

# POLITECNICO DI TORINO

**Corso di Laurea Magistrale  
in Ingegneria Matematica**

Tesi di Laurea Magistrale

**La realtà aumentata come guida nei processi industriali di  
assemblaggio e disassemblaggio**



**Relatore**

prof. Andrea Sanna

**Candidato**

Mirko Crobu

Anno Accademico 2019-2020

# Indice

Elenco delle tabelle	4
Elenco delle figure	5
<b>1 Introduzione</b>	<b>7</b>
1.1 Contesto Applicativo . . . . .	7
1.2 Motivazioni e obiettivi . . . . .	9
1.3 Organizzazione dei capitoli . . . . .	9
<b>2 Stato dell'Arte</b>	<b>10</b>
2.1 Manutenzione e assemblaggio in ambito industriale . . . . .	10
<b>3 La Realtà Aumentata</b>	<b>18</b>
3.1 Un po' di storia . . . . .	18
3.2 Reality–virtuality continuum . . . . .	20
3.3 Funzionamento della realtà aumentata . . . . .	21
3.4 Tipologie di realtà aumentata . . . . .	23
3.5 Tipologie di display . . . . .	25
3.6 Projection mapping nei compiti di assemblaggio . . . . .	29
<b>4 Requisiti del progetto</b>	<b>30</b>
4.1 Postazione di lavoro . . . . .	31
4.2 Il sistema di proiezione e rilevamento dell'interazione . . . . .	31
4.3 Le informazioni virtuali . . . . .	32
4.4 L'interazione con l'applicazione . . . . .	32
4.5 Le procedure . . . . .	33
4.6 Il ruolo dell'operatore . . . . .	34
4.7 La disposizione degli oggetti . . . . .	34

<b>5</b>	<b>Progettazione e sviluppo</b>	<b>35</b>
5.1	Architettura hardware . . . . .	36
5.1.1	Optoma W400 . . . . .	37
5.1.2	Kinect V2 . . . . .	38
5.1.3	Il sensore di profondità e la camera RGB . . . . .	39
5.1.4	Kinect for Windows SDK 2.0 . . . . .	40
5.1.5	L'unità di elaborazione . . . . .	40
5.1.6	Composizione dell'area di lavoro . . . . .	41
5.2	Architettura software . . . . .	43
5.2.1	Analisi preliminare . . . . .	43
5.2.2	Prima fase . . . . .	44
5.2.3	Seconda fase . . . . .	44
5.3	Le aree di interazione . . . . .	45
5.4	Il rilevamento delle mani . . . . .	46
5.5	Integrazione in VRStar . . . . .	50
5.6	La procedura . . . . .	52
5.6.1	Il design . . . . .	52
5.6.2	I task . . . . .	55
5.6.3	Struttura dei task . . . . .	56
5.6.4	LogicOperator: la gestione dei task . . . . .	57
5.7	La creazione della scena virtuale . . . . .	58
5.8	La projection mapping . . . . .	59
5.9	Considerazioni generali . . . . .	61
<b>6</b>	<b>Risultati</b>	<b>62</b>
6.1	Test in fase di progettazione . . . . .	62
6.1.1	Test sull'interazione . . . . .	63
6.1.2	Il problema dei falsi positivi . . . . .	64
6.1.3	Intelligibilità delle informazioni . . . . .	64
6.2	Test finali . . . . .	66
6.2.1	Misurazione dei parametri oggettivi . . . . .	67
6.2.2	Test soggettivi . . . . .	67
6.2.3	Valutazione SUS . . . . .	75
<b>7</b>	<b>Conclusioni e sviluppi futuri</b>	<b>76</b>
	<b>Ringraziamenti</b>	<b>78</b>
	<b>Bibliografia</b>	<b>79</b>

# Elenco delle tabelle

5.1	Caratteristiche Display/immagine del Optoma W400. . . . .	37
5.2	Caratteristiche Ottico del Optoma W400. . . . .	38
5.3	Caratteristiche principali della Kinect V2. . . . .	39
5.4	Funzionamento del LogicOperator . . . . .	57
6.1	SUS: interpretazione dei punteggi. . . . .	75

# Elenco delle figure

1.1	L'applicazione pokémon Go. . . . .	8
2.1	KARMA: sovrapposizione di informazioni sintetiche. . . . .	11
2.2	Processo di scomposizione automatica del modello. . . . .	13
2.3	Tabella riassuntiva degli studi descritti. . . . .	16
3.1	Diagramma del Reality-Virtuality continuum. . . . .	20
3.2	Applicazione AR marker-based. . . . .	23
3.3	Applicazione AR marker-less. . . . .	24
3.4	Applicazione AR di tipo projection. . . . .	24
3.5	Applicazione AR di tipo superimposition. . . . .	25
3.6	Head-mounted display-Google Glass. . . . .	26
3.7	Tecnica optical see-through a sinistra, e video based see-through a destra. . . . .	26
3.8	Head-up display. . . . .	27
3.9	Projection display. . . . .	28
3.10	Projection display. . . . .	28
4.1	Schema della postazione di lavoro. . . . .	31
4.2	Informazione informativa a sinistra, indicativa al centro e interattiva a destra. . . . .	32
4.3	Esempio di procedura. . . . .	33
5.1	Schema dei componenti necessari. . . . .	36
5.2	Proiettore Optoma W400. . . . .	37
5.3	Sensori della Kinect. . . . .	38
5.4	La postazione di lavoro. . . . .	41
5.5	Disposizione della Kinect e del proiettore(a sinistra vista laterale, a destra vista dall'alto rispetto al muro). . . . .	42
5.6	Schema logico della prima fase. . . . .	44
5.7	Schema logico dell'applicazione in realtà aumentata. . . . .	45
5.8	determinazione delle coordinate di un'area di interazione. . . . .	46
5.9	Effetto dell'algoritmo al variare della distanza. . . . .	47
5.10	Le mani sono rilevate solo se aderenti al piano. . . . .	48

5.11	Pulitura dell'immagine: filtro di erosione + filtro di dilatazione.	48
5.12	Il centroide della mano dopo la pulitura dell'immagine. . . . .	49
5.13	Possibili casi in cui l'area interattiva può essere selezionata. . .	50
5.14	Il bottone "Next" è legato ad un'area interattiva. . . . .	51
5.15	In verde le aree interattive, in rosso l'area selezionata. . . . .	52
5.16	La ricostruzione in scala. . . . .	53
5.17	Design di un task. . . . .	54
5.18	Animazione in basso a sinistra e highlight dei componenti coinvolti. . . . .	54
5.19	Struttura dei task utilizzata per il progetto. . . . .	56
5.20	Gli oggetti sono stati usati come riferimento per la calibrazione.	59
5.21	Creazione della scena sullo scene editor. . . . .	60
6.1	Aree di test. . . . .	63
6.2	Asserzione 1: "mi piacerebbe utilizzare spesso questo sistema".	69
6.3	Asserzione 2: "ho trovato il sistema inutilmente complesso". . .	69
6.4	Asserzione 3: "penso che il sistema fosse facile da usare". . . .	70
6.5	Asserzione 4: "penso che avrei bisogno del supporto di un tecnico per poter usare questo sistema". . . . .	70
6.6	Asserzione 5: "ho trovato che le varie funzioni di questo sistema erano ben integrate". . . . .	71
6.7	Asserzione 6: "penso che ci sia troppa incoerenza in questo sistema". . . . .	71
6.8	Asserzione 7: "immagino che la maggior parte delle persone imparerebbe a usare questo sistema molto velocemente". . . .	73
6.9	Asserzione 8: "Ho trovato il sistema molto complicato da usare".	73
6.10	Asserzione 9: "mi sentivo molto sicuro nell'utilizzare il sistema".	74
6.11	Asserzione 10: "ho avuto bisogno di imparare molte cose prima di poter iniziare con questo sistema". . . . .	74

# Capitolo 1

## Introduzione

### 1.1 Contesto Applicativo

La realtà aumentata è una tecnologia che ci permette di arricchire le informazioni acquisite tramite i nostri sensi, con informazioni aggiuntive generate artificialmente. Già da tempo, colossi aziendali come Amazon, Google, Microsoft, Facebook, Nokia, Unity, Unreal, Samsung, impiegano importanti risorse nello sviluppo e nella ricerca riguardanti la realtà aumentata. Anche il numero di medie e piccole imprese che puntano su questa nuova tecnologia è in continuo aumento. La veloce diffusione ed evoluzione dell'AR avuta negli ultimi anni è dovuta anche alla continua evoluzione delle tecnologie hardware. Grazie alla potenza raggiunta da tablet e smartphone la realtà aumentata è oggi alla portata di tutti e ad un prezzo accessibile dalla maggior parte delle famiglie. Esistono comunque hardware dedicati in grado di rendere l'esperienza AR ancora migliore come per esempio Microsoft HoloLens, Magic Leap Lightwear, Google Glass. La realtà aumentata si avvale del supporto della computer vision e della object recognition per creare applicazioni real-time interattive dalle grande potenzialità, che vengono sfruttate nei più svariati campi:

- in architettura l'AR è stata usata per visualizzare i progetti di costruzione direttamente nell'area di lavoro prima che l'edificio sia stato costruito;
- in archeologia per aiutare i ricercatori nella ricostruzione dei siti a partire da strutture esistenti e parzialmente distrutte;
- in medicina, aiutando il personale nella formazione e nel supporto in campo chirurgico;

- nell'educazione fornendo agli studenti informazioni aggiuntive come foto, video, o simulazioni tramite modelli 3D durante la lettura di un libro;
- nel turismo sono state create applicazioni che forniscono informazioni in tempo reale sui monumenti o sui siti presi in esame;
- nel mondo dei video game, con numerosi giochi, tra cui il più famoso Pokémon Go.



Figura 1.1. L'applicazione pokémon Go.

Essa viene usata anche in campi più delicati come la difesa, la navigazione e l'aviazione.

La realtà aumentata viene sempre più utilizzata anche in campo industriale, rivoluzionando la progettazione, la produzione e la gestione del processo produttivo, imponendosi come una delle principali tecnologie artefici della quarta rivoluzione industriale che si presenta col nome di Industry 4.0. Gli operatori, attraverso questa nuova tecnologia, sono in grado di velocizzare i processi produttivi, avendo la possibilità di usufruire delle informazioni necessarie in maniera diretta, veloce ed efficiente. Gli utenti supportati da applicazioni di realtà aumentata necessitano di meno assistenza da parte di personale più esperto e i tempi per la loro formazione o il tempo necessario per i compiti di assemblaggio diminuisce notevolmente .

Tra le varie tecniche di realtà aumentata utilizzate in ambito industriale vi è la realtà aumentata spaziale, conosciuta anche col nome di projection mapping. Essa consiste nella proiezione delle informazioni direttamente sul mondo reale (quindi senza usufruire di un display per visualizzare le informazioni sintetiche sovrapposte alla scena reale). Questo è possibile grazie

all'uso di proiettori classici, proiettori laser o simili. Attraverso un sistema di acquisizione come una camera, e un sistema di manipolazione dei dati, la computer vision o la modellazione 3D è possibile adattare la proiezione alla forma e alla scala della superficie target di proiezione. Tale superficie può essere un qualsiasi oggetto di forma non regolare come un'automobile, un edificio o il corpo umano. In ambito industriale tale sistema è utilizzato per visualizzare determinate informazioni direttamente sui prodotti, macchinari o qualsiasi altra installazione presente. In tal modo, gli operatori sono assistiti nei loro compiti, guidati nelle operazioni senza che debbano distogliere lo sguardo dal loro lavoro o adoperare le mani per accedere alle informazioni tramite i mezzi classici (cartacei, display elettronici...).

## 1.2 Motivazioni e obbiettivi

Il fascino della realtà aumentata e le potenzialità che questa tecnologia può offrire mi ha spinto ad affrontare questo progetto di tesi proposto da Illogic. Inoltre la possibilità di sviluppare tale progetto all'interno di un'azienda offre un'esperienza propedeutica per il mondo del lavoro. Lo scopo di questo progetto di tesi è quello di sviluppare un prototipo in grado di supportare un operatore nei compiti di assemblaggio seguendo i principi dettati dalla projection mapping.

## 1.3 Organizzazione dei capitoli

Nel secondo capitolo si espone lo stato dell'arte della realtà aumentata applicata in ambito industriale per il supporto e la manutenzione.

Nel terzo capitolo si parla della realtà aumentata in generale: si racconta la storia, si espongono i principi di funzionamento, i pregi e i difetti.

Nel quarto capitolo si descrivono i requisiti di progetto.

Nel quinto si espongono le scelte, i passi, le implementazioni e le considerazioni che hanno portato alla creazione del prototipo.

Il sesto è espone i risultati derivanti dai test oggettivi e soggettivi a cui è stato sottoposto il prototipo.

Nel settimo capitolo sarà invece il capitolo delle conclusioni in cui i risultati vengono commentati, vengono descritti limiti e punti di forza del prototipo, nonché i possibili sviluppi futuri.

