηση της θέσης και του προσανατολισμού μιας συσκευής στο χώρο με βάση κάποιο σημείο αναφοράς ονομάζεται ***ego-motion***.

### Οπτική Οδομετρία (*Visual Odometry*) [27]–[29]

Με τον όρο **οπτική οδομετρία** αναφερόμαστε σε διάφορες τεχνικές οδομετρίας οι οποίες ως βάση έχουν την εξαγωγή στοιχείων μέσα από βίντεο (με μία ή περισσότερες κάμερες), ούτως ώστε με τη χρήση διάφορων τεχνικών επεξεργασίας εικόνας και υπολογιστικής όρασης, να προκύψει μία καλή εκτίμηση της θέσης και του προσανατολισμού - που τυπικά ονομάζουμε **στάση, πόζα** (**pose**) - του συστήματος με βάση ένα αρχικό σημείο αναφοράς.

Τέτοιες τεχνικές φέρουν πολύ καλά αποτελέσματα όσον αφορά την ακρίβεια, αλλά δεν είναι πάντα εύρωστες, διότι εξαρτώνται σημαντικά από τον εντοπισμό σημείων ενδιαφέροντος (*feature points*) στην εικόνα-βίντεο. Αν λοιπόν τέτοια σημεία δε μπορούν να εντοπιστούν στο βίντεο - λόγου χάρη η κάμερα να κοιτάει έναν άσπρο τοίχο χωρίς ιδιαίτερα χαρακτηριστικά - τότε οι μέθοδοι αυτοί αδυνατούν να δώσουν ένα καλό αποτέλεσμα.

H οπτική οδομετρία χωρίζεται *κυρίως* σε 2 διαφορετικούς τύπους:

* ***Monocular***, όπου το σύστημα περιέχει μόνο μία κάμερα.
* **Stereo**, όπου το σύστημα έχει δύο κάμερες.

Επιπλέον, υπάρχουν δύο κύριες κατηγορίες τεχνικών:

* **Βάσει σημείων ενδιαφέροντος (*Feature-Based*)**, όπου τα σημεία ενδιαφέροντος εξάγονται από το βίντεο και έπειτα εντοπίζονται μέσα στο βίντεο.
* **Άμεση μέθοδος (*Direct method*)**, όπου χρησιμοποιείται ή ένταση του κάθε εικονοστοιχείου (*pixel*) ως είσοδος στον αλγόριθμο.

Υπάρχουν και υβριδικές μέθοδοι που συνδυάζουν τους δύο τρόπους.

Ένας γενικός αλγόριθμος οπτικής οδομετρίας βάσει σημείων ενδιαφέροντος παρουσιάζεται παρακάτω. Είναι σημαντικό όμως να τονίσουμε πως το πεδίο είναι ραγδαία εξελισσόμενο και είναι αρκετά δύσκολη η κατηγοριοποίηση και παρουσίαση όλων των αλγορίθμων:

1. Εξαγωγή εικόνων εισόδου μέσω μονής κάμερας, *stereo* κάμερας ή παγκατευθυντικής (*omnidirectional*) κάμερας
2. Διόρθωση εικόνας: επεξεργασία της εικόνας ώστε να αφαιρεθούν διάφορες παραμορφώσεις, όπως η παραμόρφωση από το φακό
3. Ανίχνευση σημείων ενδιαφέροντος: Ορισμός τελεστών ενδιαφέροντος, η συσχέτιση των σημείων ενδιαφέροντος μεταξύ των διαφορετικών εικόνων (*frames*) στο βίντεο και η δημιουργία του πεδίου οπτικής ροής (*optical flow, Lucas-Kanade* [30]*)*
4. Ανίχνευση σφαλμάτων και διόρθωση με χρήση του πεδίου οπτικής ροής
5. Εκτίμηση της κίνησης της κάμερας από την οπτική ροή
6. Είτε με φίλτρο Κάλμαν
7. Είτε με την ελαχιστοποίηση μιας συνάρτησης κόστους με σκοπό να βρεθούν οι γεωμετρικές και τρισδιάστατες ιδιότητες των σημείων ενδιαφέροντος.
8. Περιοδική ανασύσταση των εντοπισμένων σημείων ώστε να διατηρηθεί η ανίχνευση σε όλη την εικόνα.

Figure 1 Monocular Vio Flow Chart Example[31]

### Αδρανειακή Οδομετρία (Inertial Odometry) [32], [33]

Αδρανειακή Πλοήγηση (*Inertial Navigation*)

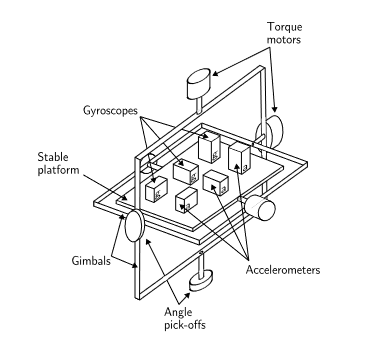
Με τον όρο **αδρανειακή οδομετρία** αναφερόμαστε στις τεχνικές οι οποίες ως βάση έχουν την επεξεργασία δεδομένων από διάφορους αισθητήρες αδράνειας (επιταχυνσιόμετρο, γυροσκόπιο, μαγνητόμετρο) ούτως ώστε να πάλι να προκύψει μία εκτίμηση της θέσης και του προσανατολισμού της συσκευής στο χώρο με βάση ένα αρχικό σημείο αναφοράς.

Τέτοιες τεχνικές είναι εύρωστες, διότι οι αισθητήρες παρέχουν συνεχώς δεδομένα, αλλά έχουν συνήθως μικρή ακρίβεια λόγω του θορύβου των μετρήσεων. Ένα κλασικό και εύκολο παράδειγμα για να καταλάβουμε τη σημασία του θορύβου και την ανάγκη αφαίρεσής του -όσο το δυνατόν γίνεται-είναι το εξής:

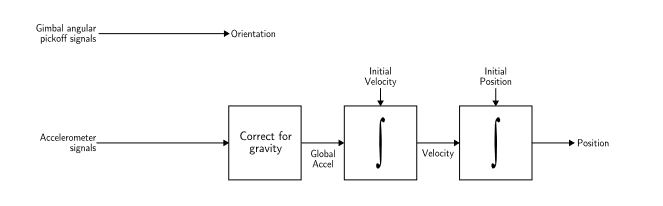
*Έχοντας κάποιες μετρήσεις από ένα επιταχυνσιόμετρο και μία αρχική θέση θέλω να βρω τη θέση της συσκευής στο χώρο.*

Μία άμεση λύση είναι η αριθμητική ολοκλήρωση της επιτάχυνσης ώστε να βρούμε την ταχύτητα και η αριθμητική ολοκλήρωση της ταχύτητας ώστε να βρούμε τη θέση. Ενώ η λύση θεωρητικά είναι άρτια, δε λαμβάνει υπόψιν τον θόρυβο που εισάγει στο σύστημα ένα επιταχυνσιόμετρο. Ένα τέτοιο όργανο, όσο καλό και να είναι, πάντα θα παρουσιάζει κάποιο θόρυβο στις μετρήσεις του. Αποτέλεσμα είναι να ολοκληρώνεται και ο θόρυβος - και μάλιστα στο τετράγωνο - και να έχουμε τη λεγόμενη **ολίσθηση** (*drift*), όπου η θέση στην αρχή φαίνεται να υπολογίζεται σωστά, αλλά μετά μετακινείται προς κάποια κατεύθυνση επιταχύνοντας ολοένα και περισσότερο. Εν ολίγης, ένα τέτοιο σύστημα δεν είναι ευσταθές και χρειάζεται επεξεργασία των μετρήσεων, φιλτράρισμα και καλή βαθμονόμηση ώστε να προκύψουν τα καλύτερα δυνατά αποτελέσματα.

Τα συστήματα αυτά υλοποιούνται με τη χρήση μίας **Μονάδας Αδρανειακής Μέτρησης** (***IMU***). Μία τέτοια μονάδα αποτελείται τυπικά από διάφορους αδρανειακούς αισθητήρες οι οποίο βασίζονται στην τεχνολογία *MEMS.* Παρακάτω απεικονίζεται σε σχεδιάγραμμα ένα τυπικό *IMU* καθώς και ένας απλουστευμένος αλγόριθμος της αδρανειακής οδομετρίας.



ΕΙΚΟΝΑ 7 Stable Platform IMU [31]

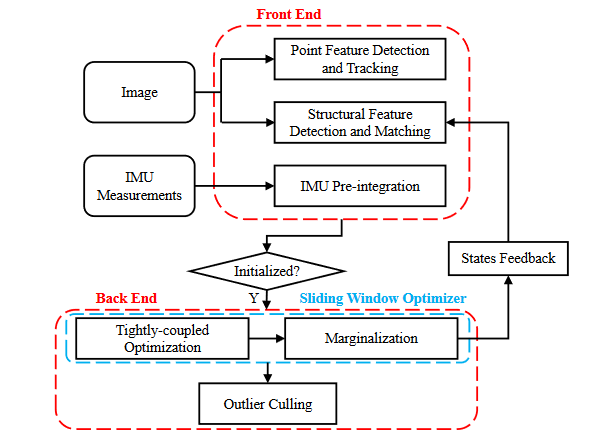


ΕΙΚΟΝΑ 8 Σύστημα Αδρανειακής Πλοήγησής [33]

### Οπτική Αδρανειακή Οδομετρία (*Visual Inertial Odometry*) [27]

Όπως γίνεται εύκολα αντιληπτό από το όνομα, η οπτική αδρανειακή οδομετρία συνδυάζει την οπτική οδομετρία με την αδρανειακή πλοήγηση σε μια προσπάθεια εξάλειψης των αδυναμιών των δύο αυτών τεχνικών.

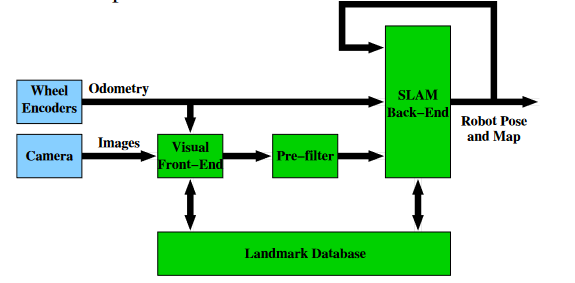
Με χρήση και των 2, το σύστημα γίνεται πιο εύρωστο και πιο ακριβές.



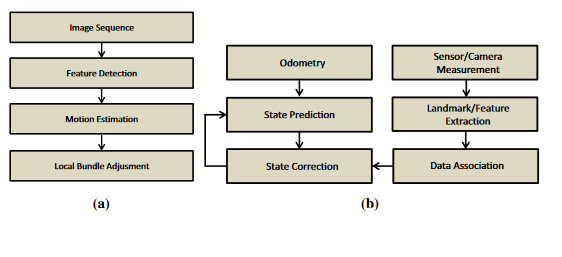
### 1.4.5 Ταυτόχρονος Εντοπισμός και Χαρτογράφηση (*Simultaneous Localization and Mapping*) – *SLAM* [34], [35]

Οι παραπάνω τεχνικές είναι αρκετά επιτυχείς στο να προσδιορίσουν τη θέση και τον προσανατολισμό μέσω κυρίως της μεταβολής διάφορων ποσοτήτων, είτε αυτές είναι θέσεις των σημείων ενδιαφέροντος είτε η θέση που προέκυψε από την επεξεργασία των μετρήσεων του επιταχυνσιόμετρου. Οι τεχνικές αυτές περισσότερο πετυχαίνουν μια τοπική συνέχεια στη διαδρομή που φαίνεται να ακολουθεί η συσκευή.