# Capitolo 2

## Stato dell'Arte

### 2.1 Manutenzione e assemblaggio in ambito industriale

Si può affermare che ai giorni nostri, nel 2020, la quarta rivoluzione industriale, conosciuta più comunemente con il termine Industry 4.0, non sia più solo materia di discussione ma è diventata realtà. Le nuove tecnologie che caratterizzano questo fenomeno sono ormai mature e vengono utilizzate sempre più frequentemente nell'industria. Grazie all'avvento dell'Industry 4.0, il processo produttivo industriale è stato rivoluzionato; la figura del lavoratore è cambiata e il lavoro è diventato più efficiente sotto molti punti di vista: velocità di produzione, ottimizzazione e sicurezza per citarne alcuni. Una delle principali tecnologie protagoniste della Industry 4.0, è la realtà aumentata. I risultati raggiunti oggi sono frutto di un lungo percorso iniziato a fine degli anni sessanta, ma che solo agli inizi degli anni novanta possiamo vedere tale tecnologia applicata nell'industria come supporto nei compiti di assemblaggio e manutenzione. Nel 1992 due ricercatori della Boeing, Thomas Caudell e D.W. Mizell, usarono per la prima volta il termine "Augmented Reality". La compagnia chiese loro di sostituire le tavole di compensato che raffiguravano le istruzioni di cablaggio con un Head-mounted display in grado di proiettare direttamente sul pavimento le informazioni in forma digitale. Le tavole di compensato erano costose da gestire e richiedevano frequenti e complesse modifiche. La loro conversione in forma digitale permise di abbassare i costi e velocizzare i tempi di modifica delle informazioni[1]. Nel 1993, i ricercatori della Columbia University Steven Feiner, Blair MacIntyre e Dorée Seligmann prototiparono KARMA: il suo scopo era quello assistere l'utente nel processo

di assemblaggio di una stampante. attraverso un un caschetto “see-through”, il sistema sovrapponeva la procedura di rimozione di un cassetto della stampante direttamente sull’oggetto fisico. Per visualizzare le informazioni nel punto esatto, il cassetto della stampante è stato fornito di marker colorati[2].

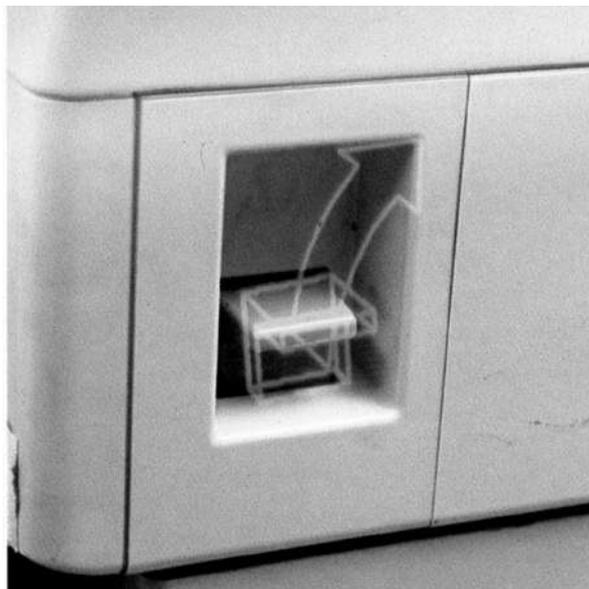


Figura 2.1. KARMA: sovrapposizione di informazioni sintetiche.

Nel 1996, Jun Rekimoto e Yuji Ayatsuka, ricercatori dei Sony Computer Science Laboratories, crearono un sistema innovativo che prevedeva l’utilizzo di marker per il riconoscimento e la determinazione della posizione 3D di oggetti reali attraverso una camera. Il loro lavoro fu preso come modello per tutti i seguenti lavori di realtà aumentata basata su marker[3]. La tecnica dei marker viene largamente usata in quegli anni, la maggiore precisione fornita lato computer vision, ha incoraggiato lo sviluppo di applicazioni in realtà aumentata[5]. Nella seconda metà degli anni novanta l’importanza di tecnologie basate su realtà aumentata, realtà virtuale e loro varianti, aumenta notevolmente: nel 1997 la Cannon Inc. assieme al governo giapponese avviarono uno dei più importanti progetti dell’epoca istituendo il Mixed Reality Systems Laboratory per la ricerca di tali tecnologie in ambito industriale. Nel 2002 un altro importante gruppo, finanziato dal governo tedesco, dà il via al progetto ARVIKA il cui scopo è lo sviluppo e la ricerca di applicazioni di realtà aumentata utili per la produzione e l’assistenza nel settore automobilistico e aerospaziale. Siamo ormai negli anni in cui si punta su tali tecnologie

per rivoluzionare l'industria del futuro. Sempre più, grandi e piccole imprese, istituti di ricerca e università di tutto il mondo si occupano di realtà aumentata in tutte le loro varianti. Complici di tale diffusione, sono anche i risultati ottenuti dagli studi atti a capire quanto la realtà aumentata possa incidere positivamente e quanto sia realmente vantaggiosa come strumento di supporto nei compiti di assemblaggio rispetto ai metodi tradizionali. Nel 1999 A.C. Boud, D.J. Haniff, C. Baber, and S.J. Steiner dell'università di Birmingham eseguirono un esperimento per cercare di capire le potenzialità e le limitazioni che tale strumento possa offrire. Essi effettuarono test su gruppi composti da cinque persone. Ogni gruppo è stato dotato di una tecnologia diversa che permetteva ai tester di usufruire delle informazioni utili per il compito di assemblaggio: sistemi tradizionali per il primo gruppo, sistema di Realtà Virtuale desktop con mouse 3D per spostarsi nella scena sintetica per il secondo gruppo e un sistema di realtà aumentata fornito da un caschetto montato sulla testa per terzo gruppo. I tester hanno trovato scomodo il sistema di realtà virtuale perchè potevano lavorare con una sola mano, mentre l'altra mano era dedicata all'interazione con la scena sintetica attraverso il mouse 3D. Inoltre non avendo un feedback tattile durante lo svolgimento dei task di assemblaggio, l'interazione risultava scomoda. C'è stato invece un riscontro positivo per i tester che hanno usato la tecnologia di realtà aumentata per svolgere il compito, evidenziando i tempi più bassi di completamento di tutti i task rispetto ai primi due gruppi[4]. Nel 2002 Arthur Tang, Charles Owen, Frank Biocca e Weimin Mou della Michigan State University hanno condotto uno studio affidando ad un gruppo di persone un compito di assemblaggio seguendo istruzioni cartacee e confrontando i risultati ottenuti con un secondo gruppo supportato dalla realtà aumentata per eseguire il medesimo compito. Il risultato ottenuto dai ricercatori ha evidenziato che la sovrapposizione delle istruzioni digitali tramite realtà aumentata riduce il tasso di errore del 82% e lo sforzo mentale diminuisce in quanto viene scaricato sull'elaboratore[6]. La riduzione del tasso di errore e dei tempi di lavoro sono una costante nei risultati degli studi di questi anni. Ma non solo, si è potuto appurare che gli operatori non trovano troppa difficoltà nell'apprendere l'uso di questa tecnologia per svolgere i compiti di assemblaggio[7][8]. I ricercatori della King Mongkut's University of Technology in Thailandia Nattapol Pathomaree e Siam Charoenseang, hanno condotto uno studio per capire se la realtà aumentata usata come strumento per apprendere le procedure di assemblaggio risulti efficace. I risultati dei loro studi sono stati molto positivi evidenziando un apprendimento più veloce per gli operatori che imparavano la procedura attraverso il visore[9].

Già dai primi anni del 2010, la realtà aumentata è ormai considerata una tecnologia matura. Grazie al potenziamento dell'hardware e della computer vision, aumenta anche l'affidabilità degli algoritmi di tracking. Acquistano più visibilità le applicazioni di realtà aumentata che non si servono di marker per individuare i punti di proiezione delle informazioni sintetiche sui display. Nel 2013 Hugo Alvarez, Iker Aguinaga, Diego Borro, provenienti dall'Università di Navarra, creano un framework finalizzato al riconoscimento degli oggetti attraverso il loro modello 3D memorizzato nel sistema come insieme di mesh triangolari non texturizzate che descrivevano ogni parte dell'oggetto preso in esame[10]. Ad arricchire il mondo della realtà aumentata sono

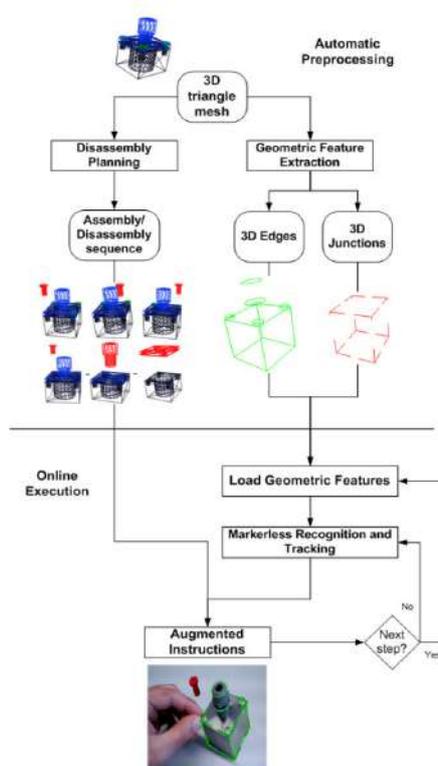


Figura 2.2. Processo di scomposizione automatica del modello.

anche i dispositivi mobili come smartphone e tablet. Essendo sempre più potenti e grazie a risoluzioni del display e della camera maggiori, sono ora idonei per lo sviluppo di applicazioni in realtà aumentata per l'industria. Nel 2015 Il Professore A. Sanna assieme ai suoi collaboratori F. Manuri, F. Lamberti, G. Paravati, P. Pezzolla condussero uno studio finalizzato a capire quanto i mobile device potessero essere utili come supporto agli operatori

nei compiti di manutenzione e assemblaggio. Il loro lavoro aveva come scopo anche quello di capire in che misura tale tecnologia potesse essere utile per persone considerate non del settore. Il loro studio ha evidenziato che dispositivi mobili come tablet e smartphone possono essere uno strumento utile per lo sviluppo di applicazioni di realtà aumentata per il supporto nei compiti di assemblaggio e manutenzione. La ricerca rivelò anche che a trarre maggior beneficio da tali applicativi sono le persone con meno esperienza in tali compiti, in cui si vede una maggiore riduzione dell'errore e risparmio di tempo[11]. In questi ultimi anni si cominciano anche gli studi che prevedono l'uso di intelligenza artificiale per effettuare la parte di object recognition nelle applicazioni di realtà aumentata. Questa tecnologia è sempre più usata ed è in forte crescita. Nel 2018 i ricercatori Xuyue Yin, Xiumin Fan, Jiajie Wang, Rui Liu and Qiang Wang della Shanghai Jiao Tong University, hanno testato le prestazioni di una deep neural network Faster R-CNN. La Faster R-CNN è una delle reti più avanzate utilizzate per il riconoscimento di oggetti. Eseguita la fase di addestramento della rete, i ricercatori hanno sviluppato un'applicazione in realtà aumentata per svolgere un compito di assemblaggio usando come parte di object recognition la Faster R-CNN. La rete è stata usata per il riconoscimento di connettori elettrici aerospaziali, categorizzati come oggetti di piccola taglia. Il test ha avuto ottimi risultati, evidenziando una precisione del 94% della rete e una velocità media di 200 millisecondi soddisfacendo i requisiti di un'applicazione real-time.

Nella seguente tabella sono riassunte le informazioni riguardanti gli studi appena presentati:

Anno	Titolo	Contesto	Hardware utilizzato	Caso di studio	Risultati
1992	Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes[1]	Aerospaziale	Display head-mounted	manutenzione assemblaggio	-fruibilità delle informazioni migliore. -mani libere durante l'accesso alle informazioni.
1993	KARMA: Knowledge-based Augmented Reality for Maintenance Assistance[2]	Columbia University Computer Graphics and User Interfaces Lab	Display head-mounted "see-through"	manutenzione assemblaggio	-mani libere durante l'accesso alle informazioni.

1996	CyberCode: Designing Augmented Reality Environments with Visual Tags[3]	Interaction Laboratory Sony Computer Science Laboratories	Display "see-through"	Sviluppo sistema per il riconoscimento di oggetti e la loro posizione nello spazio 3D	-guida sui principi necessari per l'uso di marker nella Realtà aumentata.
1999	Virtual Reality and Augmented Reality as a Training Tool for Assembly Tasks[4]	University of Birmingham The School of Manufacturing and Mechanical Engineering	Display head-mounted + mouse 3D Context-free AR	Valutazione performance e confronto tra l'utilizzo di AR, VR e istruzioni cartacee 2D in compiti di assemblaggio	RV e AR migliorano le prestazioni rispetto ai metodi tradizionali
2001	Computer Vision for Guiding Manual Assembly[5]	Computer Science and Engineering, Pennsylvania State University	Display head-mounted	Sviluppo framework per il rilevamento della scena di assemblaggio attraverso la computer vision	sviluppo di un framework per il rilevamento di oggetti in un contesto di assemblaggio tramite AR e marker
2002	Experimental Evaluation of Augmented Reality in Object Assembly Task[6]	Michigan State University	Display head-mounted VR full imersion System	confronto tra metodi tradizionali e metodi che prevedono la realtà aumentata per svolgere compiti di assemblaggio.	-riduzione tasso di errore durante il compito di assemblaggio; -meno sforzo mentale da parte dei soggetti
2002	AR-based training and support of assembly workers in automobile industry[7]	University of Paderborn	Display head-mounted	Sviluppo di un prototipo che sfrutta la realtà aumentata per l'industria dell'automobile	-riduzione tasso di errore - ottimizzazione del sistema produttivo
2002	Augmented Reality (AR) for Assembly Processes - An Experimental Evaluation [8]	Aachen University of Technology	Display head-mounted	valutazione efficacia dei sistemi di realtà aumentata per compiti di assemblaggio	- buon adattamento alla nuova tecnologia con tempi di apprendimento bassi

2005	Augmented Reality for Skill Transfer in Assembly Task[9]	King Mongkut's University of Technology Thonburi	Display head-mounted	valutazione efficacia dei sistemi di realtà aumentata come sistema di apprendimento di procedure di assemblaggio	- test positivi, altra percentuale di Skill Transfer - tester soddisfatti e a loro agio con tale tecnologia - tempi di esecuzione ridotti
2013	Providing Guidance for Maintenance Operations Using Automatic Markerless Augmented Reality System[10]	CEIT and Tecnun University of Navarra, Spain	non specificato	Sviluppo algoritmo per il tracking degli oggetti senza l'utilizzo di marker	sviluppo framework completo che viene eseguito in tempo reale e tenta di facilitare il lavoro dei lavoratori, sostituendo la documentazione e cartacea e senza l'uso di marker
2015	Using Handheld Devices to Support Augmented Reality-based Maintenance and Assembly Tasks[11]	Dipartimento di Automatica e Informatica, Politecnico di Torino	tablet	valutazione efficacia di sistemi di realtà aumentata su dispositivi mobili come sistema di apprendimento di procedure di assemblaggio	-maggiori benefici in termini sia di probabilità d'errore che di risparmio di tempo per soggetti che non conoscono la procedura di assemblaggio
2018	An automatic interaction method using part recognition based on deep network for augmented reality assembly guidance[12]	Shanghai Jiao Tong University	Display head-mounted	metodo di riconoscimento di oggetti tramite una deep neural network	ottime prestazioni della rete neurale sia in termini di tempo che di accuratezza.