Σκοπός του **SLAM** είναι η χαρτογράφηση του χώρου. Για να αποκτήσει πλήρη αυτονομία ένα ρομπότ πρέπει να έχει την ικανότητα να εξερευνά το χώρο χωρίς εξωτερική παρέμβαση από κάποιον χρήστη. Για να επιτευχθεί αυτό, το ρομπότ δημιουργεί έναν αξιόπιστο χάρτη και εντοπίζει τον εαυτό του στον χάρτη.

Η έννοια του ***vSLAM*** (*visual SLAM*) συνδέεται άμεσα με το *VO-VIO*. Αυτό συμβαίνει διότι οι αλγόριθμοι που αναπτύχθηκαν για την οδομετρία ενός ρομπότ ή μιας οποιασδήποτε συσκευής μπορούν να επεκταθούν ώστε να χαρτογραφήσουν το χώρο. Ταυτόχρονα όμως, αυτή η χαρτογράφηση βοηθά και το υπόλοιπο σύστημα να διορθώσει τυχόν σφάλματα που προέκυψαν από την οδομετρία. Ούτως ή άλλως, τα σημεία ενδιαφέροντος τα οποία εντοπίζει το *VIO* σύστημα αποτελούν σημεία ενδιαφέροντος που μπορεί να αποθηκεύσει το *vSLAM*σύστημα.

ΕΙΚΟΝΑ 9 Block Διάγραμμα Ενός Συστήματος vSLAM [35]



ΕΙΚΟΝΑ 10 Block Διάγραμμα για (α) Vo kai (B) SLAM

### Εφαρμογές

Τα παραπάνω χρησιμοποιούνται σε διάφορες εφαρμογές. Μερικές από αυτές είναι:

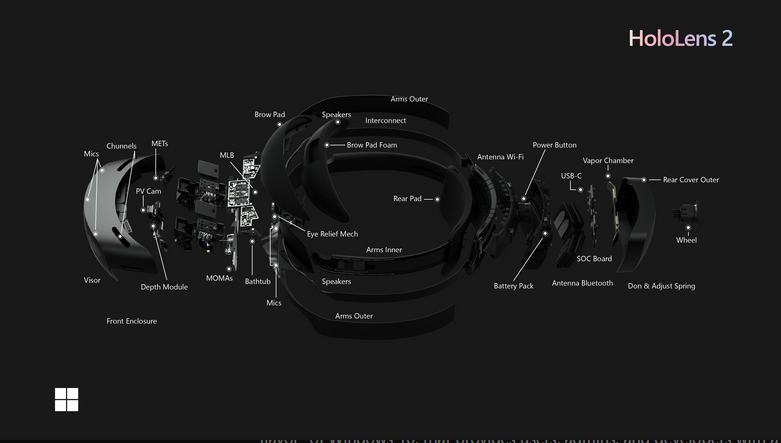
1. Στο *Mars Rovers*, όπου το αυτοκινούμενο όχημα χρησιμοποιεί τους αλγορίθμους ώστε να εξερευνήσει τον Άρη αυτόνομα. [29]
2. Στην εύρεση της τοποθεσίας μίας συσκευής σε περιπτώσεις όπου δεν υπάρχει καλό σήμα GPS. [36]
3. Στα έξυπνα κινητά με σκοπό τις εφαρμογές επαυξημένης πραγματικότητας.
4. Στα *HMDs* μικτής πραγματικότητας ώστε να γνωρίζουν τη θέση και τον προσανατολισμό τους στο χώρο. Ένα από αυτά τα *HMDs* είναι και το *Hololens.*

## Microsoft Hololens 2 [37]

Τα *Hololens* είναι μια σειρά έξυπνων γυαλιών μικτής πραγματικότητας, τα οποία και παράγει η *Microsoft*. Ξεκίνησαν με τα *Hololens 1* το 2016 και το 2019 εξέδωσαν μία νέα έκδοση, τα *Hololens 2,* τα οποία χρησιμοποιούνται μέχρι και σήμερα. Δυστυχώς τα *Hololens 1* πλέον δεν έχουν ιδιαίτερη υποστήριξη από τη *Microsoft* ως προς το λογισμικό ανάπτυξης εφαρμογών.

Στην ενότητα αυτή θα γίνει μία σύντομη περιγραφή των σημαντικότερων εξαρτημάτων της συσκευής καθώς και των εξαρτημάτων τα οποία είναι ιδιαίτερης σημασίας λόγω του θέματος της διπλωματικής. Επιπλέον, καταγράφονται οι διάφοροι τρόποι αλληλεπίδρασης τους οποίους προσφέρει η συσκευή στο χρήστη.

### Περιγραφή συσκευής



ΕΙΚΟΝΑ 11 Αποδόμηση της Συσκευής

To *Microsoft Hololens 2* είναι ένα *HMD* το οποίο λειτουργεί ως ολογραφικός υπολογιστής. Είναι η εξέλιξη του *Hololens 1*, με σκοπό να κάνει την εμπειρία πιο άνετη και πιο συναρπαστική, ώστε ο χρήστης να πιστεύει πως όντως βρίσκεται σε έναν εκτεταμένο κόσμο. Το λειτουργικό του είναι τα *Windows Holographic OS*, το οποίο βασίζεται στα *Windows 10*, δίνοντάς έτσι στον σχεδιαστή μια πλατφόρμα εύρωστη, ασφαλή και αποδοτική.

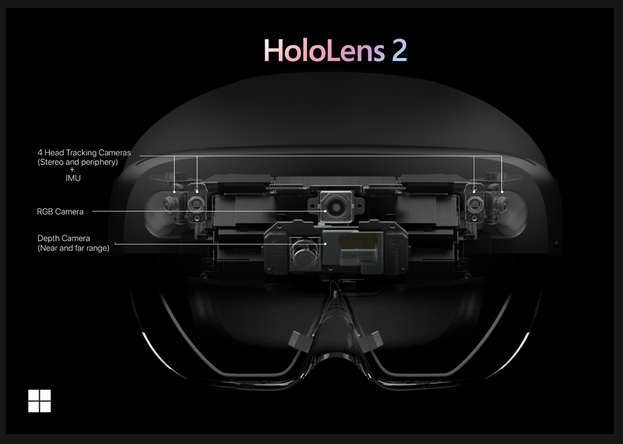
### Βασικά εξαρτήματα

* ***Visor*:** Η λεγόμενη προσωπίδα. Περιέχει τους αισθητήρες και τις οθόνες. Δίνεται η δυνατότητα περιστροφής της ακόμα και όταν φοράει κάποιος χρήστης τη συσκευή.
* ***Headband:* Μία ζώνη που περιβάλλει το κεφάλι και χρησιμοποιείται για να στηρίζεται η συσκευή πάνω σε αυτό. Ο χρήστης μπορεί να χρησιμοποιήσει μία ροδέλα που βρίσκεται στο πίσω μέρος της ώστε να σφίξει ή χαλαρώσει τη ζώνη αυτή, ώστε να τη νιώθει άνετα.**
* ***Brightness buttons*: Όταν ο χρήστης φοράει τη συσκευή, τα κουμπιά που ρυθμίζουν τη φωτεινότητα βρίσκονται στην αριστερή μεριά του *headband* κοντά στον κρόταφό του.**
* ***Volume buttons*: Με τη σειρά τους, τα κουμπιά που ρυθμίζουν την ένταση του ήχου βρίσκονται στη δεξιά μεριά του headband, κοντά στον κρόταφό του.**
* ***Power button*: Το κουμπί που ενεργοποιεί ή απενεργοποιεί τη συσκευή βρίσκεται στη δεξιά πλευρά του πίσω εξωτερικού καλύμματος.**
* ***USB-C port*: Μία θύρα *USB type C* βρίσκεται κάτω από το *power button*. Μέσω αυτής ο σχεδιαστής μπορεί να συνδέσει τη συσκευή στον υπολογιστή του εάν θέλει να ανεβάσει την εφαρμογή του.**

### Αισθητήρες

Οι αισθητήρες που βρίσκονται πάνω στη συσκευή αποτελούν πολύ σημαντικό κομμάτι της, γιατί όπως θα δούμε, είναι τα εξαρτήματα τα οποία υλοποιούν τους διάφορους αλγορίθμους επεξεργασίας εικόνας, υπολογιστικής όρασης και οπτικής οδομετρίας των οποίων η περιγραφή έγινε σε προηγούμενες ενότητες.

* ***Head tracking*:** Για τον εντοπισμό της κεφαλής, χρησιμοποιούνται 4 κάμερες ορατού φωτός. Αυτές οι κάμερες είναι που χρησιμοποιούνται για την υλοποίηση του *VIO-SLAM* ώστε η συσκευή να ξέρει τη θέση της και τον προσανατολισμός της στο χώρο.
* ***Eye tracking:*** 2 κάμερες υπέρυθρου φωτός χρησιμοποιούνται για τον εντοπισμό του βλέμματος του χρήστη.
* ***Depth:*** Ένας αισθητήρας βάθους, o οποίος χρησιμοποιείται για τη χαρτογράφηση του χώρου.
* ***IMU*:** Μία μονάδα μέτρησης της αδράνειας, η οποία περιλαμβάνει επιταχυνσιόμετρο, γυροσκόπιο και μαγνητόμετρο. Όπως είδαμε παραπάνω, χρησιμοποιείται στην οπτική οδομετρία για τη βελτίωση της εκτίμησης της θέσης και του προσανατολισμού της συσκευής.
* ***Camera*:** Μία κάμερα με χαρακτηριστικά 8-MP, 1080p30 βίντεο.



ΕΙΚΟΝΑ 12 Κύριοι Αισθητήρες του Hololens 2

Επιπλέον, αναφέρεται πως τα ηχεία εφαρμόζουν τεχνολογία χωρικού ήχου, προσφέροντας έτσι μια πιο ρεαλιστική ακουστική εμπειρία, αφού ο ήχος δίνει την εντύπωση πως έρχεται από διάφορα σημεία στο χώρο, ακριβώς όπως συμβαίνει στην πραγματικότητα.

Τέλος, για τη χωρική χαρτογράφηση του χώρου τα *Hololens* χρησιμοποιούν τον αισθητήρα βάθους και τις υπέρυθρες κάμερες. Ο αισθητήρας βάθους στέλνει ακτίνες στο χώρο, στις υπέρυθρες συχνότητες, οι οποίες και αντανακλώνται πάνω στα αντικείμενα. Έπειτα επιστρέφουν στα γυαλιά και συλλέγονται από τις υπέρυθρες κάμερες. Μέσω του χρόνου που έκανε να επιστρέψει μία ακτίνα, τα γυαλιά μπορούν να εκτιμήσουν την απόσταση των αντικειμένων από τον χρήστη και εν τέλει να δημιουργήσουν τον εικονικό χάρτη του περιβάλλοντος. Η διαδικασία αυτή είναι επαναληπτική.

### Τρόποι αλληλεπίδρασης – ελέγχου του *HoloLens 2* [38]

Το *HoloLens 2* προσφέρει διάφορους τρόπους αλληλεπίδρασης με το μικτό περιβάλλον που δημιουργεί. Σε γενικές γραμμές, η χρήση μοιάζει περισσότερο με τη χρήση ενός κινητού τηλεφώνου, όπου ο χρήστης χρησιμοποιεί περισσότερο τα χέρια του και λιγότερο εξωτερικές συσκευές, όπως ένα ποντίκι ή ένα πληκτρολόγιο.