Figura 2.3. Tabella riassuntiva degli studi descritti.

Come si può notare da questa analisi, la realtà aumentata porta grandi vantaggi nel processo produttivo industriale. Essa offre un ottimo supporto per le operazioni di assemblaggio, manutenzione e addestramento per gli operatori. Gli utilizzatori hanno bisogno di una fase di familiarizzazione con la nuova tecnologia, ma che in generale è ben accolta dagli utenti novizi della realtà aumentata. Anche se questa prima fase di apprendimento porta ad un rallentamento del processo produttivo si hanno comunque dei vantaggi in seguito: come si è visto, i tempi di produzione diminuiscono, così come il tasso di errore e lo sforzo mentale. C'è da dire che applicazioni di questo genere necessitano di essere progettate al meglio per poter beneficiare dei loro vantaggi: un'interazione immediata, semplice e intuitiva sono sicuramente requisiti fondamentali. L'importantissimo vantaggio di poter operare con entrambe le mani libere ha spinto i ricercatori e le imprese a puntare maggiormente sugli head-mounted display. Si sono comunque rivelati ottimi dispositivi anche tablet e smartphone. Per quanto si siano raggiunti ottimi risultati in ambito industriale, la realtà aumentata presenta comunque degli svantaggi: le persone sottoposte ai test hanno sottolineato il peso di questi device e il fatto che il campo visivo è ridotto creando un certo disagio. E' doveroso sottolineare che c'è bisogno di fare passi avanti anche per quanto riguarda i tempi di reattività, precisione nella object recognition e visualizzazione dei contenuti. Sono queste quindi le sfide del futuro che sicuramente potremo apprezzare nei prossimi anni.

## Capitolo 3

# La Realtà Aumentata

### 3.1 Un po' di storia

Come molte delle tecnologie presenti oggi, anche la realtà aumentata è stata predetta da un'opera di fantasia. Nel libro *The Master Key: an Electrical Fairy Tale* pubblicato nel 1901, Frank L. Baum racconta la storia di un ragazzino che trova un paio di occhiali tecnologici in grado di rivelare aspetti caratteristici della personalità di un individuo. Il ragazzino, indossando gli occhiali, vedrà tali caratteristiche descritte direttamente sulla fronte delle persone. Possiamo considerare “The Sword of Damocles” il primo dispositivo di realtà aumentata mai inventato. Parliamo del 1968 quando un professore dell'università di Harvard (Ivan Sutherland) assieme ad un suo studente (Bob Sproull) svilupparono una nuova tecnologia utilizzando un head-mounted display con il quale un utente poteva vedere semplici forme geometriche sovrapposte alla scena reale. Seguirono altre ricerche e studi in merito, ma solo nei primi anni 90 la realtà aumentata iniziò ad essere adoperata in ambito industriale. Il termine “Augmented Reality” è stato usato per la prima volta da Thomas P. Caudell, ricercatore della Boeing, per lo sviluppo di un'applicazione in grado di dare supporto agli operai nei compiti di assemblaggio[1]. In questi anni, la realtà aumentata è ancora materia da laboratori di ricerca e viene usata soprattutto in ambito industriale, medico e militare. Bisogna aspettare il 1999 per vedere la realtà aumentata divenire accessibile a tutti: il professor Hirokazu Kato rilascia la libreria open-source ARToolKit. Da questo momento la tecnologia entra anche nel mondo dell'intrattenimento. Nei primi anni duemila inizia una diffusione della realtà aumentata. Essa inizia ad essere usata per grandi eventi: nella stagione NFL del 2003, Sportvision mostra al pubblico un sistema capace di evidenziare la prima e la decima

linea di campo nel video proposto dalla famosa Skycam. L'IKEA propone un'applicazione in grado di visualizzare gli articoli del proprio catalogo sistemati a piacere del cliente all'interno della proprio alloggio. Gli applicativi nati e sviluppati in quegli anni, hanno spinto le grandi aziende a creare dispositivi dedicati per la realtà virtuale. Possiamo assistere nel 2013 alla nascita dei Google Glass. Gli occhiali della Google trasformarono la fantascienza in realtà. Il dispositivo era in grado di mostrare direttamente sulle lenti immagini e video, la cui qualità era simile a quella offerta da un display di 25 pollici in HD ad una distanza di circa 2 metri. Il prezzo di lancio del dispositivo si aggirava intorno ai 1500 \$ e offriva una connessione diretta con lo smartphone via Bluetooth. Attraverso tale interazione gli occhiali potevano usufruire del GPS del telefono. Ma il prezzo ancora alto per tale dispositivo ha impedito una larga diffusione. Nel 2016 si assiste ad un vero e proprio boom della realtà aumentata: sono presenti in commercio ormai diversi visori a prezzi molto più accessibili rispetto ai Google Glass. I principali sono sicuramente gli Oculus Rift sviluppati da Oculus VR e gli HoloLens sviluppati da Microsoft. I due dispositivi sono in grado di creare rispettivamente immagini stereoscopiche e immagini 3D. Se nel 2015 poteva esserci ancora qualcuno che non conoscesse il termine "realtà aumentata", nel 2016 grazie ai nuovi visori e soprattutto all'applicazione Pokemon Go, tutti quanti sentono parlare di questa nuova tecnologia. Pokemon Go è un'app sviluppata per smartphone Android e IOS che si basa sulla realtà aumentata geolocalizzata attraverso il GPS e sviluppato da Niantic. Il gioco ha avuto enorme successo, diventando l'applicazione più scaricata dall'App Store di Apple durante la sua prima settimana di pubblicazione. Pokemon Go ha infranto anche altri due record: per un periodo è stato il gioco con più player online attivi contemporaneamente e l'applicazione più usata, battendo addirittura social network come Facebook e Twitter. Per quanto riguarda il futuro, le previsioni sono positive e si prevede una sempre maggiore diffusione e uso di tale tecnologia nei più svariati settori. L'azienda israeliana Lumus, ha pubblicato una ricerca in cui afferma che i videogiochi in realtà virtuale potrebbero raggiungere un fatturato di 11,6 miliardi di dollari entro il 2025. Nel settore sanitario si arriverebbe a 5,1 miliardi, 4,7 miliardi di dollari nel settore ingegneristico.

## 3.2 Reality–virtuality continuum

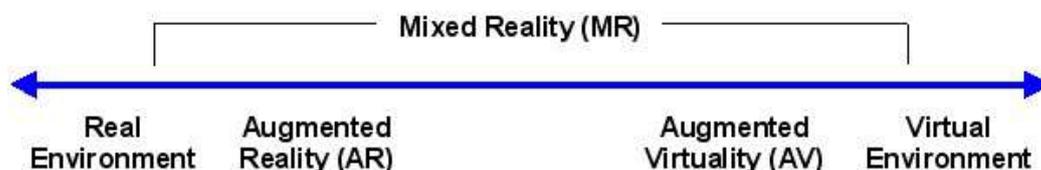


Figura 3.1. Diagramma del Reality-Virtuality continuum.

Il “Reality-Virtuality continuum” è un concetto espresso da Paul Milgram in una pubblicazione nel 1994 (“Augmented Reality: A class of displays on the reality-virtuality continuum.”). Con tale nozione il professore ci aiuta a dare una definizione e una classificazione migliore della realtà aumentata tra le varie tecnologie ad essa associate. Come si può notare dal grafico sopra esposto, tra il mondo reale e il mondo completamente virtuale non vi è una sottile linea di confine, ma vi sono degli stadi intermedi rappresentati appunto dalla realtà aumentata e dalla virtualità aumentata. La realtà virtuale indica un mondo completamente sintetico privo di elementi reali. L’ augmented reality e la augmented virtuality sono due tecnologie che prevedono la sovrapposizione di input virtuali e reali dando vita alla “mixed reality”. Augmented virtuality prevede un mondo virtuale in cui sono sovrapposte informazioni reali. Un esempio potrebbe essere la visualizzazione delle proprie mani in un mondo completamente virtuale. Vale invece l’inverso per la realtà aumentata: sono le informazioni virtuali ad essere sovrapposte nel mondo reale.

### 3.3 Funzionamento della realtà aumentata

Lo scopo della realtà aumentata, a differenza della realtà virtuale, non è quello di convincere l'utente di essere realmente presente nel mondo sintetico. Nella realtà aumentata l'utente è consapevole di essere presente nel mondo reale che è "aumentato" attraverso informazioni sintetiche. Tali informazioni possono essere essenzialmente di 5 tipi e possono essere utilizzate singolarmente o simultaneamente:

- immagini;
- animazioni;
- video;
- modelli 3D;
- audio;

Tali informazioni devono essere collocate con criterio nel campo visivo dell'utente. La AR può essere visualizzata in modi diversi, attraverso head-mounted display, dispositivi portatili come tablet e smartphone, superfici reali trasformate in display, ologrammi e altre varie tecniche sperimentali ancora relegate all'ambito di ricerca. A seconda dei casi quindi, anche il procedimento di creazione della realtà aumentata sarà diverso. Le tecnologie principali usate per creare applicazioni di questo genere sono:

1. **localizzazione simultanea e mappatura(SLAM)**: è una tecnica nata nell'ambito della robotica che permette la determinazione della posizione del dispositivo mentre viene rilevato l'ambiente circostante;
2. **rilevamento della profondità**: meccanismo con il quale si calcola la distanza tra un dispositivo e un determinato oggetto reale;
3. **camera e sensori**: questi componenti sono importanti per rilevare le interazioni degli utenti e catturare tutte le informazioni sul mondo circostante come luminosità, posizione degli oggetti in relazione alle coordinate del dispositivo, mappatura del mondo reale ecc...;
4. **capacità di processamento**: le applicazioni in AR hanno bisogno di unità di processamento quale CPU, RAM, GPS, WiFi e tutte le componenti comuni ad un vero e proprio elaboratore. Questo perché hanno bisogno di elaborare una certa quantità di informazioni, non solo in input,

ma anche in output in modo da poter fornire all'utente le informazioni sovrapposte al mondo reale;

5. **Proiezione:** è la tecnologia usata che, a seguito dell'acquisizione dei dati e dell'elaborazione, fornisce all'utente le informazioni sintetiche. Questo può avvenire attraverso auricolari, o display come nel caso di smartphone e tablet, oppure su qualsiasi superficie idonea nel caso della projection mapping;

### 3.4 Tipologie di realtà aumentata

Possiamo individuare principalmente 4 tipi di realtà aumentata. Ogni variante usa tecniche diverse, sebbene il fine ultimo resta sempre quello di sovrapporre informazioni sintetiche alle informazioni provenienti dal mondo reale.

- **Marker-Based:** Sfrutta tecniche di image recognition provenienti dalla computer vision per riconoscere il marker posizionato nel mondo reale. Un marker non è altro che un'immagine familiare all'applicazione di AR. Essa può essere un QR code, una forma geometrica (avente un particolare colore atto a marcare il contrasto di tale figura), una classica immagine. L'applicazione conosce tutte le caratteristiche dell'immagine: forma, colori, dimensioni. Grazie a queste informazioni, il software attraverso la camera potrà individuare il marker nel mondo reale, riuscendo a calcolare orientamento e distanza dal dispositivo. In questo modo è possibile renderizzare immagini o i modelli 3D con le giuste proporzioni, in una specifica posizione e con l'orientamento corretto.



Figura 3.2. Applicazione AR marker-based.

- **Marker-less:** attraverso un GPS, una bussola, un giroscopio e l'accelerometro, l'applicativo in AR è in grado di determinare la posizione dell'utente nello spazio reale. Risulta quindi possibile posizionare il contenuto virtuale in una determinata area dello spazio. Grazie a questa tecnologia è possibile creare navigatori in AR, applicazioni che forniscono informazioni su negozi o aree archeologiche solamente inquadrando la città attraverso il dispositivo.



Figura 3.3. Applicazione AR marker-less.

- **Projection Mapping:** attraverso un proiettore, le informazioni sintetiche vengono proiettate su qualsiasi superficie che può diventare uno schermo interattivo. La parte di computer vision gioca un ruolo importante nel caso della interattività.



Figura 3.4. Applicazione AR di tipo projection.

- **Superimposition Based:** fornisce una vista "alternativa" dell'oggetto target aumentandolo o sostituendolo completamente con una sua versione che mostra informazioni non recepibili a occhio nudo. Anche in questo contesto la Computer Vision gioca un ruolo importante: è essenziale il riconoscimento dell'oggetto per poterlo "aumentare".