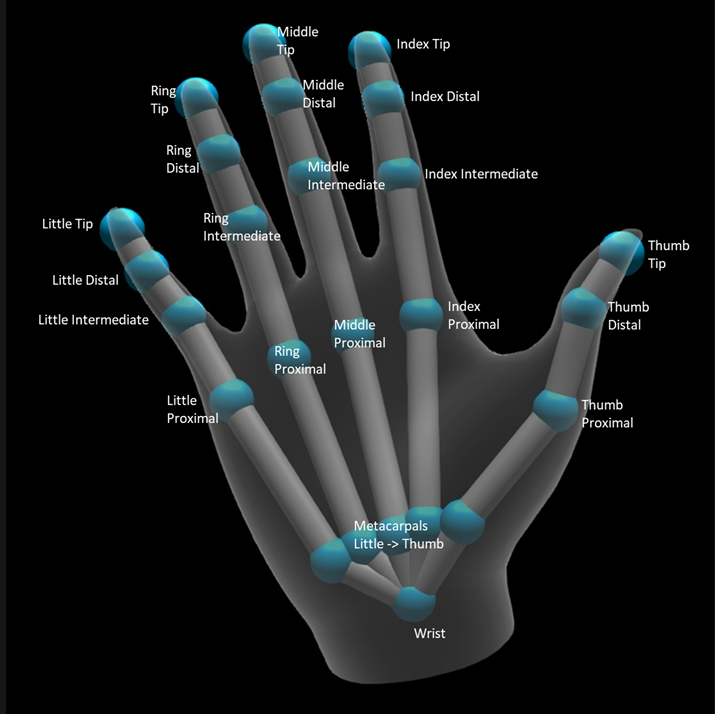
Παρακάτω παρουσιάζονται τα βασικά στοιχεία που έπειτα συνδέονται για τους διαφορετικούς τρόπους αλληλεπίδρασης.

* ***Gaze*-Βλέμμα** [39], [40]**:** Το «βλέμμα» του χρήστη. Το εικονικό βλέμμα του χρήστη είναι ουσιαστικά μια εικονική κουκίδα η οποία αναγράφεται στο οπτικό του πεδίο. Η κουκίδα αυτή ελέγχεται με 2 διαφορετικούς τρόπους, και τυπικά μπορεί να θεωρηθεί πως είναι το αντίστοιχο του κέρσορα του υπολογιστή:
* Η κουκίδα μετακινείται με βάση την κίνηση του κεφαλιού.
* Η κουκίδα μετακινείται με βάση το που βλέπει ο χρήστης. Βέβαια σε αυτήν την περίπτωση η *Microsoft* προτείνει να μην είναι ορατή η κουκίδα, διότι η κίνησή της βάσει του βλέμματος καθιστά κουραστικό στο χρήστη το γραφικό που βλέπει.
* ***Hand tracking*** [41]**:** Τα *HoloLens 2* προσφέρουν τη δυνατότητα ανίχνευσης και των 2 χεριών μέσω διάφορων τεχνικών υπολογιστικής όρασης και μηχανικής εκμάθησης. Ο χρήστης τοποθετεί τα χέρια του μπροστά από τις κάμερες της συσκευής ενώ η ίδια η συσκευή τρέχει συνεχώς έναν αλγόριθμο ανίχνευσης των χεριών. Όποτε εντοπιστεί ένα χέρι, το *Hololens* δημιουργεί ένα πλέγμα – σύνολο σημείων στο χώρο τα οποία και αποτελούν το εικονικό χέρι. Παρότι η ανίχνευση είναι αρκετά εύρωστη, παρατηρείται μικρή καθυστέρηση στην απόκριση των κινήσεων του χρήστη.

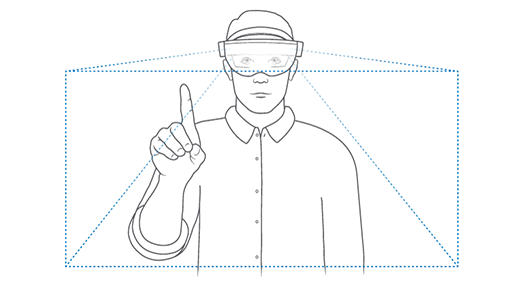
Το εικονικό χέρι που δημιουργείται μπορεί αν αλληλοεπιδράσει με τα εικονικά αντικείμενα με διάφορους τρόπους. Υπάρχει η δυνατότητα να τα αγγίξει, να τα περιστρέψει, να τα μετακινήσει με αρκετά οικείο τρόπο, περίπου όπως θα αντιμετώπιζε και ένα πραγματικό αντικείμενο. Με αυτόν τον τρόπο ο χρήστης μπορεί να πατήσει και σε διάφορα κουμπιά σε κάποιο μενού.

Τέλος, πέραν από την αλληλεπίδραση από κοντά (*near interaction*), το χέρι παρέχει και τη δυνατότητα αλληλεπίδρασης από μακριά (*far interaction*). Σε αυτήν την περίπτωση εμφανίζεται μια ακτίνα σε μια προκαθορισμένη γωνία από το ανιχνευμένο χέρι με την οποία μπορεί σαν δείξεις σε απομακρυσμένα αντικείμενα.

σκελετός και σημεία ενδιαφέροντος του χεριού



ΕΙΚΟΝΑ 13 Σκελετός και σημεία ενδιαφέροντος του χεριού



ΕΙΚΟΝΑ 14 Πλαίσιο Ανίχνευσης Χεριού [42]

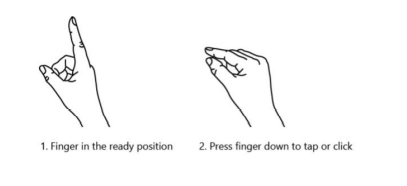
* **Χειρονομίες (*Gestures*):** Μέσω της ανίχνευσης του χεριού, καθώς και άλλων τεχνικών επεξεργασίας εικόνας, η συσκευή παρέχει τη δυνατότητα αναγνώρισης χειρονομιών. Παρακάτω θα αναφέρουμε τις βασικές χειρονομίες. Παρόλα αυτά, για κάποιον πιο εξειδικευμένο με την ανάπτυξη εφαρμογών στο *HoloLens 2*, υπάρχει και η δυνατότητα επέκτασης της αναγνώρισης αυτής με επιπλέον χειρονομίες.
* **Ηχητικές εντολές**[43]**:** Το *HoloLens 2* παρέχει τη δυνατότητα αναγνώρισης μιας πληθώρας φωνητικών εντολών, σε αρκετές μάλιστα γλώσσες. Οι φωνητικές εντολές είναι πολύ χρήσιμες όταν ο χρήστης δε θέλει να κουράζει το χέρι του. Παρόλα αυτά, κατά τη διάρκεια της διπλωματικής παρατηρήθηκε πως μερικές φορές αναγνώριζε λέξεις στα ελληνικά ως κάποια φωνητική εντολή στα αγγλικά.

Επιπλέον, παρέχονται διάφορα γενικευμένα API ώστε να επεκταθούν αυτές οι λειτουργίες. Τα API αυτά δεν αφορούν μόνο το *HoloLens 2,* αλλά και άλλες συσκευές μικτής πραγματικότητας της *Microsoft*. [44]

**Βασικές Χειρονομίες** [42]

Στο *HoloLens 2* παρέχονται δύο βασικές χειρονομίες που πρέπει να γνωρίζει ο χρήστης για την περιήγηση στο μικτό περιβάλλον.

1. ***Air tap*** [38], κατά το οποίο ο χρήστης δείχνει με το δείκτη περίπου προς το ταβάνι και έπειτα να αγγίξει τον αντίχειρα με το δείκτη του, σηκώνοντάς το κατευθείαν. Αυτή η χειρονομία αντιστοιχεί στο αριστερό κλικ του ποντικιού. Υπάρχει προφανώς και η ενδιάμεση κατάσταση στην οποία ο χρήστης παραμένει στη θέση 2, σα να μην έχει απελευθερώσει το χέρι του από το αριστερό ποντίκι. Αυτό ονομάζεται ***tap and hold***.



ΕΙΚΟΝΑ 15 Θέσεις ετοιμότητας και θέση tap

1. **Χειρονομία Έναρξης (*Start gesture*)**, η οποία επιτρέπει στο χρήστη να εμφανίσει το κύριο μενού καθώς και να το εξαφανίσει. Ο χρήστης πρέπει να βάλει το χέρι του στο πλαίσιο ανίχνευσης με τη μεριά της παλάμης να κοιτάει την κάμερα. Έπειτα πατάει στον καρπό του, στο σημείο όπου θα έχει εμφανιστεί ένα γραφικό που υποδηλώνει το μενού έναρξης. Έτσι το μενού εμφανίζεται ή εξαφανίζεται αναλόγως του τι κατάσταση βρισκόταν πριν. Το μενού αυτό ο χρήστης μπορεί να το διαχειριστεί πατώντας στα εικονικά κουμπιά που εμφανίζονται.

Υπάρχει και η **χειρονομία έναρξης με το ένα χέρι**, όπου πάλι ο χρήστης πρέπει να έχει το χέρι του στο πλαίσιο ανίχνευσης με τη μεριά της παλάμης να κοιτάει την κάμερα και έπειτα να κλείσει αντίχειρα και δείκτη όπως στη χειρονομία *air tap*.

ΕΙΚΟΝΑ 16 Start and One Handed Start gestures

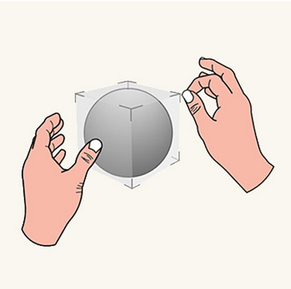


**Δείκτες, επιλογή και βασική αλληλεπίδραση** [45]

Σύμφωνα με τα παραπάνω, ο χρήστης στο *HoloLens 2* έχει τις εξής 2 κατηγορίες αλληλεπίδρασης:

1. **Χειρισμός από κοντά (*Near manipulation*)**

Η αλληλεπίδραση αυτή ενεργοποιείται όταν το αντικείμενο με το οποίο γίνεται η αλληλεπίδραση βρίσκεται έως και περίπου 50 εκατοστά από τον χρήστη. Στην αλληλεπίδραση αυτή ο χρήστης μπορεί να πιάσει το αντικείμενο, να το μετακινήσει, να το περιστρέψει και να αλλάξει το μέγεθός του. Η ίδια αλληλεπίδραση χρησιμοποιείται ώστε ο χρήστης να πατήσει ένα κουμπί ή μία επιλογή που εμφανίζεται κοντά του.

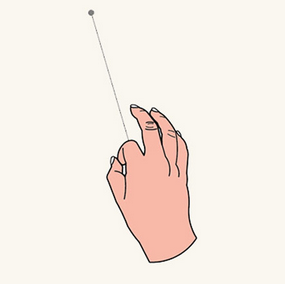
****

ΕΙΚΟΝΑ 17 Χειρισμός από κοντά

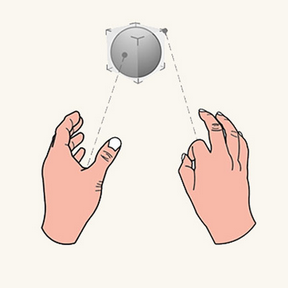
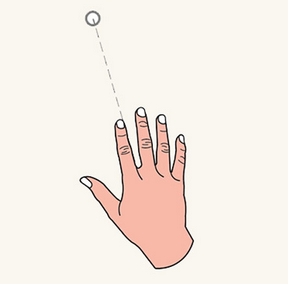
1. **Χειρισμός από μακριά (Far manipulation)**

Ο χειρισμός από μακριά ενεργοποιείται όταν το αντικείμενο με το οποίο γίνεται η αλληλεπίδραση βρίσκεται μακρύτερα των 50 εκατοστών από τον χρήστη. Στην αλληλεπίδραση αυτή έχουμε δύο τύπους δεικτών.