Figura 3.5. Applicazione AR di tipo superimposition.

### 3.5 Tipologie di display

I display usati per le applicazioni di realtà aumentata possono essere di vario genere, aventi caratteristiche diverse tra loro. La scelta del display è un fattore fondamentale per la qualità di un'applicazione in AR.

- **Head-Mounted Displays:** Head-Mounted Displays: ne esistono principalmente due categorie. Gli optical see-through sono provvisti di due display trasparenti posti di fronte agli occhi attraverso il quale l'utente può vedere il mondo reale. Attraverso un sistema header tracking è possibile collocare le informazioni sintetiche in determinate posizioni nello spazio. Attraverso un monitor le immagini verranno proiettate sui piccoli display posti davanti agli occhi, creando quindi una sovrapposizione di informazioni del tipo virtuale-reale. I video based see-through funzionano con lo stesso principio di header tracking per collocare le informazioni sintetiche nelle giuste coordinate spaziali, ma il display in questo caso è un vero e proprio monitor che combina le informazioni sintetiche con le immagini provenienti dal mondo reale catturate attraverso una camera.



Figura 3.6. Head-mounted display-Google Glass.

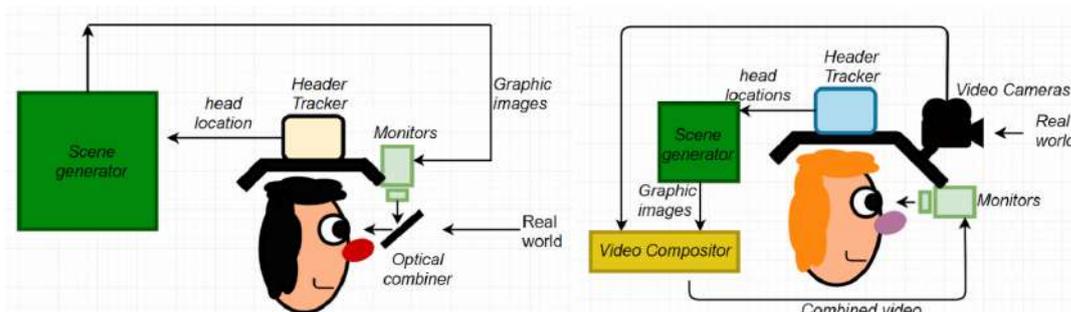


Figura 3.7. Tecnica optical see-through a sinistra, e video based see-through a destra.

- **Head-up displays:** sono composti da un display trasparente posto tra l'utente e la scena reale osservata. Il display non è necessariamente legato ai movimenti del corpo dell'utente come nel caso degli HMD. Nati nel mondo dell'aviazione, sono molto diffusi in campo automobilistico dove vengono utilizzati per mostrare al guidatore informazioni utili per la guida direttamente sul parabrezza in modo che l'utente non sia costretto a distogliere lo sguardo dalla strada.
- **Hand-Held Display:** i dispositivi più rappresentativi di questa categoria sono sicuramente tablet e smartphone. Si tratta di device che permettono di visualizzare un mondo in AR a discrezione dell'utente. Esso ha infatti il controllo manuale del display, potendo quindi controllare non solo l'orientamento e il punto di vista del mondo reale aumentato,



Figura 3.8. Head-up display.

ma anche la sua posizione all'interno del proprio campo visivo. Attraverso questi tipi di display l'utente può facilmente "passare" dal mondo aumentato al mondo reale.

- **Spatial Displays:** la scena reale da aumentare diventa essa stessa il display. Attraverso un proiettore le informazioni vengono proiettate direttamente su qualsiasi superficie piana o irregolare. Questo è possibile grazie a particolari tecniche di proiezione che distorcono l'immagine al punto da poter compensare le irregolarità presentate dal piano di proiezione.



Figura 3.9. Projection display.



Figura 3.10. Projection display.

## 3.6 Projection mapping nei compiti di assemblaggio

Nel capitolo precedente abbiamo analizzato come la realtà aumentata si sia evoluta in ambito industriale e come si sia pian piano affermata fino a diventare una preziosa risorsa per le aziende. Dai vari studi sono però emersi due importanti svantaggi: il peso dei device e il ridotto campo visivo. La *spatial augmented reality*, conosciuta più comunemente come *projection mapping*, risolve questi due problemi: infatti questa variante prevede che le informazioni sintetiche non vengano visualizzate attraverso un display *see-through*, ma proiettate direttamente sull'oggetto target, adattando video o immagini a superfici irregolari. Il piano di proiezione viene trasformato in un display interattivo. L'operatore è libero dal peso dei devices potendo inoltre, sfruttare a pieno il proprio campo visivo. La *projection mapping* è largamente utilizzata nel campo dello spettacolo, e ci sono veramente poche pubblicazioni riguardo al suo utilizzo nel campo industriale come strumento di supporto per l'assemblaggio, il training e la manutenzione. Nel 2015 i ricercatori Leonardo Rodrigueza, Fabian Quintb, Dominic Gorecky, David Romeroc, Héctor R. Sillerc della Universidad Autónoma de Nuevo León in Messico, hanno affrontato il tema, sviluppando un prototipo e verificandone i risultati. Per il loro compito si sono serviti di un proiettore e una base di supporto che lo tenesse ad una distanza fissa dal piano di lavoro, un laptop per i calcoli necessari e una linea di dispositivi di rilevamento automatico del movimento delle mani. Il processo di assemblaggio prevedeva 13 step da attuare su un porta biglietti da visita. I ricercatori, esaminando i risultati, hanno potuto affermare che tale tecnologia adempia al suo compito con successo, e che sia in grado di migliorare i processi d'industriali. Dal loro lavoro sono però emerse alcune limitazioni:

- le superfici riflettenti riducono drasticamente la qualità delle informazioni proiettate;
- un alto grado del field of view riduce la qualità delle immagini;
- un elevato angolo di inclinazione del proiettore rispetto al piano di proiezione rende le informazioni distorte.

Nonostante queste limitazioni, i vantaggi offerti dalla *projection mapping* hanno spinto le aziende a investire su tale tecnologia nell'ambito dell'assemblaggio, manutenzione e training per l'industria.

## Capitolo 4

# Requisiti del progetto

Lo scopo del progetto di tesi è quello di progettare e realizzare un'applicazione in realtà aumentata in grado di guidare un operatore nelle operazioni di assemblaggio in una determinata linea di produzione industriale. Al giorno d'oggi molte operazioni legate alla linea di produzione devono ancora essere eseguite manualmente dagli operatori. In tale contesto, i fattori che possono rallentare la linea di produzione e causare un aumento dei costi produttivi sono molteplici: l'operatore necessita di essere formato su tali procedure. Durante il periodo di apprendimento la produzione è bloccata. Un altro fattore molto importante è rappresentato dall'errore umano che in caso di lavori manuali può essere rilevante per i tempi di produzione. Altri fattori sono legati alle grandi quantità di informazioni che devono essere consultate dagli operatori in procedure molto complesse. I tempi di consultazione delle informazioni servite tramite materiale cartaceo o altri sistemi classici non solo risultano spesso inefficienti e lenti, ma apportano all'operatore anche un alto livello di sforzo mentale che si ripercuote, in un secondo momento, sulle prestazioni durante le fasi della procedura. I motivi sopra elencati sono serviti da spunto per lo sviluppo di questo progetto di tesi. L'applicazione si deve servire di tecniche di projection mapping per perseguire il suo scopo. I motivi che hanno portato a questa scelta derivano dai vantaggi offerti dalla projection mapping. L'operatore sarà in grado di usufruire del supporto della realtà aumentata senza dovere rinunciare all'uso delle mani nel caso di dispositivi hand-held, o a operare con un campo visivo ridotto nel caso di un head-mounted display. L'applicazione mira al miglioramento sia dei tempi di produzione e della diminuzione dell'errore umano durante le fasi di assemblaggio. Per raggiungere tale scopo l'applicazione deve essere in grado di fornire all'operatore informazioni in tempo reale sui compiti da

svolgere. Essa deve fornire passo passo le istruzioni necessarie all'utente per l'avanzamento nel processo di assemblaggio e guidarlo nel riconoscere le varie parti descritte dalle istruzioni.

## 4.1 Postazione di lavoro

La postazione di lavoro deve prevedere un piano di lavoro su cui l'operatore può eseguire la procedura di assemblaggio su un oggetto. Tale piano deve avere una base di larghezza pari a 1,60 m e un'altezza pari a 60 cm. All'interno dell'aria descritta dal piano di lavoro devono essere proiettate tutte le informazioni sintetiche rivolte all'operatore.

## 4.2 Il sistema di proiezione e rilevamento dell'interazione

Il sistema di proiezione deve essere posto ad un'altezza di almeno 1,50 m al di sopra del piano di lavoro. La camera di profondità e la camera RGB devono essere poste ad un'altezza tale da non invadere il campo d'azione dell'operatore. Il sistema di elaborazione delle informazioni, rappresentato da un calcolatore, non deve stare sul piano di lavoro.

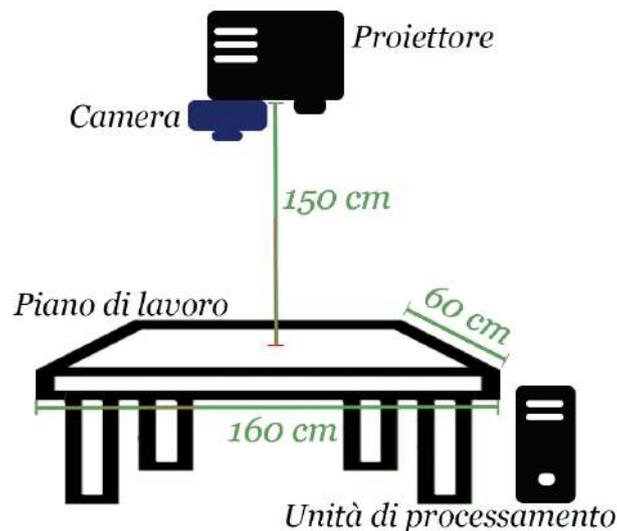


Figura 4.1. Schema della postazione di lavoro.

### 4.3 Le informazioni virtuali

Le informazioni sintetiche devono essere proiettate dal sistema direttamente sul piano di lavoro o sugli oggetti presi in esame per la procedura. Le figure proiettate dall'applicazione possono essere di tre tipi:

- informative: sono istruzioni testuali, immagini o animazioni che forniscono all'operatore istruzioni o spiegazioni riguardo i task;
- indicative: sono forme che segnalano all'operatore i componenti descritti dalle figure informative;
- interattive: sono figure, come per esempio bottoni, che l'operatore può usare per interagire con il sistema.



Figura 4.2. Informazione informativa a sinistra, indicativa al centro e interattiva a destra.

### 4.4 L'interazione con l'applicazione

Deve essere presente un sistema di rilevamento dell'operato dell'utente. In particolare l'operatore deve essere in grado di interagire con l'applicativo attraverso bottoni proiettati sul piano. Le aree di interazione possono avere dimensioni diverse e regolabili a seconda dell'istruzione da eseguire. Esse possono essere posizionate su qualunque parte del piano di lavoro, comprese aree irregolari appartenenti all'oggetto o alle sue componenti.

## 4.5 Le procedure

Le procedure sono suddivise in diversi task. Per ogni task il sistema deve fornire all'operatore un pannello informativo contenente un testo descrittivo dell'operazione da svolgere, gli indicatori necessari per aiutare l'operatore ad individuare le parti utilizzate durante la procedura, una o più animazioni se ritenute necessarie, un bottone per la conferma di completamento del task da parte dell'operatore. I task devono seguire sempre lo stesso schema:

1. caricamento dell'istruzione;
2. proiezione della descrizione dell'operazione;
3. highlight delle componenti protagoniste;
4. proiezione di eventuali animazioni esplicative;
5. l'operatore esegue l'operazione richiesta;
6. l'operatore seleziona il tasto "next" per confermare l'avvenuta operazione;
7. caricamento dell'istruzione successiva.



Figura 4.3. Esempio di procedura.

## 4.6 Il ruolo dell'operatore

L'operatore deve essere in grado di operare libero da qualunque sorta di dispositivo indossabile. Non deve essere provvisto di guanti colorati o altri tipi di marker per poter interagire con l'applicazione. Secondo i principi della realtà aumentata, deve essere in grado di visualizzare le informazioni direttamente sul piano di lavoro o sovrapposte agli oggetti presenti.

## 4.7 La disposizione degli oggetti

Gli oggetti su cui devono essere eseguite le procedure di assemblaggio devono avere una posizione prefissata al momento della creazione della procedura, in modo da fornire all'operatore una configurazione iniziale che sia la migliore possibile per poter operare.

## Capitolo 5

# Progettazione e sviluppo

In questo capitolo sono descritte le scelte implementative seguite per lo sviluppo di questo progetto. Si rimanda al capitolo 6 per l'analisi dei limiti, dei problemi e i possibili sviluppi futuri riguardanti questo progetto di tesi.

Il primo step per lo sviluppo dell'applicazione è stato quello di studiare i requisiti del progetto in modo da poter determinare l'architettura hardware e software dell'applicazione. Per quanto riguarda l'architettura software si è scelto di seguire la struttura offerta dal game engine proprietario di Illogic: VRStar. In questo modo è stato possibile usufruire delle funzionalità presenti nella piattaforma e integrare questo lavoro come componente stesso del software aziendale. La scelta dell'hardware utilizzato è ricaduta sull'uso della Kinect V2, il proiettore Optoma W400 e un portatile ASUS perchè tali strumentazioni erano già presenti in azienda e offrono tutte le funzionalità necessarie a soddisfare i requisiti richiesti dal progetto.

## 5.1 Architettura hardware

Dall'analisi dei requisiti del progetto sono state identificate le seguenti componenti hardware:

1. un'unità di elaborazione;
2. un'unità di proiezione;
3. una camera RGB;
4. una camera di profondità;

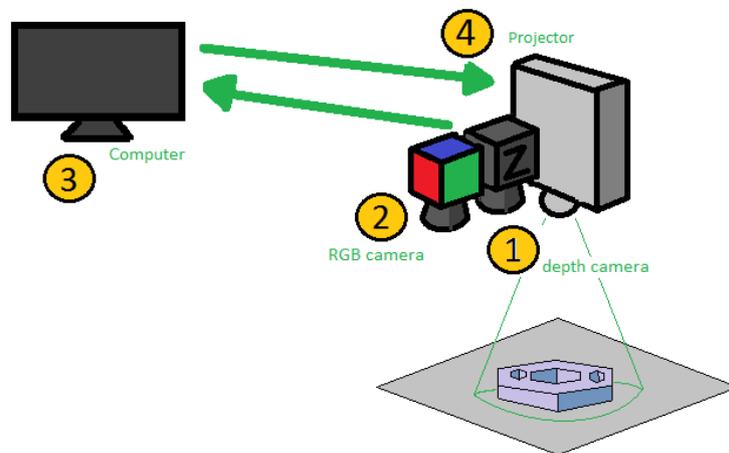


Figura 5.1. Schema dei componenti necessari.

### 5.1.1 Optoma W400



Figura 5.2. Proiettore Optoma W400.

Il proiettore usato per questo progetto di tesi è l'Optoma W400 anch'esso fornito da Illogc. L'Optma 400 si presta benissimo il lavoro richiesto da questo progetto di tesi. Esso è HD ready, dotato di due altoparlanti e le sue dimensioni sono abbastanza contenute (298 mm x 230 mm x 96 mm). Di seguito vengono esposte le principali caratteristiche del proiettore:

Tecnologia display	DLP
Risoluzione nativa	WXGA (1280x800)
Luminosità	4.000 lumen
Rapporto di contrasto	22.000
Aspetto nativo	16:10
Aspetto - compatibile	4:3,16:9
Dimensione schermo	0.68m ~8.04m (26.87" ~316.5") (diagonal)

Tabella 5.1. Caratteristiche Display/immagine del Optoma W400.

Rapporto di tiro	1.19:1 - 1.54:1
Distanza Proiettore (m)	1.2m - 12m
Zoom	1.3
Tipo Zoom	Manuale
Focale (mm)	16.91 ~21.6
Offset nativo	112%

Tabella 5.2. Caratteristiche Ottico del Optoma W400.

### 5.1.2 Kinect V2

Kinect versione 2 è un dispositivo di motion detection sviluppato e commercializzato da Microsoft. Esso è nato come accessorio per il gaming, sviluppato inizialmente per la Xbox 360, si è presto affermato come dispositivo di ricerca nei più svariati campi grazie alle funzionalità e alle caratteristiche offerte dal prodotto. La Kinect v2 è in grado di rilevare i movimenti del corpo umano e tradurli in eventi specifici all'interno dei giochi per Kinect o altri tipi di applicativi. Il dispositivo gestisce quattro differenti sistemi hardware rappresentati da una camera RGB, un sensore di profondità, un proiettore a raggi infrarossi e un set di microfoni per l'audio.



Figura 5.3. Sensori della Kinect.

Come si evince dalla figura sopra esposta, la camera di profondità e la camera RGB si trovavano in posizioni spaziali diverse. Questo implica un diverso punto di vista da parte delle due camere.

RGB stream	1024x1080 px
Depth range	0.50 m - 8 m
Depth sensor type	Time of Flight
Depth stream	512x424 px
Infrared stream	512x424px
Audio stream	4 mic 48 kHz
USB	3.0
# Body trackerd	6
# joints	25

Tabella 5.3. Caratteristiche principali della Kinect V2.

### 5.1.3 Il sensore di profondità e la camera RGB

La camera RGB ha una risoluzione di 1024x1080 pixel e lavora a 30 frame per secondo. Il suo field of view è di 84.1° x 53.8°. Il sensore di profondità ha una risoluzione di 512x424 pixel e lavora a 30 frame per secondo. Il field of view della depth camera è di 70 gradi in orizzontale e 60 gradi in verticale. Al suo interno la camera di profondità contiene un sensore monocromatico CMOS e un proiettore a luce infrarossi. Attraverso questo sistema la Kinect è in grado di creare immagini 3D di ciò che il sensore riprende. La misurazione della distanza viene calcolata grazie al "tempo di volo" che impiegano i raggi infrarossi ad arrivare sulla superficie, essere riflessi e tornare indietro sul sensore. Questa tecnica di misurazione soffre di alcune restrizioni: infatti non è possibile misurare con precisione superfici con particolari caratteristiche riflettenti come oggetti trasparenti, specchi o materiali estremamente lucidi, superfici nere. La depth camera ha un raggio di azione che va dai 50 cm fino agli 8 metri, anche se si ha una buona precisione entro i 4 metri. Il sensore di profondità cattura 217 088 valori di profondità, cioè pari alla risoluzione di 512x424 offerta dalla camera. I valori di profondità, ognuno dei quali memorizzato su 16 bit, sono organizzati in una matrice(512x424) e possono essere interpretati come un'immagine in scala di grigi in cui l'intensità del colore rappresenta la distanza in millimetri tra la camera e la superficie catturata nell'immagine.

### 5.1.4 Kinect for Windows SDK 2.0

Il Kinect for Windows Software Development Kit di Microsoft offre tutte le funzionalità per interfacciarsi con la Kinect v2. I seguenti requisiti di sistema sono necessari per l'utilizzo del SDK:

- sistema operativo Windows 8 o più recente;
- processore 64-bit (x64);
- 4 GB di memoria o più;
- Physical dual-core 3.1 GHz o un processore più veloce;
- USB 3.0 controller dedicato per la Kinect;
- DX11 capable graphics adapter;

### 5.1.5 L'unità di elaborazione

Per lo sviluppo del codice e il testing dell'applicazione è stato usato un PC portatile ASUS avente le seguenti caratteristiche:

- processore Intel® Core™ i7-6500U CPU @ 2,50 GHz 2.59 GHz;
- memoria RAM da 12 GB;
- sistema operativo Windows 10 Pro 64 Bit;
- scheda grafica GeForce GTX 950.

Le caratteristiche offerte da questo PC soddisfano perfettamente i requisiti minimi necessari per l'utilizzo della Kinect V2 e l'uso del relativo SDK esposti nel precedente paragrafo.

### 5.1.6 Composizione dell'area di lavoro

Per poter soddisfare i requisiti riguardanti la postazione di lavoro esposti nel paragrafo 2 del capitolo 3 riguardante i requisiti del progetto è stata costruita la seguente struttura.



Figura 5.4. La postazione di lavoro.

La disposizione della Kinect, del proiettore e la scelta delle misure sono frutto di vari test eseguiti usando il pavimento come base d'appoggio per i dispositivi e la parete come piano di proiezione. Lo scopo di tale simulazione è stato quello di trovare le misure e le posizioni più adatte per la Kinect e il proiettore in modo da soddisfare i requisiti di progetto e avere la resa migliore da parte dei dispositivi in base alle loro caratteristiche. Nello schema sottostante sono riassunte le misure risultanti dal test e che sono poi state usate per fissare i due dispositivi sul muro al di sopra del tavolo usato come piano di lavoro.

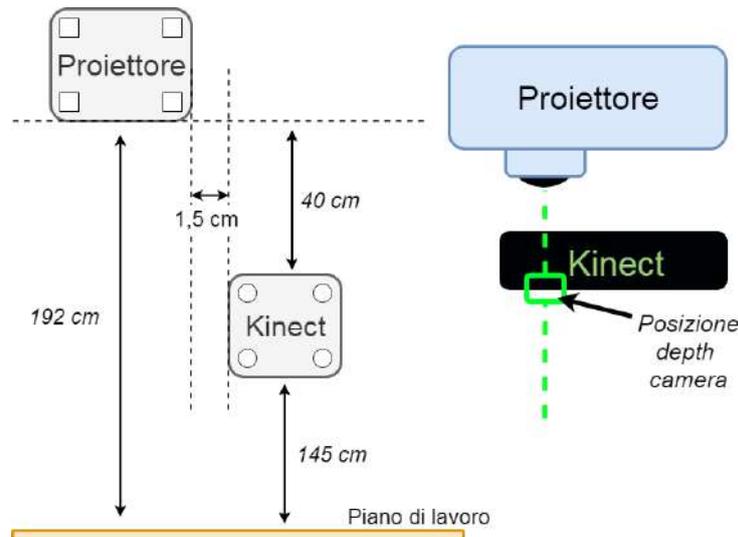


Figura 5.5. Disposizione della Kinect e del proiettore (a sinistra vista laterale, a destra vista dall'alto rispetto al muro).

Le misure espresse in figura 5.5 sono relative alle basi delle staffe dei due dispositivi. In questo modo è stato possibile fissare con precisione i dispositivi sul muro. Le staffe usate per sostenere la Kinect e il proiettore offrono un discreto grado di mobilità che ha permesso successive regolazioni anche dopo il fissaggio sul muro. In base a questa disposizione, è stato possibile soddisfare i requisiti del progetto relativi all'area di proiezione che deve essere coperta. Attraverso questa disposizione l'area coperta dal proiettore è di circa 150 cm di larghezza e 84 cm di altezza. La scelta della disposizione della Kinect ha invece permesso al sensore di profondità di incorniciare l'intera area di proiezione coperta dal proiettore. I cavi necessari all'alimentazione e al passaggio dei dati dei due dispositivi sono stati racchiusi in una canaletta di plastica che scorre lungo il muro in modo da avere la vista del piano di lavoro libera. Il tavolo sottostante ai due dispositivi, è stato coperto da una tovaglia bianca in modo da far risaltare colori e forme dettati dalla realtà aumentata.

## 5.2 Architettura software

In questo paragrafo si espone l'architettura del software, i ragionamenti che hanno portato a determinate scelte implementative e la selezione degli strumenti utilizzati.

### 5.2.1 Analisi preliminare

Il sistema deve essere in grado di eseguire differenti procedure. Nasce quindi la necessità di avere un'architettura capace di accettare in input una procedura generica ed eseguirla. È essenziale la presenza di una fase preliminare in cui avviene la creazione della procedura. Anche la gestione delle aree interattive richiede una particolare attenzione: le aree devono essere gestite in fase di costruzione della relativa procedura. In modo da poter soddisfare le esigenze di progetto, sono state individuate tre componenti essenziali:

1. un'applicazione per la definizione delle aree sensibili e dell'immagine di riferimento (necessaria in seguito per la rilevazione delle mani durante la procedura);
2. uno scene editor per la creazione delle scene virtuali e la gestione della logica riguardanti le procedure;
3. l'applicazione in realtà aumentata che supporta l'operatore nel compito di assemblaggio.

Gli strumenti di sviluppo utilizzati sono stati i seguenti:

- sistema operativo: Windows 10;
- ambiente di sviluppo: Visual Studio;
- linguaggio di programmazione: C++;
- librerie e framework esterni: OpenCV, OpenSceneGraph.

### 5.2.2 Prima fase

In questa fase viene catturata l'immagine di profondità che verrà usata come riferimento nella fase successiva e vengono definite quali aree della scena reale dovranno essere interattive durante la procedura. Questo viene fatto tramite il primo applicativo che, attraverso la camera RGB della Kinect, permette all'utente di selezionare tali aree direttamente su un'immagine. Le coordinate estrapolate in questo step vengono usate da particolari attori appartenenti alla scena virtuale. La scena virtuale è stata creata tramite lo scene editor proprietario di Illogic. In questa fase viene definita anche la procedura con i vari sotto task che dovrà eseguire l'operatore. Attraverso l'editor è gestita anche la logica necessaria per gestire la scena durante tutta la procedura.

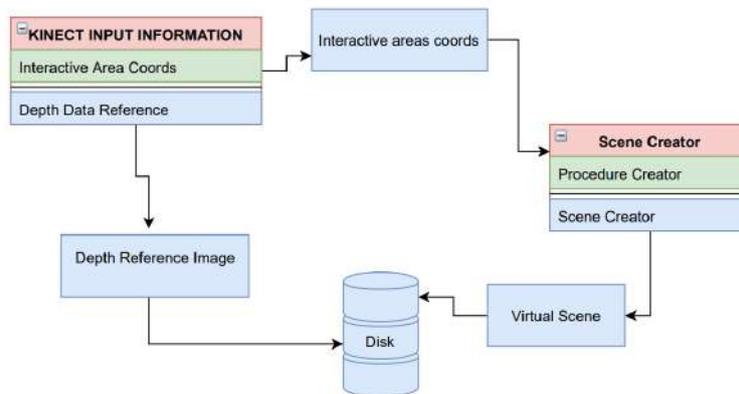


Figura 5.6. Schema logico della prima fase.

### 5.2.3 Seconda fase

L'applicazione in realtà aumentata riceve in input uno streaming continuo di informazioni di profondità da parte della Kinect. Il software carica da disco la scena virtuale compresa delle informazioni riguardo le aree interattive e l'immagine di profondità create nella prima fase. Tramite tutte queste informazioni, a seguito di un processamento delle immagini, il software è in grado di identificare in ogni momento la posizione delle mani dell'operatore e gestire eventuali interazioni col sistema. In base ai risultati derivanti dal processamento il sistema gestito dal game engine VRStar proietterà le relative informazioni atte a virtualizzare l'area di lavoro.