* Το κεφάλι ή το βλέμμα ως δείκτης.
* Μία ακτίνα που ξεκινά υπό μια μικρή γωνία από το κέντρο της παλάμης του χεριού και καταλήγει σε κάποιο αντικείμενο. Σημειωτέων πως εδώ μπορούν να αλληλοεπιδράσουν και τα 2 χέρια με το περιβάλλον.

Ανάλογα τώρα με το τι είδους αλληλεπίδραση είναι, ο χρήστης πάλι μπορεί να επιλέξει κάτι (με τη χρήση του *air tap*), να μετακινήσει, περιστρέφει ή αλλάξει μέγεθος σε ένα αντικείμενο καθώς και να το φέρει πιο κοντά.

ΕΙΚΟΝΑ 18 Κατάσταση Επικύρωσης (Commit State)

****

ΕΙΚΟΝΑ 19 Χειρισμός από κοντά

ΕΙΚΟΝΑ 20 Ελεύθερη Κατάσταση Δείκτη (Pointing State)

1. **Επιλογή με φωνητική εντολή**, κατά την οποία ο χρήστης κοιτάει μία επιλογή και την επικυρώνει τυπικά λέγοντας **Select** ή οποιαδήποτε άλλη εντολή έχει επιλέξει ο σχεδιαστής.
2. **Επιλογή με βλέμμα**, όπου ο χρήστης κοιτάει μία επιλογή για ένα χρονικό διάστημα – είτε με το δείκτη κεφαλής ή με το δείκτη βλέμματος – και η επιλογή επικυρώνεται μετά το πέρας του.

## 1.6 Πλατφόρμα ανάπτυξης Unity [46]

H Unity αποτελεί μία πλατφόρμα ανάπτυξης δισδιάστατων (***2D***) και τρισδιάστατων (***3D***) εφαρμογών. Υποστηρίζει, μεταξύ άλλων, ανάπτυξη εφαρμογών για υπολογιστή (*Windows, Linux, MacOS*), για UWP (Universal Windows Applications), για Android και iOS. Διατίθεται από το 2005 και είναι γραμμένη σε C++ (backend-runtime) και σε C# (backend-scripts). Η χρήση της διευκολύνει την ανάπτυξη εφαρμογών τις οποίες πρέπει να υποστηρίζουν διαφορετικές πλατφόρμες, αφού με τον ίδιο κώδικα μπορεί η εφαρμογή να τρέξει σχεδόν σε οποιαδήποτε από τις παραπάνω πλατφόρμες αναφέρθηκαν, με ελάχιστες ή και καμία αλλαγή.

Στην παρακάτω ενότητα αναπτύσσονται κάποια βασικά στοιχεία της Unity, καθώς και κάποια πιο εξειδικευμένα που αφορούν την ανάπτυξη *XR* εφαρμογών και την επέκταση της Unity πέραν της C#.

### 1.6.1 Βασικά στοιχεία

Η Unity, καθώς χτίστηκε για να βοηθήσει εξαρχής στην ανάπτυξη παιχνιδιών και έπειτα και σε γενικότερες εφαρμογές, παρέχει ένα περιβάλλον ανάπτυξης υψηλού επιπέδου. Δε χρειάζεται ο προγραμματιστής / σχεδιαστής να προγραμματίσει την εφαρμογή του από την αρχή, αλλά δουλεύει πάνω σε ένα πλαίσιο το οποίο ήδη του προσφέρει έτοιμα πολλά πράγματα. Ένα από τα σημαντικότερα είναι πως δε χρειάζεται να γράψει ούτε υλοποίηση για το πως σχεδιάζονται τα γραφικά στην οθόνη, αλλά ούτε να ασχοληθεί ιδιαίτερα με τα μαθηματικά τρισδιάστατου χώρου, περιγραφή των οποίων έγινε σε προηγούμενη ενότητα.

Το βασικό δομικό στοιχείο της Unity είναι το ***GameObject***. Αποτελεί μία οντότητα του περιβάλλοντος της Unity της οποίας τη συμπεριφορά μπορεί να καθορίσει ο προγραμματιστής μέσω των ***C# Scripts*** που γράφει. Ένα τυπικό *C# Script* το οποίο τοποθετείται πάνω σε ένα *gameobject* κληρονομεί από την κλάση ***MonoBehaviour***, η οποία παρέχει κάποιες βασικές λειτουργίες που έχει οποιαδήποτε συμπεριφορά. Όταν αυτό το *script* τοποθετείται πάνω σε ένα *gameobject* τότε λέμε πως το gameobject έχει αποκτήσει ένα ***Component***. H *Unity* παρέχει διάφορες συναρτήσεις ώστε τα *components* πάνω σε κάθε *gameobject* να έχουν τη δυνατότητα επικοινωνίας μεταξύ τους. Με αυτό τον τρόπο προωθείται το **αρθρωτό μοντέλο προγραμματισμού** (***component-oriented)***, το οποίο κατά μία έννοια είναι μία επέκταση του αντικειμενοστραφούς. Σε αυτό το μοντέλο προσπαθεί ο προγραμματιστής να περιγράψει το σύστημα το οποίο θέλει να φτιάξει με βάση τα βασικά δομικά στοιχεία του, τα οποία έπειτα συνεργάζονται μεταξύ τους για να εκτελέσουν μια λειτουργία. Έτσι αποφεύγεται η ύπαρξη μονολιθικών κλάσεων οι οποίες εκτελούν πολλαπλές λειτουργίες και η κάθε λειτουργία περιγράφεται μεμονωμένα.

Μία επιπλέον καλή τακτική στη Unity είναι οι κλάσεις να κληρονομούν το πολύ από μία κλάση βάση και όχι παραπάνω.

Να σημειωθεί επιπλέον πως ένα οποιοδήποτε *gameobject* έρχεται πάντα με ένα προκαθορισμένο *component*, το ***Transform***, το οποίο περιγράφει τη θέση, την περιστροφή και το μέγεθος του *gameobject* στον χώρο.

Με όσα αναφέρθηκαν παραπάνω κάποιος μπορεί να σχεδιάσει *scripts* και εφαρμογές στη *Unity* τα οποία είναι ευκολότερα τροποποιήσιμα και ευανάγνωστα.

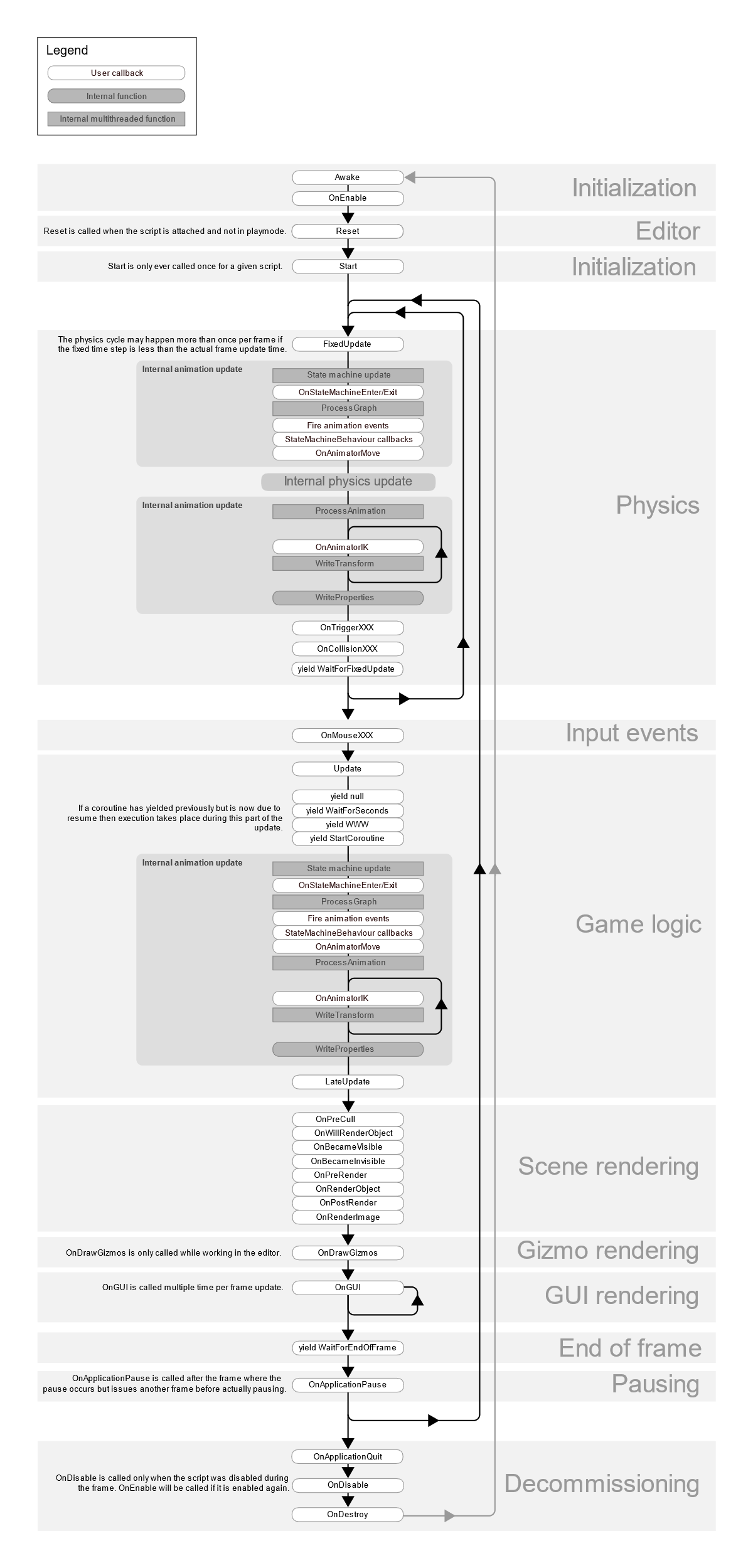
Τέλος, είναι σημαντικό να επισημάνουμε πως η πλατφόρμα πάνω στην οποία λειτουργεί το *Hololens 2* είναι η ***UWP (Universal Windows Platform)***, σχεδιασμένη από τη *Microsoft* ώστε να γράφονται εφαρμογές με τον ίδιο κώδικα για διαφορετικές συσκευές που υποστηρίζουν αυτήν την πλατφόρμα.

### 1.6.2 Κύκλος ζωής μιας εφαρμογής Unity [47]

Ένα οποιοδήποτε *component* έχει έναν κύκλο ζωής. Πέραν τον κύκλο ζωής του *component*, μια εφαρμογή στη *Unity* έχει διάφορα συστήματα τα οποία εκτελούν το δικό τους κύκλο ζωής. Παρακάτω αναφέρονται τα σημαντικότερα στάδια του κύκλου ζωής ενός *component* καθώς και παρουσιάζεται ένα αναλυτικό διάγραμμα ροής το οποίο περιγράφει πλήρως τα κύρια συστήματα της Unity.