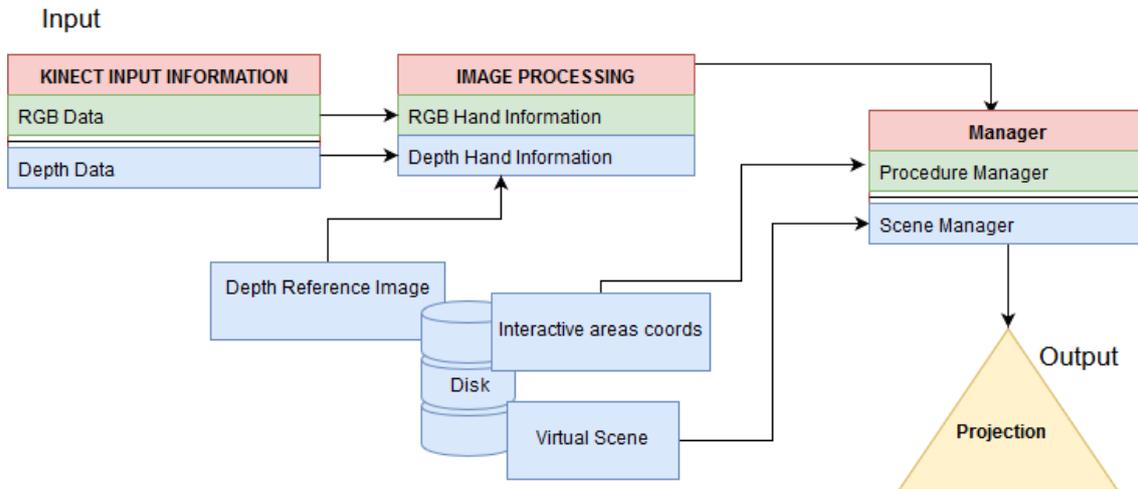


Figura 5.7. Schema logico dell'applicazione in realtà aumentata.

### 5.3 Le aree di interazione

Il primo passo, è stato quello di creare un'applicazione desktop in grado di definire le aree della scena reale che dovranno essere interattive durante lo svolgimento della procedura da parte dell'operatore. Inoltre il software definisce la matrice delle informazioni di profondità necessarie durante la fase di assembling. Per la definizione delle aree interattive è stato creato un tool che offre all'utente la possibilità di tracciare, attraverso il mouse, tali regioni direttamente su un'immagine RGB proveniente dalla Kinect. La Kinect V2 ha una camera RGB avente risoluzione 1920 x 1080 pixel. La camera di profondità ha una risoluzione di 512 x 424 pixel, e come esposto nel paragrafo 5.2 riguardante l'architettura della Kinect, le due camere hanno punti di vista diversi, essendo distanziate tra loro di qualche centimetro. Risulta quindi necessaria una mappatura tra il punto di vista e la risoluzione della camera RGB e il punto di vista e la risoluzione della camera di profondità. In tal modo è possibile avere una corrispondenza tra le informazioni provenienti dai due sensori. Per risolvere questo problema ci si è serviti delle funzioni offerte dal SDK per la Kinect v2 di Microsoft. In questo modo l'utente può scegliere le aree interattive data l'immagine RGB del piano di lavoro. L'applicazione grazie anche all'uso della libreria open-souce OpenCV, è in grado di fornire le coordinate per ogni area e le misure di profondità essenziali per il funzionamento del prototipo. Una volta calcolate le aree, è stato necessario creare un "Actor Component" di VRStar in grado di gestire le coordinate ricavate

nel passo precedente. Il componente può essere "agganciato" ad un attore di tipo bottone che può quindi gestire il comportamento.



Figura 5.8. determinazione delle coordinate di un'area di interazione.

## 5.4 Il rilevamento delle mani

Lo scopo di questo modulo è quello di identificare, frame per frame e in real-time, la posizione delle mani dell'operatore nello spazio 3D. Le mani vengono identificate attraverso le variazioni di distanza che avvengono tra il sensore di profondità e il piano di lavoro. Questo modulo è essenziale per gestire l'interazione uomo-macchina durante le fasi di assemblaggio.

Per rilevare le variazioni di profondità che avvengono nella scena inquadrata, e quindi sul piano di lavoro, ci si è serviti della camera di profondità. Per comprendere il procedimento seguente basti sapere che il sensore di profondità della Kinect restituisce un'immagine avente risoluzione  $512 \times 424$  pixel in cui ogni valore dei pixel indica la distanza (espressa in millimetri) che intercorre tra il sensore e i piani delle superfici inquadrato.

Il procedimento di rilevamento delle variazioni di profondità delle mani nell'area di lavoro segue i successivi passi:

1. durante la definizione delle aree interattive viene acquisita un'immagine di riferimento da parte del sensore di profondità e salvata in memoria. L'immagine viene catturata in assenza dell'utente dalla scena. In questo modo si ottiene una misura "pulita" della scena.
2. Il sensore di profondità cattura continuamente delle misure durante lo svolgimento della procedura da parte dell'utente. In-time, per ogni frame catturato dal sensore, l'applicazione fa un confronto tra le misure

presenti nell'immagine di riferimento e le misure rilevate nel frame corrente. Da questo confronto emergono le coordinate dei pixel che hanno valore di profondità diverso dall'immagine di riferimento. In base a tali informazioni viene creata un'immagine binaria della stessa risoluzione. Verranno colorati di bianco i pixel in cui il valore di profondità è variato di una certa soglia, i restanti verranno colorati di nero. In questo modo si è ricavata un'immagine che evidenzia le differenze di profondità nella scena.



Figura 5.9. Effetto dell'algoritmo al variare della distanza.

3. L'obiettivo è quello di rilevare quando la mano tocca una superficie in modo da simulare il tocco di un pulsante che secondo i requisiti del progetto deve essere una proiezione virtuale sul piano di lavoro. Per questo motivo è necessario rilevare le variazioni dei valori di profondità che avvengono entro un certo range al di sopra del piano in cui è proiettato il pulsante. È stato quindi aggiunto un ulteriore filtro all'algoritmo che permette di evidenziare di bianco solo i punti in cui la variazione di profondità varia di un massimo di 4 cm dall'immagine di riferimento. In questo modo se l'operatore passasse la mano o il braccio sopra il bottone senza toccare il piano, il bottone non verrà attivato.

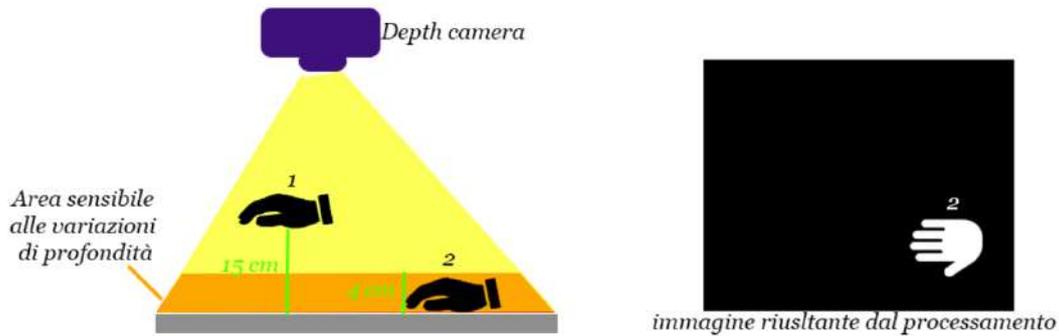


Figura 5.10. Le mani sono rilevate solo se aderenti al piano.

4. È stata necessaria una finitura dell'immagine per minimizzare il rumore introdotto dal sensore di profondità. Si è notato che il sistema introduce un notevole errore di misurazione ad ogni frame. Data l'imprecisione della misura, che tra un frame e il seguente può variare fino a qualche cm, si è deciso di alzare l'aria sensibile di rilevazione di un 1 cm dalla superficie di lavoro. Essendo il sistema ora capace di rilevare variazioni comprese tra 1 cm e 4 cm rispetto al piano di lavoro, il rumore è stato ridotto notevolmente ma non del tutto. Il passo successivo è stato quello di usare gli operatori morfologici: in particolare è stata effettuata un'operazione di erosione per eliminare il rumore seguita da un'operazione di dilatazione per ristabilire il più possibile le dimensioni delle mani alla misura originale.



Figura 5.11. Pulitura dell'immagine: filtro di erosione + filtro di dilatazione.

5. Ogni blob relativo alla corrispondente mano, ha bisogno di essere identificato spazialmente. attraverso una coordinata, è possibile capire se una mano entra all'interno di un area sensibile. Per ottenere la coordinata della mano si è deciso di calcolare il centroide del blob.

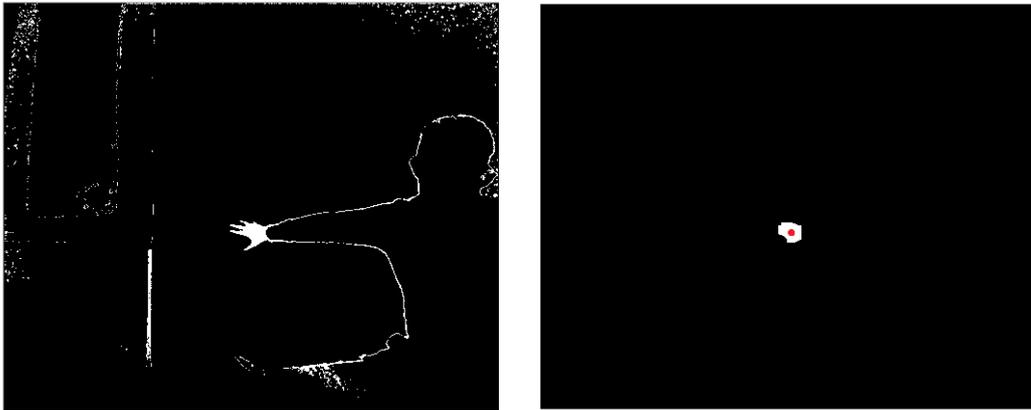


Figura 5.12. Il centroide della mano dopo la pulitura dell'immagine.

A questo punto dell'algoritmo si è deciso di testare il sistema in modo da capire quale sia il grado di precisione ottenibile. Si rimanda al paragrafo 6.1.1 (Test sull'interazione) per una descrizione più accurata riguardo tali test. Sono state definite diverse aree interattive di diversa misura. Si è potuto verificare che con questo sistema si ha una buona precisione nel selezionare un bottone se l'area relativa non sia più piccola di un quadrato avente il lato di 4 cm. Risulta invece difficoltoso lavorare con aree di dimensione minore. Per questo motivo si è deciso di seguire il seguente ragionamento per migliorare la precisione dell'algoritmo.

A partire dal centroide, viene costruito un quadrato il cui lato misura 6 cm. Il quadrato in questione rappresenta l'area coperta dal palmo di una mano media. In questo modo attraverso operazioni di intersezione e inclusione è possibile selezionare anche i bottoni più piccoli con maggiore facilità.

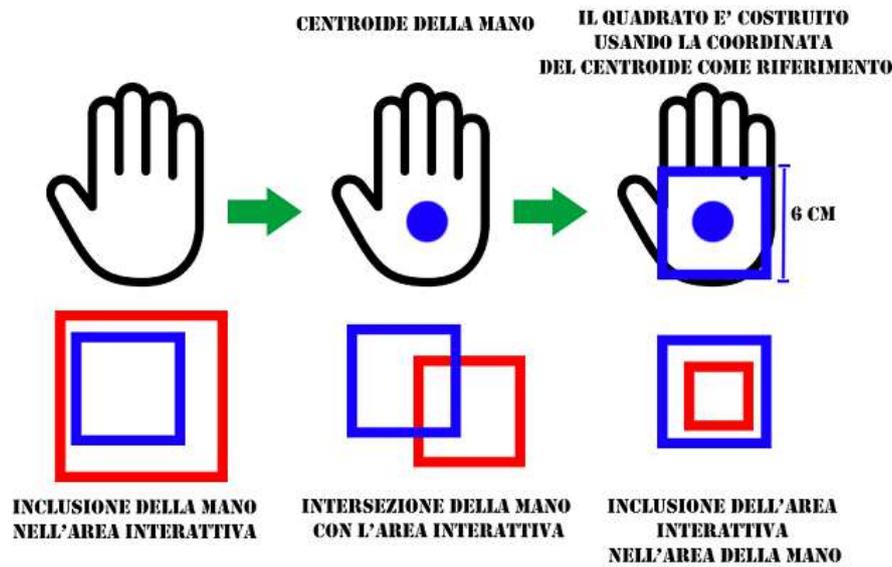


Figura 5.13. Possibili casi in cui l'area interattiva può essere selezionata.

## 5.5 Integrazione in VRStar

Gli algoritmi precedentemente esposti, sono stati in seguito strutturati in modo da poter essere integrati all'interno del motore proprietario di Illogic. È stato creato un manager e un componente denominati rispettivamente "HotAreaManager" e "HotAreaComponent".

- **HotAreaComponent**: esso è un componente esposto assieme agli altri oggetti di tipo component presenti nello scene editor di VRStar. Il component creato può essere associato ad un attore in scena di tipo bottone. Esso integra l'attore con quattro parametri che rappresentano le coordinate delle aree che noi vogliamo rendere interattive e un parametro "HotArea name" che identifica l'area. Sostanzialmente ogni bottone rappresenta una singola area interattiva e il suo comportamento è legato a ciò che succede all'interno dell'area sensibile.

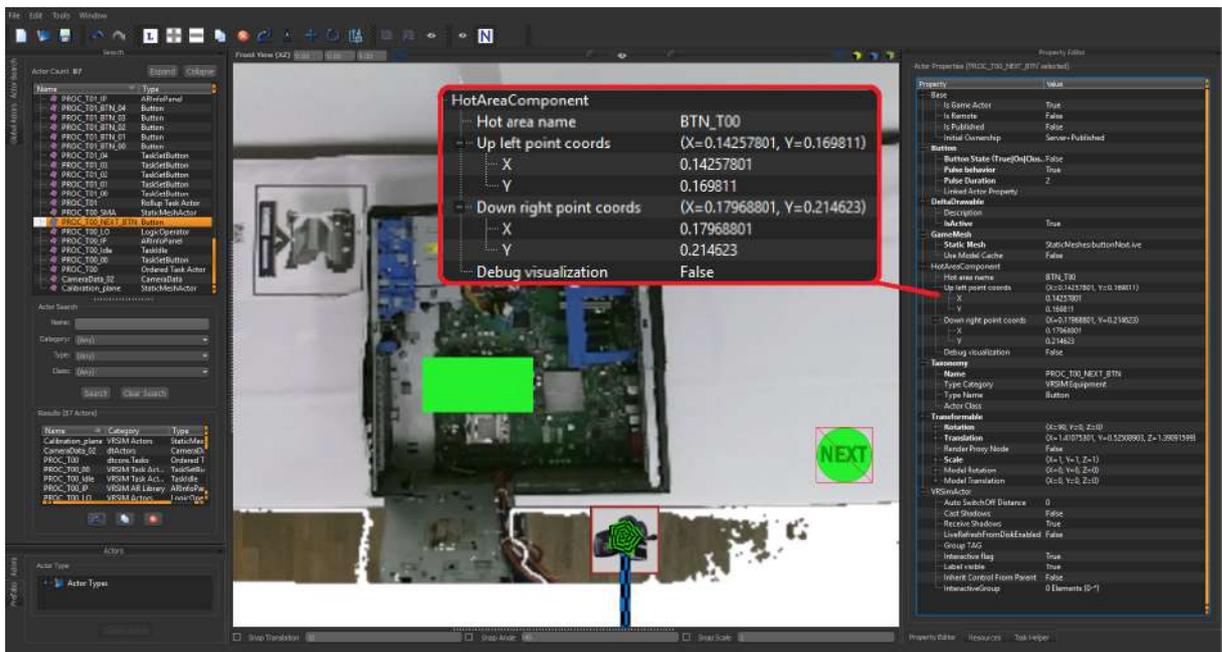


Figura 5.14. Il bottone "Next" è legato ad un'area interattiva.

- HotAreaManager:** I manager sono moduli dell'architettura di VRStar che gestiscono gli attori in scena in base al verificarsi di determinati eventi. HotAreaManager supervisiona la posizione delle mani dell'operatore durante la procedura, e ha visibilità su tutti gli attori a cui è associato un HotAreaComponent. Esso ha tutte le informazioni necessarie per poter capire quando le mani finiscono all'interno delle aree sensibili e scatenare un'evento sull'attore corrispondente all'area occupata dalla mano. Nella figura seguente è rappresentato il lavoro che esegue il manager ad ogni frame.

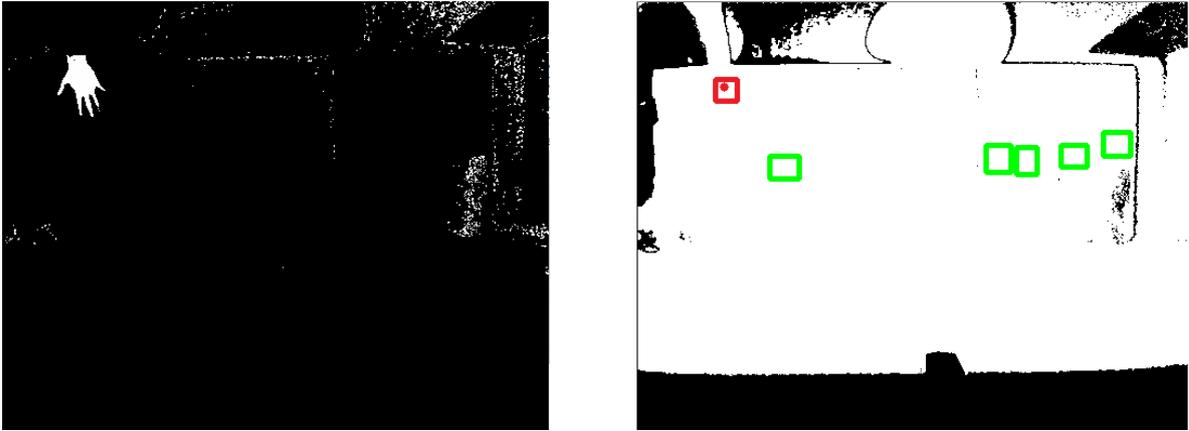


Figura 5.15. In verde le aree interattive, in rosso l'area selezionata.

## 5.6 La procedura

Obiettivo di questa tesi è la creazione di un prototipo in grado di assistere un operatore durante un procedimento di assemblaggio. Per questo motivo, una volta creato il sistema in grado di gestire l'interazione dell'utente, è stata creata una procedura in modo da testare il software. Per tale compito è stato scelto di creare una procedura in grado di guidare l'utente nell'assemblaggio di un PC.

### 5.6.1 Il design

L'applicazione deve essere in grado di fornire all'utente tutte le informazioni necessarie per eseguire i vari task. Le informazioni devono essere dirette e di immediata comprensione. Per questo motivo si è deciso di usare sia informazioni testuali per descrivere i task, sia informazioni indicative come un highlight sulle parti interessate. Nel caso di istruzioni più articolate si è deciso di proiettare anche delle animazioni create appositamente dai 3D artist di Illogic. Per un miglior risultato si è creato uno schema in scala dell'area di lavoro e si è lavorato su Photoshop CS6 per capire quali posizioni e quali proporzioni dovessero avere i vari componenti virtuali che dovranno aumentare la scena.

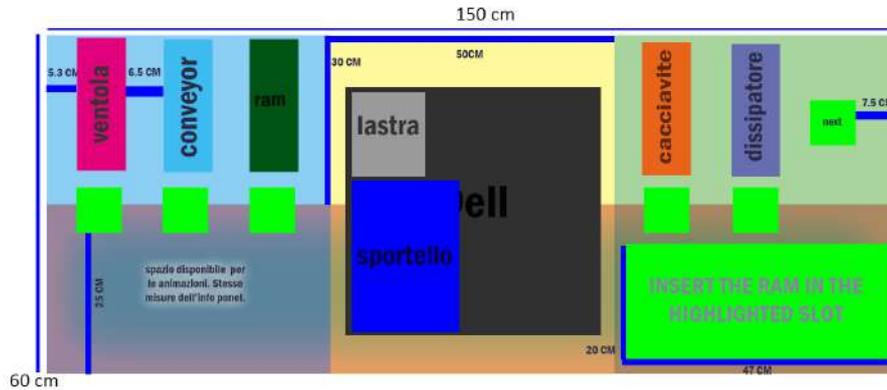


Figura 5.16. La ricostruzione in scala.

Questo lavoro è risultato molto utile in fase di costruzione della scena rendendo più semplici le fasi successive del lavoro. La scelta dei colori delle informazioni sintetiche non è casuale. È stato usato il verde per illuminare le componenti coinvolte nell'operazione e il pannello che descrive l'istruzione da eseguire in quanto l'utente medio associa inconsciamente tale colore a qualcosa che "non è dannoso" e che quindi si può toccare o fare. Nel caso di procedure che prevedano parti che non dovrebbero essere toccate, come per esempio componenti molto caldi, si può usare il rosso per evidenziarli, in modo da mettere in allerta l'utente sul "divieto" di toccare tale componente. L'arancione è stato scelto per le frecce che indicano il verso in cui deve essere eseguita un'operazione. L'arancione è un colore caldo che ha effetto stimolante (quindi adatto al suggerimento di esecuzione di un'azione) e una forte visibilità.



Figura 5.17. Design di un task.

Il tasto "NEXT" sta vicino al pannello delle istruzioni, che dovrebbe essere la prima informazione da considerare alla proiezione del task.



Figura 5.18. Animazione in basso a sinistra e highlight dei componenti coinvolti.

Le animazioni riguardanti il montaggio saranno proiettate tra i componenti e il case del PC. Ad ogni task all'utente sarà indicato sia che componente prendere sia dove esso debba essere inserito. Le scelte di disposizione delle informazioni sono frutto di vari test (effettuati in fase di sviluppo) su persone

che non conoscevano la procedura e che non hanno esperienza su applicazioni di questo tipo. Nonostante le animazioni coprano una buona parte del piano di lavoro, esse non venivano notate se proiettate a destra del PC. Questo, per quanto possa sembrare strano, è del tutto normale: una persona non esperta concentra la propria attenzione sul componente principale e i pezzi da assemblare, dando maggiore importanza all'area centrale del proprio campo visivo più di quanto farebbe in situazioni in cui è richiesta meno concentrazione.

### **5.6.2 I task**

La procedura è composta da dieci task organizzati nel seguente modo:

1. posizionare il PC nella posizione indicata dal software;
2. controllare che tutte le componenti che devono essere assemblate siano presenti su piano di lavoro;
3. rimuovere il pannello laterale dal case;
4. alzare la cover in modo da accedere alla scheda madre;
5. inserire la ram;
6. inserire la ventola;
7. montare il dissipatore;
8. inserire il convogliatore dell'aria;
9. abbassare la cover,
10. montare il pannello laterale del case;

### 5.6.3 Struttura dei task

I task sono organizzati secondo la seguente struttura ad albero:

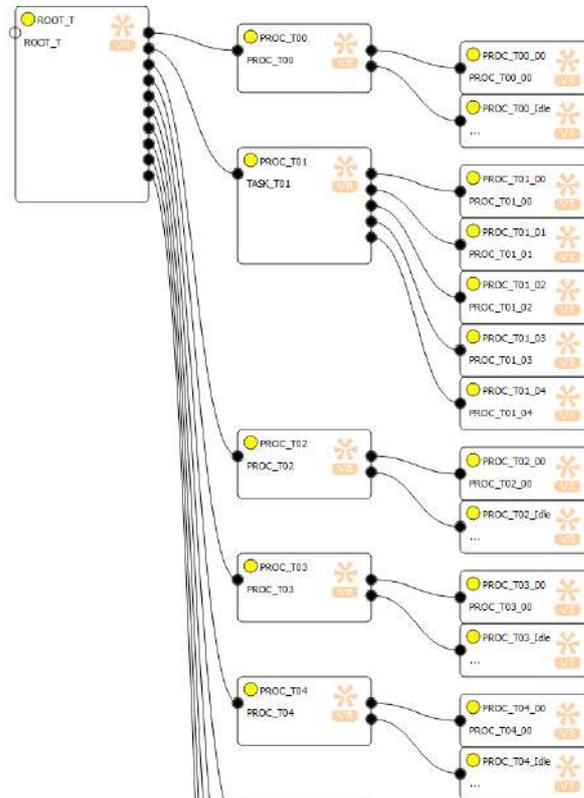


Figura 5.19. Struttura dei task utilizzata per il progetto.

In VRStar sono presenti attori di tipo task di varie tipologie. Per la creazione della procedura del progetto di tesi sono stati usati i seguenti attori di tipo task:

- **TaskSetButton:** sono legati alla proprietà di un attore di tipo button. Quando il bottone cambia stato il task a cui è legato viene considerato completato. Ai TaskSetButton sono legati gli attori che si riferiscono alle aree interattive;
- **RollUp:** sono attori di tipo task che risultano completati quando tutti i suoi sotto-task vengono completati. I sotto-task possono essere completati in ordine casuale.
- **Ordered:** Anche questo tipo di attore risulta completato quando tutti i suoi sotto-task vengono completati. In questo caso i sotto-task devono

essere completati secondo un ordine preciso. Nella procedura in esame viene usato per racchiudere l'intera struttura di task in modo che siano visualizzati dall'utente nell'ordine corretto;

- TaskIdle: sono task che risultano completati dopo il trascorrere di un determinato intervallo di tempo a partire dalla loro attivazione. Questo attore è stato usato per inserire una pausa di un secondo prima che venga caricata l'istruzione successiva.

#### 5.6.4 LogicOperator: la gestione dei task

Lo scopo del LogicOperator è quello di pilotare il comportamento degli attori in scena. Esso può prendere in input i valori delle proprietà degli attori, eseguire un'operazione di AND oppure di OR e in base al risultato modificare il valore delle proprietà di una lista di attori selezionati come output; Questo tipo di attore è stato usato per attivare e disattivare le varie componenti visuali in base al completamento dei task. Una volta che un task viene completato, vengono disattivate tutte le componenti visuali legate a tale task e attivate quelle relative al task successivo. Ogni task ha il proprio LogicOperator che effettua l'operazione appena descritta. Di seguito è riportata la tabella della verità utilizzata dal LogicOperator per pilotare gli attori durante i task.

LogicOperator N		
Math Operator Type: MULTIPLIER (AND)		
Input		Output
Task_N-1.Complete	!(Task_N.Complete)	value
0	!0 = 1	0
1	!0 = 1	1
1	!1 = 0	0

Tabella 5.4. Funzionamento del LogicOperator

Il LogicOperator N prende in input il valore della proprietà "Complete" del task N-1 e del task N. La proprietà "Complete" vale 0 se il task non è ancora stato completato e 1 quando il task è stato completato. Il valore della proprietà "Complete" del task N ricevuto in input viene negato. In questo modo si ha il valore 1 in output solo quando il task N-1 è completato e il task N non è completato. Il valore ottenuto in output viene impostato come valore nella proprietà "IsActive" di ogni attore legato al task N. In questo

modo gli attori relativi al task N saranno attivi e visibili solo quando il task N-1 è stato completato e il task N è da completare. In tutti gli altri casi gli attori saranno inattivi.