**Σημαντικά γεγονότα / μέθοδοι στα *components***

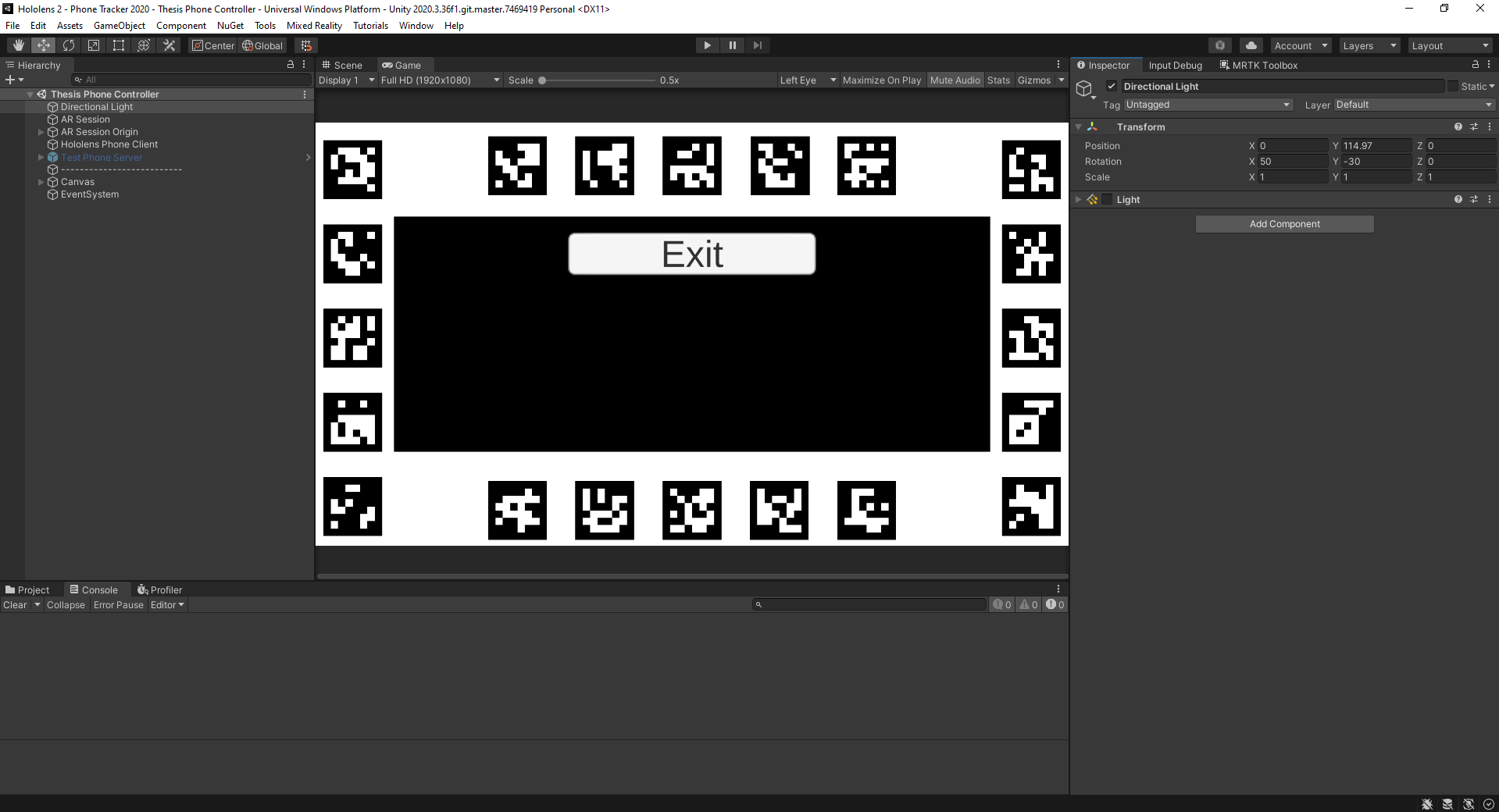
1. ***Awake:*** Η μέθοδος *Awake* καλείται όταν δημιουργείται για πρώτη φορά ένα *gameobject* το οποίο έχει πάνω ένα *component* με αυτή τη συνάρτηση. Δεν ξανακαλείται για το υπόλοιπο της ζωής του *gameobject*.
2. ***OnEnable:*** Η μέθοδος καλείται όταν δημιουργείται για πρώτη φορά ένα *gameobject* και είναι και ενεργό. Έπειτα καλείται όταν το *component* για οποιοδήποτε λόγο μεταβεί από την ανενεργό στην ενεργό κατάσταση.
3. ***OnDisable:*** Η μέθοδος καλείται όταν ένα *component* μεταβαίνειγια οποιοδήποτε λόγο από την ενεργό στην ανενεργό κατάσταση
4. ***Start:*** H μέθοδος καλείται όταν δημιουργηθεί ένα *gameobject* με τουλάχιστον ένα *component* που την υλοποιεί. Δεν ξανακαλείται κατά τη διάρκεια ζωής ενός *gameobject* και πάντα καλείται μετά την ***Awake***, ώστε τα διάφορα *components* να μπορούν να συνεργαστούν καλύτερα.
5. ***Update:*** Αφορά τη λογική του παιχνιδιού και τρέχει σε κάθε *frame*, με το ρολόι της εφαρμογής. Η μεταβολή του χρόνου δηλαδή μεταξύ κάθε *Update* δεν είναι σταθερή.
6. ***Fixed Update:***Αφορά την προσομοίωση της φυσικής και τρέχει με σταθερό ρολόι, ώστε η προσομοίωση να παρέχει όσο το δυνατόν πιο έγκυρα αποτελέσματα γίνεται



### 

ΕΙΚΟΝΑ 21 Κύκλος ζωής ενός Monobehaviour

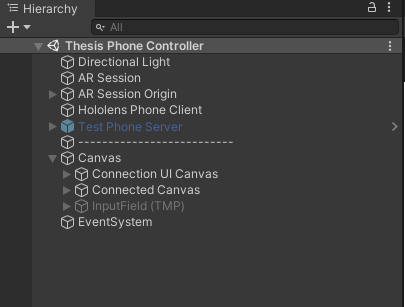
### 1.6.3 Βασικό Περιβάλλον Ανάπτυξης (Editor)



ΕΙΚΟΝΑ 22 Editor Unity

Στην παραπάνω εικόνα παρουσιάζεται το περιβάλλον ανάπτυξης (***Editor***) της Unity. Τα διάφορα παράθυρα μπορούν να επανατοποθετηθούν από τον χρήστη του *Editor* για την πιο εύκολη διαχείρισή του. Συνοπτικά τα πιο σημαντικά παράθυρα είναι:

* ***Project:*** Κάθε εφαρμογή στη *Unity* δημιουργεί ένα φάκελο που λέγεται ***Assets***. Μέσα σε αυτόν τον φάκελο βρίσκονται όλα τα αρχεία και τα δεδομένα, καθώς και τα *scripts* τα οποία χρησιμοποιεί ο προγραμματιστής / σχεδιαστής της εφαρμογής. Υπάρχει μία μπάρα αναζήτησης για διευκόλυνση εντοπισμού ενός επιθυμητού αρχείου και διάφορες άλλες δυνατότητες που στοχεύουν στη διευκόλυνση του χρήστη.
* ***Console (Κονσόλα):*** Αυτό το παράθυρο το χρησιμοποιεί ο σχεδιαστής ώστε να εκτυπώσει οτιδήποτε δεδομένα θέλει κατά τη διάρκεια της εφαρμογής. Είναι ουσιαστικά ένας τρόπος εντοπισμού σφαλμάτων (***debugging***), καθώς παρέχεται η δυνατότητα εποπτείας της εσωτερικής κατάστασης των διάφορων *components.*
* ***Profiler:*** O profiler χρησιμεύει στην επισκόπηση της λειτουργίας της εφαρμογής. Δείχνει πόση ώρα χρειάζονται διάφορα συστήματα της *Unity* για να εκπληρώσουν κάποιον κύκλο λειτουργίας καθώς και η λογική που έχει περιγράψει ο σχεδιαστής σε *components*.
* ***Hierarchy (Ιεραρχία):*** Η ιεραρχία είναι άκρως σημαντική, διότι δείχνει τα διάφορα *gameobjects* που έχουν τοποθετηθεί στην εφαρμογή. Παρατηρείται πως υπάρχει μια σχέση γονέα – παιδιού μεταξύ των *gameobjects*, σε μορφή δένδρου. Αυτή η σχέση υπάρχει και για την καλύτερη οργάνωση του σχεδιασμού αλλά και γιατί προσφέρει άλλα πλεονεκτήματα, όπως την αλλαγή του συστήματος αναφοράς του τρισδιάστατου χώρου σε ένα παιδί στο σύστημα αναφοράς του γονέα. Επιπλέον, δίνει τη δυνατότητα πρόσθεσης, αφαίρεσης και επιλογής *gameobjects*.

******

ΕΙΚΟΝΑ 23 Η ιεραρχία των gameobjects

* ***Inspector:*** Ο *inspector* αποτελεί τον τρόπο εποπτείας των διάφορων *components* που βρίσκονται πάνω σε ένα *gameobject*. Εδώ μπορεί ο χρήστης της *Unity* να προσθαφαιρέσει components καθώς και να κάνει αλλαγή στην αρχικοποίησή τους.
* ***Scene View – Game View:*** Το *Scene View* δίνει τη δυνατότητα περιήγησης στη σκηνή. Παρέχονται οι δυνατότητες τροποποίησης βασικών στοιχείων ενός *gameobject* - θέσης, περιστροφής και μεγέθους – μέσω ενός γραφικού εργαλείου. H κάμερα σε αυτό το view είναι ελεύθερη.0

Με την εναλλαγή στο Game View ο σχεδιαστής βλέπει ακριβώς το πως θα είναι το παιχνίδι κατά τη στιγμή της εκκίνησης, σύμφωνα πάντα με τον τρόπο που έχει ρυθμίσει την κάμερα.

Να σημειωθεί πως η γίνεται παρέχει πολλές από τις λειτουργίες της σε μορφή πακέτων, τα οποία και κατεβάζεις από τον ***Package Manager***.

### 1.6.4 *Plug-ins* [48]

H *Unity* προσφέρει τη δυνατότητα στον σχεδιαστή της χρήσης εξωτερικών βιβλιοθηκών, γραμμένες ακόμα και σε διαφορετική γλώσσα προγραμματισμού. Αυτό είναι ένα πολύ χρήσιμο εργαλείο όταν δεν αρκεί η C# για την υλοποίηση ενός αλγορίθμου ή όταν δεν υπάρχουν οι κατάλληλες βιβλιοθήκες για την ανάπτυξη του. Τα *plug-ins* χωρίζονται σε 2 κατηγορίες:

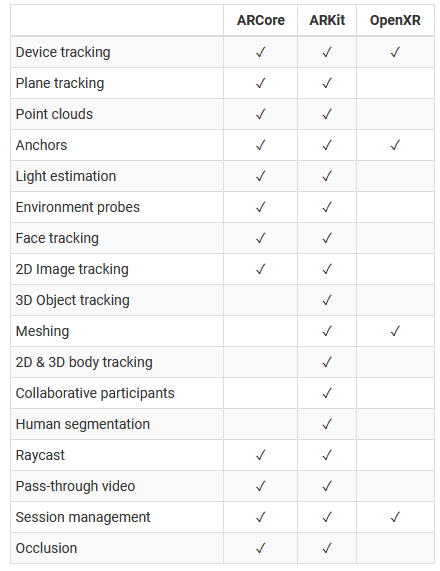
1. ***Managed plug-ins***: Είναι *compiled* σε *.NET assemblies*. Περιέχουν ουσιαστικά μόνο *.NET* κώδικα, που σημαίνει πως δε μπορούν να χρησιμοποιήσουν περαιτέρω χαρακτηριστικά πέραν αυτών που παρέχει το *.NET*. Η διαφορά τους με τα απλά *scripts* της *Unity* είναι ό,τι γίνονται *compile* εξωτερικά από τη *Unity*, σε αντίθεση με τα scripts που αποθηκεύονται μέσα στη *Unity* ως πηγαίος κώδικας.
2. ***Native plug-ins***: Αποτελούν κώδικα γραμμένο σε μια άλλη συνήθως γλώσσα προγραμματισμού, όπως παραδείγματος χάρη C++. Γίνονται *compile* εξωτερικά της *Unity* και δίνουν, μεταξύ άλλων, τη δυνατότητα κλήσης στο λειτουργικό σύστημα καθώς και συγκεκριμένα για μια πλατφόρμα χαρακτηριστικά τα οποία αλλιώς δεν είναι διαθέσιμα στη *Unity*.

Τα *plug-ins* τοποθετούνται σε συγκεκριμένους υπο-φακέλους στη *Unity* με συγκεκριμένα ονόματα για την κάθε πλατφόρμα. Έρχονται σε διάφορες μορφές, όπως ***.dll, .so*** κ.ο.κ

### 1.6.5 *AR Foundation* [49]

Ένα από τα πακέτα που παρέχει η *Unity* είναι το *AR Foundation*. Είναι ένα πακέτο ανάπτυξης εφαρμογών για περιβάλλοντα επαυξημένης πραγματικότητας το οποίο στοχεύει στην ενοποίηση των διαφορετικών υλοποιήσεων λογισμικού *AR* γύρω από ένα ενιαίο *API*. Η δυνατότητα αυτή είναι εξαιρετικά σημαντική, διότι μπορεί κάποιος πολύ εύκολα να αλλάξει την υλοποίηση διάφορων αλγορίθμων (όπως του *VIO* που αναφέρθηκε σε προηγούμενη ενότητα) με ελάχιστο κόπο. Τα πακέτα λογισμικού τα οποία ενοποιεί το συγκεκριμένο *API* είναι:

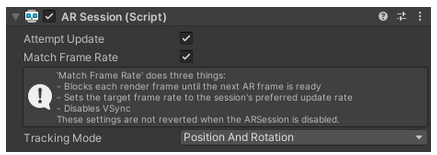
1. *Google ARCore για Android*
2. *Apple ARKit για iOS*
3. *Magic Leap για το Magic Leap*
4. *WindowsXR ή OpenXR για προϊόντα της Microsoft.*



ΕΙΚΟΝΑ 24 Χαρακτηριστικά των διάφορων plug-in

***AR Session***

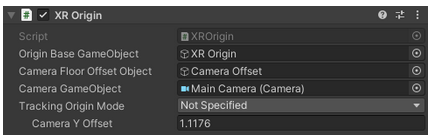
Αποτελεί το *component* το οποίο ελέγχει τον κύκλο ζωής μιας εφαρμογής επαυξημένης πραγματικότητας μέσω του *AR Foundation*



ΕΙΚΟΝΑ 25 AR Session component

***XR Origin***

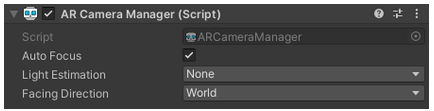
To *component* αυτό παρέχει στα ανιχνεύσιμα χαρακτηριστικά του χώρου, όπως για παράδειγμα επίπεδες επιφάνειες και σημεία ενδιαφέροντος, την τελική θέση, περιστροφή και μέγεθος στη σκηνή ώστε να συνδυάζονται με τον πραγματικό κόσμο.



ΕΙΚΟΝΑ 26 XR Origin Component

***AR Camera Manager***

To *component* διαχειρίζεται κάποια χαρακτηριστικά της κάμερας της συσκευής επαυξημένης πραγματικότητας. Επιπλέον υπάρχει ένα component για απενεργοποίηση της εικόνας της κάμερας στη συσκευή (*AR Camera Background*). Η δυνατότητα αυτή χρησιμεύει στα έξυπνα κινητά και όχι σε *HMDs* όπως το *Hololens*.



ΕΙΚΟΝΑ 27 AR Camera Component

***AR Input Manager***

To component αυτό ενεργοποιεί την ανίχνευση του κόσμου. Δεν έχει κάποια ιδιαίτερα χαρακτηριστικά.

### 1.6.6 MRTK (Mixed Reality Toolkit) [50]

Το *MRTK-Unity* είναι μία εργαλειοθήκη η οποία παρέχεται ως *package* για τη *Unity* για τη διευκόλυνση της ανάπτυξης και του σχεδιασμού εφαρμογών επαυξημένης πραγματικότητας. Προσφέρει στο σχεδιαστή μια συλλογή από *gameobjects* και *components*. Να σημειωθεί πως το *plug in* αυτό είναι παρόμοιο με το *AR Foundation,* αλλά είναι αποκλειστικά από τη *Microsoft* και ως εκ τούτου είναι καλύτερο να χρησιμοποιηθεί αυτό. Παρουσία και των 2 στην εφαρμογή δε φαίνεται να παρουσιάζει κάποιο πρόβλημα, παρόλα αυτά αρκεί η χρήση του ενός ώστε να μπορείς να χτίσεις μία *AR* εφαρμογή. Αναφερόμαστε συγκεκριμένα στο *MRTK 2.8* διότι αυτό χρησιμοποιήθηκε στα πλαίσια της διπλωματικής. Παρόλα αυτά γνωστοποιείται πως υπάρχει μία καινούρια έκδοση, το *MRTK 3.*

Οι πλατφόρμες και οι συσκευές που υποστηρίζονται είναι οι εξής:

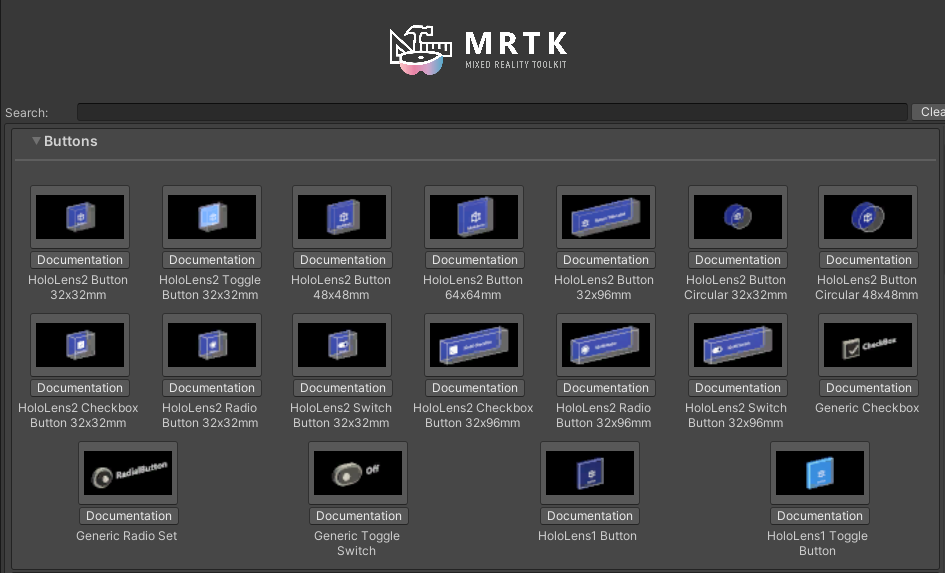


ΕΙΚΟΝΑ 28 Πλατφόρμες και Συσκευές MRTK

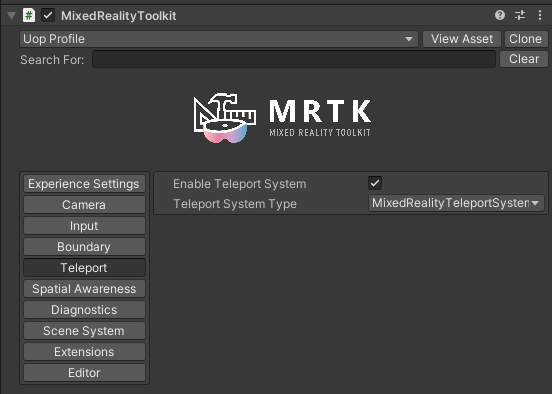
Η εγκατάστασή του γίνεται εύκολα από το *Mixed Reality Feature Tool*, ένα εργαλείο που παρέχει η *Microsoft* και όχι από τον *Package Manager* της *Unity*. Παρόλα αυτά θεωρείται ως ένα πακέτο το οποίο μετά εμφανίζεται στον *manager*.

Μετά την εγκατάστασή του εμφανίζεται ένα *tab* στον *Unity Editor* το οποίο παρέχει αρκετά εργαλεία, καθώς και παρέχονται αρκετά components. Μερικά από αυτά είναι:

1. Διευκόλυνση του τελικού build της εφαρμογής από τη Unity στο Hololens.
2. Ρύθμιση του *project* με τις καταλληλότερες ρυθμίσεις για την εκάστοτε εφαρμογή.
3. Εύκολη εναλλαγή διάφορων προφίλ, ανάλογα με τον τύπο εφαρμογής και τη συσκευή.
4. Διάφορα έτοιμα *UI Components* για τη δημιουργία διεπαφής χρήστη.
5. Παροχή των *component Solver* τα οποία παρέχουν τη δυνατότητα τοποθέτησης διάφορων στοιχείων με βάση τη θέση, το κεφάλι, το χέρι του χρήστη κ.ο.κ



ΕΙΚΟΝΑ 29 MRTK Toolbox - Components



ΕΙΚΟΝΑ 30 Ρυθμίσεις του MRTK

### 1.6.7 *Microsoft Visual Studio* [51]

To *Microsoft Visual Studio* είναι τυπικά ο *IDE* / code *editor* που χρησιμοποιείται για ανάπτυξη εφαρμογών στη *Unity*. Εκτός της *Unity*, μπορεί να χρησιμοποιηθεί και για την ανάπτυξη *plug-in*, αφού περιέχει διάφορους *compilers*.

Συγκεκριμένα για το *HoloLens*, η *Unity* παράγει εξωτερικά ένα *Visual Studio Project* το οποίο μετά ο σχεδιαστής μπορεί να κάνει *compile* και να το φορτώσει στη συσκευή, είτε με ασύρματη σύνδεση (*Wi-fi*) είτε ενσύρματα μέσω μίας θύρας *USB*. H διαδικασία αυτή όμως δεν είναι σύντομη, γι’ αυτό ο σχεδιαστής κρίνεται αναγκαίο να χρησιμοποιεί το ***Holographic Remoting***, ένα *API* το οποίο καθιστά ικανή την επικοινωνία *Unity Editor* με το HoloLens, ούτως ώστε η εφαρμογή τρέχει στον υπολογιστή και επικοινωνεί με τη συσκευή, ανταλλάσσοντας πληροφορίες σχετικά με το χώρο και την αναγνώριση χειρονομιών. Ο σχεδιαστής εν τέλη βλέπει μια προσομοίωση της εφαρμογής στο *Hololens* χωρίς να χρειάζεται να τη φορτώσει σε αυτό.

# Κεφάλαιο 2 – Σχεδίαση πλαισίου επικοινωνίας μεταξύ ελεγχόμενης συσκευής και ελεγκτή

Στο παρόν κεφάλαιο αναπτύσσεται η σχεδίαση της επικοινωνίας μεταξύ ενός εξωτερικού ελεγκτή και μίας συσκευής την οποία επιθυμούμε να ελέγξουμε. Στη συγκεκριμένη περίπτωση, η ελεγχόμενη συσκευή θα είναι το HoloLens 2 και ο ελεγκτής ένα έξυπνο κινητό. Παρουσιάζεται στην αρχή ο σκοπός και η πιθανή χρησιμότητα μιας τέτοιας επικοινωνίας. Έπειτα δίνεται ο σχεδιασμός ενός συστήματος ***serialization*** και ενός συστήματος επικοινωνίας μεταξύ των δύο συσκευών. Στη συνέχεια γίνεται αναφορά στο *Input System* της *Unity,* το οποίο και χρησιμοποιείται για την εύκολη κωδικοποίηση σημάτων του ελεγκτή. Τέλος, συνδυάζονται όλα αυτά για τη δημιουργία ενός πελάτη (*client*) και ενός διακομιστή (server), που θα τρέχουν στον ελεγκτή και στην ελεγχόμενη συσκευή αντίστοιχα.

Η σχεδίαση έγινε στη *Unity 2020.3 (*προτείνεται για το HoloLens το 2022), με το *MRTK 2.8*.

## 2.1 Σκοπός

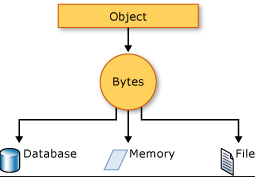
Όπως είδαμε στο Κεφάλαιο 1, το *HoloLens 2* προσφέρει πολλούς και διαφορετικούς τρόπους αλληλεπίδρασης. Ενώ είναι όλοι εξαιρετικά χρήσιμοι και σχεδιασμένοι ώστε ο χρήστης να νιώθει πως βρίσκεται σε ένα εκτεταμένο περιβάλλον, πέραν του πραγματικού, έρχονται με κάποια μειονεκτήματα.

1. Οι τρόποι αλληλεπίδρασης που αφορούν την οποιαδήποτε ανίχνευση των χεριών επιβάλλουν τα χέρια να βρίσκονται μέσα στο πλαίσιο ανίχνευσης της κάμερας. Παρότι μεγάλο, αυτό σημαίνει πως ο χρήστης πρέπει να μετακινεί συνέχεια τα χέρια του και όλο του το σώμα, μία διαδικασία που ίσως αποβεί κουραστική.
2. Οι τρόπου αλληλεπίδρασης που αφορούν αναγνώριση φωνητικών εντολών ή επιλογή μέσω δείκτη κεφαλής – βλέμματος απαιτούν περισσότερο χρόνο, καθιστώντας την εμπειρία πιο αργή.
3. Επέκταση της αλληλεπίδρασης με την εισαγωγή λειτουργιών που το HoloLens δε μπορεί να υποστηρίξει. Το κινητό παρέχει μία οθόνη αφής, αισθητήρες, δόνηση και πληθώρα άλλων λειτουργιών που μπορούν αν συνδυαστούν για τη δημιουργία ενός καινούριου τύπου αλληλεπίδρασης.

Σημειώνεται πως το πλαίσιο επικοινωνίας είναι γραμμένο σε *C#* εντός *Unity* και για τον διακομιστή αλλά και για τον πελάτη. Επιπλέον, η σχεδίαση χρησιμοποιεί μία τροποποιημένη εκδοχή ενός πλαισίου δικτύωσης ελεύθερου κώδικα (*open source*), του **MASTERSERVER FRAMEWORK** [52].

## 2.2 Κωδικοποίηση (*Serialization*) [53]

Για να είναι ικανή η οποιαδήποτε επικοινωνία μεταξύ δύο διαφορετικών συστημάτων, πρέπει να υπάρχει ένα ενιαίο πλαίσιο επικοινωνίας. Αυτό σημαίνει πως πρέπει από το εσωτερικό περιβάλλον μιας εφαρμογής πελάτη να κωδικοποιηθεί ένα αντικείμενο σε μια άλλη μορφή (**ροή από bytes**), η οποία θα σταλεί μέσω ενός δικτύου στην άλλη συσκευή, θα ληφθεί από την εφαρμογή διακομιστή και έπειτα θα αποκωδικοποιηθεί (*deserialize*) σε κάτι χρήσιμο για την εφαρμογή αυτή.



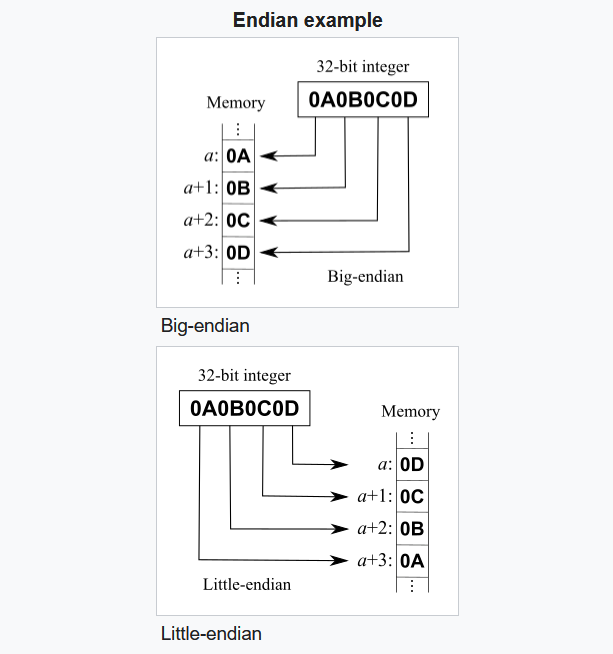
ΕΙΚΟΝΑ 31 Διάγραμμα ροής Κωδικοποίησης

### 2.2.1 Η κλάση *EndianBitConverter*

Ορίζουμε την κλάση βάση *EndianBitConverter*, της οποίας ο σκοπός είναι να παρέχει μεθόδους οι οποίες μετατρέπουν βασικούς τύπους δεδομένων, π.χ. έναν ακέραιο, σε έναν πίνακα από bytes. Επιπλέον, παρέχουν μεθόδους και για τον αντίστροφο μετασχηματισμό, τη μετατροπή δηλαδή ενός πίνακα από bytes σε ένα βασικό τύπο δεδομένων.

Η κλάση αυτή είναι αφαιρετική κλάση διότι έγινε η επιλογή παροχής της μετατροπής σε bytes κατά δύο τρόπους, με *little-endian* και με *big-endian*. Η εφαρμογή χρησιμοποιεί τον *big-endian* τρόπο. Αυτό σημαίνει πως η αποθήκευση ξεκινά από το ***least significant bit (LSB)*** και τελειώνει στο ***most significant bit (MSB)***. Οι υλοποιήσεις τις κλάσης αυτής είναι αντίστοιχα ο *LittleEndianBitConverter* και ο *BigEndianBitConverter*.

Η ανάγκη υλοποίησης των δύο αυτών διαφορετικών τρόπων αποθήκευσης προέκυψε επειδή ο ήδη υπάρχον *BitConverter* του *.NET* παίρνει το *endianness* [54] από την αρχιτεκτονική του υπολογιστή, ενώ η κωδικοποίηση απαιτεί να είναι σταθερό και στις συσκευές για λόγους επικοινωνίας.



ΕΙΚΟΝΑ 32 Endianness



ΕΙΚΟΝΑ 33 Βασικοί Τύποι Δεδομένων στη C# [55]

### 2.2.2 H κλάση *EndianBinaryWriter*

Η κλάση αυτή παρέχει μεθόδους για τη μετατροπή μιας δεδομένων σε μια σειρά από bytes. Παίρνει ως ορίσματα έναν τύπο *EndianBitConverter,* ένα *Encoding* και ένα *Stream*. To *Encoding* αφορά τον τρόπο κωδικοποίησης των αλφαριθμητικών, ενώ το *Stream* είναι μία κλάση την οποία μπορείς να χρησιμοποιήσεις για να αποθηκεύσεις και να αναζητήσεις μια αλληλουχία από bytes. Η κλάση *Stream* είναι μία αφαιρετική κλάση. Στην περίπτωσή μας χρησιμοποιούμε την υλοποίηση *MemoryStream* που δίνεται από τη βασική βιβλιοθήκη.

Οι διάφορες μέθοδοι που υλοποιούνται είναι *overloads* με χαρακτηριστική ονομασία *Write*, η οποία παίρνει ως ορίσματα όλα τους βασικούς τύπους που περιγράφηκαν παραπάνω και τους γράφει στο *Stream* μέσω του *EndianBitConverter*.

### 2.2.3 Η κλάση *EndianBinaryReader*

Αντίστοιχα η κλάση αυτή χρησιμοποιείται για τη μετατροπή μιας αλληλουχίας από bytes στους βασικούς τύπους της C#. Πάλι παίρνει ως ορίσματα έναν τύπο *EndianBitConverter*, ένα *Encoding* και ένα *Stream* και παρέχει μεθόδους για την αποκωδικοποίηση ενός μέρους μιας αλληλουχίας από *bytes*.

### 2.2.4 Η διεπαφή *ISerializablePacket*

Η διεπαφή *ISerializablePacket* ορίζει τρεις μεθόδους:

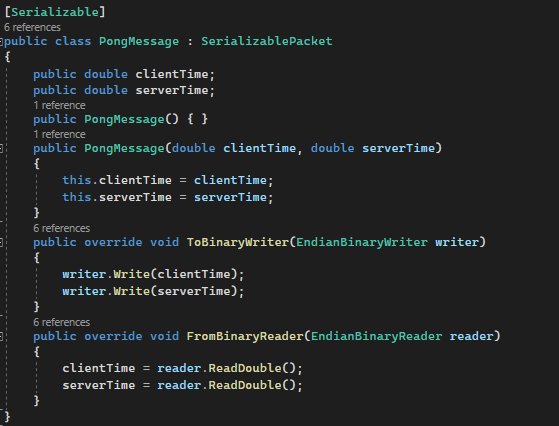
1. Μία μέθοδο για την κωδικοποίηση ενός αντικειμένου αυτής της κλάσης σε έναν *EndianBinaryWriter*.
2. Μία μέθοδο για την αποκωδικοποίηση ενός αντικειμένου αυτής της κλάσης από έναν *EndianBinaryReader*.
3. Mία μέθοδο που επιστρέφει την κλάση ως έναν πίνακα από bytes. Η μέθοδος αυτή είναι παρωχημένη, αφού μπορεί η υλοποίηση να υπάρχει πάντα από τον *EndianBinaryWriter*. Παρόλα αυτά, για λόγους πληρότητας παρέμεινε στον ορισμό της διεπαφής.

Ο σκοπός της διεπαφής είναι ο σχεδιαστής / προγραμματιστής να μπορεί να δημιουργήσει δικά του πακέτα, συγκεντρώνοντας την πληροφορία σε ένα αντικείμενο. Αυτό κάνει το σχεδιασμό πιο αποδοτικό, καθώς και τον κώδικα πιο ευανάγνωστο.

Να σημειωθεί εδώ πως ένα *ISerializablePacket* πρέπει να υλοποιηθεί από τον προγραμματιστή. Αυτό σημαίνει πως δεν υπάρχει κάποια αυτόματη μετατροπή οποιασδήποτε κλάσης υλοποιεί αυτή τη διεπαφή. Μία βελτίωση του συγκεκριμένου συστήματος κωδικοποίησης θα ήταν η εξής:

Μέσω *Reflection* [56]να διαβάζονται οι βασικοί τύποι που ορίζονται ως μέλη μίας κλάσης που υλοποιεί μια εφαρμογή και μετά μέσω τεχνικών *weaving* [57] να υλοποιείται η κλάση κατά το *compile time*. Η υλοποίηση όμως αυτής της τεχνικής είναι σχετικά πολύπλοκη και ξεφεύγει από τα πλαίσια της διπλωματικής.

Τέλος, υπάρχει μία υλοποίηση της διεπαφής, η κλάση *SerializablePacket*, η οποία είναι αφαιρετική κλάση η οποία υλοποιεί όμως τη μέθοδο 3). Αυτό συμβαίνει διότι είχαν γραφτεί ήδη αρκετά πακέτα με αυτόν τον τρόπο, και παρόλο που η μέθοδος είναι παρωχημένη, δεν ήταν επιθυμητό να δημιουργηθεί πρόβλημα στο σύστημα.



ΕΙΚΟΝΑ 34 Απλή Υλοποίηση Ενός Pong Μηνύματος

## 2.3 Ενοποιημένη Επικοινωνία

Στην ενότητα αυτή αναλύεται η σχεδίαση των κλάσεων και components που αποτελούν ένα γενικό σύστημα επικοινωνίας μεταξύ ενός πελάτη και ενός διακομιστή.

### 2.3.1 H διεπαφή IMessage

# Βιβλιογραφία

[1] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, “Augmented reality: A class of displays on the reality-virtuality continuum,” *Telemanipulator Telepresence Technol.*, vol. 2351, Jan. 1994, https://doi.org/10.1117/12.197321

[2] “Extended reality,” *Wikipedia*. Sep. 18, 2022. Accessed: Sep. 21, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Extended\_reality&oldid=1110892164

[3] “Virtual reality,” *Wikipedia*. Sep. 13, 2022. Accessed: Sep. 21, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Virtual\_reality&oldid=1110016947

[4] M. Speicher, B. Hall, and M. Nebeling, “What is Mixed Reality?,” May 2019. https://doi.org/10.1145/3290605.3300767

[5] M. Javaid and A. Haleem, “Virtual Reality applications toward medical field,” *Clin. Epidemiol. Glob. Health*, vol. 8, Dec. 2019, https://doi.org/10.1016/j.cegh.2019.12.010

[6] H. O. Barros, M. M. Soares, E. L. R. Filho, W. Correia, and F. Campos, “Virtual Reality Immersion: An Important Tool for Diagnostic Analysis and Rehabilitation of People with Disabilities,” *Lect. Notes Comput. Sci.*, pp. 337–344, 2013.

[7] “Half-Life: Alyx,” *Half-Life*. https://www.half-life.com/en/alyx (accessed Sep. 21, 2022).

[8] L. digital studio, “Beat Saber - VR rhythm game.” https://www.beatsaber.com/ (accessed Sep. 21, 2022).

[9] H. Luo, G. Li, Q. Feng, Y. Yang, and M. Zuo, “Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019,” *J. Comput. Assist. Learn.*, vol. 37, no. 3, pp. 887–901, 2021, https://doi.org/10.1111/jcal.12538

[10] “Augmented reality,” *Wikipedia*. Sep. 20, 2022. Accessed: Sep. 21, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Augmented\_reality&oldid=1111308728

[11] “Pokémon GO,” *Pokémon GO*. https://pokemongolive.com/ (accessed Sep. 21, 2022).

[12] H. Soneria, “AR Decor: Decoration using markerless Augmented Reality,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 8, no. 5, p. 2145, 2020.

[13] “Houzz - Home Design & Remodel - Εφαρμογές στο Google Play.” https://play.google.com/store/apps/details?id=com.houzz.app&hl=el&gl=GR (accessed Sep. 21, 2022).

[14] F. Tscheu and D. Buhalis, “Augmented Reality at Cultural Heritage sites,” 2016, pp. 607–619. https://doi.org/10.1007/978-3-319-28231-2\_44

[15] G. Martin *et al.*, “Use of the HoloLens2 Mixed Reality Headset for Protecting Health Care Workers During the COVID-19 Pandemic: Prospective, Observational Evaluation,” *J. Med. Internet Res.*, vol. 22, no. 8, p. e21486, Aug. 2020, https://doi.org/10.2196/21486

[16] D. F. Bosché, D. M. Abdel-Wahab, and D. L. Carozza, “Towards a Mixed Reality System for Construction Trade 1 Training 2,” 2015. https://www.semanticscholar.org/paper/Towards-a-Mixed-Reality-System-for-Construction-1-2-Bosch%C3%A9-Abdel-Wahab/f3de52d5e7f998fce95cbe7be9c73c2c0ed04607 (accessed Sep. 21, 2022).

[17] Yilei Huang, Samjhana Shakya, and Temitope Odeleye, “Comparing the Functionality between Virtual Reality and Mixed Reality for Architecture and Construction Uses,” *J. Civ. Eng. Archit.*, vol. 13, no. 7, Jul. 2019, https://doi.org/10.17265/1934-7359/2019.07.001

[18] A. Muñoz and A. Martí Testón, “HOLOMUSEUM: A HOLOLENS APPLICATION FOR CREATING EXTENSIBLE AND CUSTOMIZABLE HOLOGRAPHIC EXHIBITIONS,” Jul. 2018, pp. 2303–2310. https://doi.org/10.21125/edulearn.2018.0625

[19] W. Hou, “Augmented Reality Museum Visiting Application based on the Microsoft HoloLens,” *J. Phys. Conf. Ser.*, vol. 1237, p. 052018, Jun. 2019, https://doi.org/10.1088/1742-6596/1237/5/052018

[20] T. Davis, “Homogeneous coordinates and computer graphics,” Dec. 2001.

[21] “Homogeneous coordinates,” *Wikipedia*. May 26, 2022. Accessed: Sep. 21, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Homogeneous\_coordinates&oldid=1089902994

[22] Y.-B. Jia, “Quaternions and Rotations,” p. 12.

[23] G. S. and B. Eater, “Visualizing quaternions, an explorable video series.” https://eater.net/quaternions (accessed Sep. 22, 2022).

[24] R. Saroha, R. Bala, and S. Siwach, “Review paper on Overview of Image Processing and Image Segmentation,” p. 13, 2013.

[25] D. Ziou and S. Tabbone, “Edge Detection Techniques - An Overview,” vol. 8, Jun. 2000.

[26] “How does the Microsoft HoloLens work?,” *Quora*. https://www.quora.com/How-does-the-Microsoft-HoloLens-work (accessed Sep. 22, 2022).

[27] N. A. Ismail, T. C. Wen, S. Salam, A. M. Nawi, and S. E. N. Mohamed, “A Review Of Visual Inertial Odometry For Object Tracking And Measurement,” vol. 9, no. 02, p. 7, 2020.

[28] “Visual odometry,” *Wikipedia*. Jul. 23, 2022. Accessed: Sep. 23, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Visual\_odometry&oldid=1100024244

[29] M. Maimone, Y. Cheng, and L. Matthies, “Two years of Visual Odometry on the Mars Exploration Rovers,” *J. Field Robot.*, vol. 24, no. 3, pp. 169–186, Mar. 2007, https://doi.org/10.1002/rob.20184

[30] “OpenCV: Optical Flow.” https://docs.opencv.org/3.4/db/d7f/tutorial\_js\_lucas\_kanade.html (accessed Sep. 23, 2022).

[31] Y. Wang, L. Chen, P. Wei, and X. Lu, “Visual-Inertial Odometry of Smartphone under Manhattan World,” *Remote Sens.*, vol. 12, no. 22, p. 3818, Nov. 2020, https://doi.org/10.3390/rs12223818

[32] D. Ksentini, A. R. Elhadi, and N. Lasla, “Inertial Measurement Unit: Evaluation for Indoor Positioning,” in *2014 International Conference on Advanced Networking Distributed Systems and Applications*, Bejaia, Algeria, Jun. 2014, pp. 25–30. https://doi.org/10.1109/INDS.2014.12

[33] O. J. Woodman, “An introduction to inertial navigation,” p. 37.

[34] K. Yousif, A. Bab-Hadiashar, and R. Hoseinnezhad, “An Overview to Visual Odometry and Visual SLAM: Applications to Mobile Robotics,” *Intell. Ind. Syst.*, vol. 1, no. 4, pp. 289–311, Dec. 2015, https://doi.org/10.1007/s40903-015-0032-7

[35] N. Karlsson, E. di Bernardo, J. Ostrowski, L. Goncalves, P. Pirjanian, and M. E. Munich, “The vSLAM Algorithm for Robust Localization and Mapping,” in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, Barcelona, Spain, 2005, pp. 24–29. https://doi.org/10.1109/ROBOT.2005.1570091

[36] D. Dornellas, F. Rosa, A. Bernardino, R. Ribeiro, and J. Santos-Victor, “GPS emulation via visual-inertial odometry for inspection drones,” Dec. 2019, pp. 755–760. https://doi.org/10.1109/ICAR46387.2019.8981597

[37] scooley, “HoloLens 2 hardware.” https://learn.microsoft.com/en-us/hololens/hololens2-hardware (accessed Sep. 24, 2022).

[38] Mamaylya, “HoloLens 2 gestures (for example, gaze and air tap) for navigating a guide in Dynamics 365 Guides - Dynamics 365 Mixed Reality.” https://learn.microsoft.com/en-us/dynamics365/mixed-reality/guides/operator-gestures-hl2 (accessed Sep. 24, 2022).

[39] sostel, “Eye tracking overview - Mixed Reality.” https://learn.microsoft.com/en-us/windows/mixed-reality/design/eye-tracking (accessed Sep. 24, 2022).

[40] sostel, “Eye-gaze-based interaction - Mixed Reality.” https://learn.microsoft.com/en-us/windows/mixed-reality/design/eye-gaze-interaction (accessed Sep. 24, 2022).

[41] Sean-Kerawala, “Hand tracking in Unreal - Mixed Reality.” https://learn.microsoft.com/en-us/windows/mixed-reality/develop/unreal/unreal-hand-tracking (accessed Sep. 25, 2022).

[42] scooley, “Getting around HoloLens 2.” https://learn.microsoft.com/en-us/hololens/hololens2-basic-usage (accessed Sep. 24, 2022).

[43] evmill, “Use your voice to operate HoloLens.” https://learn.microsoft.com/en-us/hololens/hololens-cortana (accessed Sep. 25, 2022).

[44] thetuvix, “Voice input in Unity - Mixed Reality.” https://learn.microsoft.com/en-us/windows/mixed-reality/develop/unity/voice-input-in-unity (accessed Sep. 25, 2022).

[45] caseymeekhof, “Point and commit with hands - Mixed Reality.” https://learn.microsoft.com/en-us/windows/mixed-reality/design/point-and-commit (accessed Sep. 25, 2022).

[46] U. Technologies, “Unity - Manual: Unity User Manual 2021.3 (LTS).” https://docs.unity3d.com/Manual/index.html (accessed Sep. 25, 2022).

[47] U. Technologies, “Unity - Manual: Order of execution for event functions.” https://docs.unity3d.com/Manual/ExecutionOrder.html (accessed Sep. 25, 2022).

[48] U. Technologies, “Unity - Manual: Plug-ins.” https://docs.unity3d.com/Manual/Plugins.html (accessed Sep. 26, 2022).

[49] “About AR Foundation | AR Foundation | 4.1.12.” https://docs.unity3d.com/Packages/com.unity.xr.arfoundation@4.1/manual/index.html (accessed Sep. 25, 2022).

[50] polar-kev, “MRTK2-Unity Developer Documentation - MRTK 2.” https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/ (accessed Sep. 25, 2022).

[51] TerryGLee, “Code editor features - Visual Studio (Windows).” https://learn.microsoft.com/en-us/visualstudio/ide/index-writing-code (accessed Sep. 26, 2022).

[52] A. Poškevičius, “Barebones Master Server.” Sep. 20, 2022. Accessed: Sep. 26, 2022. [Online]. Available: https://github.com/alvyxaz/barebones-masterserver

[53] BillWagner, “Serialization (C#).” https://learn.microsoft.com/en-us/dotnet/csharp/programming-guide/concepts/serialization/ (accessed Sep. 26, 2022).

[54] “Endianness,” *Wikipedia*. Sep. 25, 2022. Accessed: Sep. 26, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Endianness&oldid=1112276452

[55] BillWagner, “Built-in types - C# reference.” https://learn.microsoft.com/en-us/dotnet/csharp/language-reference/builtin-types/built-in-types (accessed Sep. 26, 2022).

[56] BillWagner, “Reflection (C#).” https://learn.microsoft.com/en-us/dotnet/csharp/programming-guide/concepts/reflection (accessed Sep. 26, 2022).

[57] M. Blackstock, “Aspect Weaving with C# and .NET,” Jan. 2004.