## 5.7 La creazione della scena virtuale

Per la creazione della scena virtuale è stato usato lo scene editor proprietario di Illogic. Esso offre allo sviluppatore una serie di attori che possono essere messi in scena. Essi possono essere principalmente di tre tipi:

- attori provvisti di una componente renderizzabile come una mesh o modelli 3D complessi;
- attori che hanno solo una funzionalità logica, capaci di gestire eventi;
- attori che hanno sia intelligenza che una componente visuale.

Oltre gli attori Task e LogicOperator descritti del paragrafo precedente, per la creazione della scena sono stati usati attori di tipo 2D poiché la scena deve essere renderizzata usando una camera ortogonale. Gli attori utilizzati sono i seguenti:

- **Button:** essi vengono pilotati dai componenti che rappresentano le aree interattive nella scena reale. Possono assumere lo stato OFF oppure lo stato ON. Quando una mano entra in un'area interattiva il relativo bottone si accende. Ad un bottone può essere associata una geometria.
- **StaticMeshActor:** sono attori non interattivi che rappresentano una geometria. Questo attore è stato usato per pilotare i quad usati per illuminare i componenti. Altri StaticMeshActor sono stati usati assieme al componente "ImageSequenceComponent" per gestire le animazioni.
- **InfoPanel:** questi attori non interattivi sono specializzati per la visualizzazioni di informazioni testuali o illustrative come immagini o icone. Essi vengono usati per fornire all'utente il testo descrittivo del task.

## 5.8 La projection mapping

Per la parte di projection mapping due sono stati i punti principali da affrontare:

- la necessità di avere una camera ortogonale;
- avere una corrispondenza tra scena reale e scena virtuale.

Il primo punto è stato risolto agendo direttamente sul codice sorgente del game engine VRStar integrando il sistema della possibilità di scegliere una camera prospettica o una camera ortogonale. Per il secondo punto invece ci si è serviti del seguente procedimento.

1. Acquisizione di un'immagine del piano di lavoro compreso delle componenti necessarie per l'assemblaggio. L'immagine è acquisita dalla camera RGB della Kinect;
2. creazione di una scena virtuale che prevede una camera ortogonale posta perpendicolarmente sopra un quad avente l'immagine catturata al punto 1 come texture;
3. a questo punto è stata effettuata la calibrazione in modo che la proiezione della scena virtuale coincida con la scena reale. Una volta proiettato il quad, aggiustando i parametri della camera virtuale (piani di clipping e la posizione della camera) si è raggiunto il matching tra reale e virtuale;



Figura 5.20. Gli oggetti sono stati usati come riferimento per la calibrazione.

4. si è ora in possesso di un sistema calibrato. Il quad con la cattura della scena reale funge da "blueprint" da prendere per riferimento per la disposizione degli elementi virtuali che dovranno essere proiettati. In questo modo è stato facile rispettare posizioni e proporzioni degli elementi, avendo un buon controllo sul sistema di proiezione.

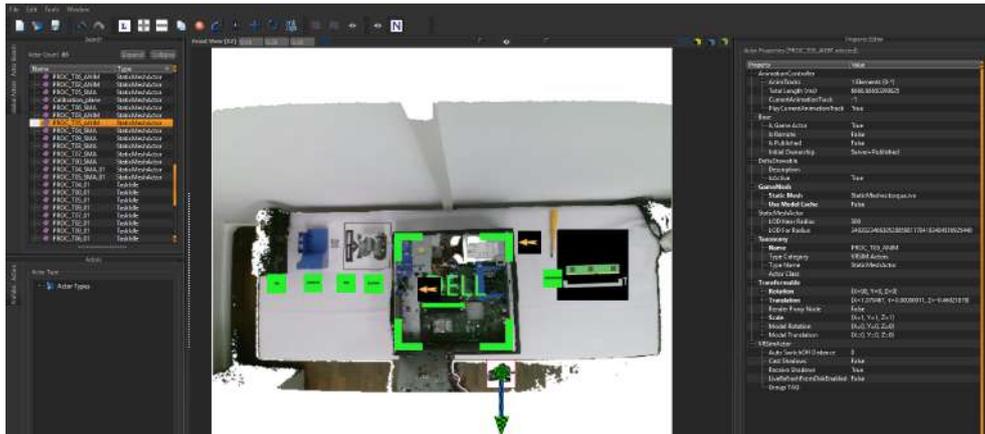


Figura 5.21. Creazione della scena sullo scene editor.

## 5.9 Considerazioni generali

Una delle criticità principali durante lo sviluppo di questo prototipo è stata quella di progettare un'architettura del sistema secondo lo stile imposto dal design del software proprietario di Illogic. Dopo una prima versione architeturale, è stato necessario comprendere e studiare la struttura di VRStar in modo da capire come predisporre e integrare il codice.

Nella fase di creazione della scena le maggiori difficoltà si sono riscontrate nella gestione dei task: le problematiche sorte sono state relative all'uso degli attori LogicOperator. Essi sono strumenti molto potenti con il quale è possibile gestire tutti gli elementi in scena con grande controllo. Di contro il loro utilizzo risulta complicato e vanno quindi compresi in maniera approfondita.

Per quanto riguarda la parte di image processing e la gestione di librerie grafiche di basso livello si è rivelata molto utile la preparazione acquisita durante questo corso di magistrale.

Per la parte di projection mapping la disposizione delle informazioni e il loro design si sono rivelati fattori fondamentali: un design ben pianificato influisce sensibilmente sulle performance dell'operatore.

# Capitolo 6

## Risultati

Una volta ultimato il software, la procedura è stata oggetto di test. Sono stati presi in considerazione sia parametri oggettivi che soggettivi. Oltre i risultati, in questo capitolo si descrivono i limiti del prototipo e possibili soluzioni.

### 6.1 Test in fase di progettazione

Durante la fase di sviluppo, sono stati eseguiti vari test per capire quale fosse l'accuratezza del prototipo riguardo l'interattività del sistema. Per tali test, si sono offerte tre persone: una di sesso femminile con nessuna conoscenza pregressa riguardo applicativi di realtà aumentata o simili, e due ingegneri informatici esperti del settore. I test descritti riguardano le aree d'interazione e il design della scena virtuale.

### 6.1.1 Test sull'interazione

L'obiettivo del primo test riguarda la precisione del sistema: misurare la dimensione minima possibile che possono avere le aree d'interazione. La kinect v2 è posizionata all'altezza di 145 cm dal piano di lavoro. Sono state create aree quadrate di diversa grandezza e una scena di test in cui sono presenti solo i bottoni. Sono state definite dieci aree interattive che ascendono in grandezza, in cui la misura del lato varia 10 cm a 1 cm.



Figura 6.1. Aree di test.

Dal test, usando il centroide come descritto nella prima parte del paragrafo 5.4 (Il rilevamento delle mani), si è potuto appurare che risulta semplice e con una precisione del 100% selezionare aree la cui lunghezza del lato non sia minore di 4 cm. Per le aree il cui lato misura 3 cm, risulta ancora possibile selezionarle ma con un po' di difficoltà, mentre per le due aree più

piccole risulta molto difficile riuscire ad attivarle. Il test ha messo in luce un importante problema e questo ha portato all'implementazione della soluzione descritta al punto 5 del paragrafo 5.4("Il rilevamento delle mani"): in breve, il centroide è stato "convertito" in un area di forma quadrata il cui lato misura 6 cm. In tal modo, attraverso operazioni di inclusione ed intersezione è possibile selezionarle anche le aree più piccole.

### **6.1.2 Il problema dei falsi positivi**

Il sistema è capace di rilevare la selezione di un bottone. La selezione può essere eseguita non solo attraverso la mano ma tramite qualsiasi oggetto che si trovi nel range di profondità compreso tra 1 cm e 4 cm al di sopra delle superfici utilizzate come riferimento per il bottone. In questo stato il bottone si comporta come un classico interruttore "tangibile": per esempio, i pulsanti della tastiera sono stati pensati per essere premuti dalle dita umane, ma nulla vieta che possano essere attivati attraverso qualsiasi oggetto solido. Tale comportamento da parte del sistema potrebbe forse essere utile in alcuni contesti, ma potrebbe essere un limite in altri: basti pensare alla situazione in cui un operatore debba spostare oggetti relativamente ingombranti o difficili da tenere sotto controllo come un mazzo di cavi. In tal contesto le aree interattive possono essere attivate accidentalmente con facilità andando così a disorientare e rallentare il lavoro dell'operatore. La soluzione più immediata che può venire in mente è quella di usare un algoritmo per il riconoscimento delle mani attraverso una camera RGB e andare a verificare le informazioni di profondità alle coordinate individuate dall'algoritmo di hand recognition.

### **6.1.3 Intelligibilità delle informazioni**

Durante la fase di sviluppo, a lavoro quasi ultimato, si è deciso di far provare il prototipo a persone senza nessuna esperienza relativa all'assemblaggio di un PC. Dai test è risultato che la sola indicazione del componente e del relativo slot in cui deve essere montato in aggiunta all'istruzione descritta testualmente non sono abbastanza esplicativi. Il passo seguente è stato quello di aggiungere delle animazioni che spiegassero in che modo deve avvenire il montaggio. Anche in questo caso non sempre le animazioni vengono comprese perfettamente dall'utente inesperto. Tali risultati suggeriscono l'importanza di fornire all'utente animazioni altamente fedeli e dettagliate sul compito da eseguire. In alcuni casi l'istruzione può essere meglio compresa attraverso un'animazione che simuli, grazie a modelli digitali, l'operazione in modo

che si possano isolare e mostrare solo determinati meccanismi. In altri casi sarebbe meglio fornire le istruzioni attraverso uno streaming video, in cui l'utente viene guidato da utenti esperti.

## 6.2 Test finali

Per la valutazione complessiva del prototipo si è deciso di raccogliere sia parametri oggettivi che parametri soggettivi. Per tali test sono stati reclutati 14 utenti che gentilmente hanno dedicato un po' del loro tempo per eseguire la procedura relativa all'assemblaggio di un PC. Il prototipo è stato testato su un gruppo di 14 persone di età compresa tra i 23 e i 44 anni. Dei 14 utenti, 5 sono di sesso femminile e 9 di sesso maschile. 5 tester provengono da un background ingegneristico di indirizzo informatico con conoscenze riguardanti la realtà virtuale e la realtà aumentata e con esperienza su applicazioni affini. Il resto dei tester ha invece una formazione umanistica o economica. La metà degli utenti non aveva mai utilizzato prima un'applicazione in realtà aumentata.

I passi della procedura, descritti dal paragrafo 6.2 del capitolo 5.

1. posizionare il PC nella posizione indicata dal software;
2. controllare che tutte le componenti che devono essere assemblate siano presenti su piano di lavoro;
3. rimuovere il pannello laterale dal case;
4. alzare la cover in modo da accedere alla scheda madre;
5. inserire la ram;
6. inserire la ventola;
7. montare il dissipatore;
8. inserire il convogliatore dell'aria;
9. abbassare la cover,
10. montare il pannello laterale del case;

I fattori di giudizio presi in considerazione per i test oggettivi sono il numero di errori durante lo svolgimento della procedura e i tempi di completamento. Per quanto riguarda i test soggettivi viene proposto agli utenti un questionario di 10 domande.

### 6.2.1 Misurazione dei parametri oggettivi

Per i test oggettivi sono stati presi in esame gli errori commessi dai tester e i tempi di esecuzione necessari per il completamento della procedura. Da tali test sono emersi i seguenti risultati. I tempi di completamento per la procedura variano tra i 2 minuti e 30 secondi circa per gli utenti più veloci e i 4 minuti e 30 secondi circa per gli utenti che non avevano mai assemblato un PC prima. Gli utenti più veloci, che sono anche quelli in possesso di esperienza pregressa in compiti simili, non hanno avuto particolari problemi nel eseguire la procedura, commettendo uno o due errori massimo. Gli errori commessi dagli utenti più esperti si sono riscontrati nel task numero 4 (alzare la cover) e nel task numero 8 (inserire il convogliatore dell'aria). Causa degli errori sono stati i dettagli: gli utenti alzavano il pannello interno fino a formare un angolo di 90 gradi. Ma solo portando il pannello oltre i 90 gradi è possibile successivamente inserire la ventola. Inserire il convogliatore dell'aria è di per sé un compito semplice: il pezzo deve scivolare nella scanalatura della ventola fino ad appoggiarsi sulla scheda madre senza impedimenti o manovre complicate. Individuare però tale scanalatura non è stato così immediato da parte dei tester più esperti.

Per quanto riguarda gli utenti senza esperienza, il numero di errori è aumentato notevolmente. In alcuni casi è stato necessario un aiuto esterno per indirizzare i tester nella giusta direzione. Un particolare errore, riscontrato nel task 5 (inserire la ram) ha rilevato un dato importante: l'animazione mostra in che modo inserire la ram, come fissarla e in quale verso essa debba essere posta. Gli utenti, non avendo nessuna esperienza a riguardo, non hanno notato assolutamente l'indicazione del verso.

### 6.2.2 Test soggettivi

A seguito del test pratico, agli utenti è stato proposto il questionario standard (SUS) System Usability Scale, creato da John Brooke nel 1986. Il questionario SUS è uno dei modi più efficaci per raccogliere dati statisticamente validi e fornire un punteggio chiaro e ragionevolmente preciso. Il punteggio SUS ha lo scopo di fornire una stima sulle prestazioni di usabilità dell'applicazione: efficacia, efficienza e facilità d'uso. Il punteggio è calcolato in base alle risposte fornite a ciascuna asserzione. un punteggio maggiore uguale a 68 indica che il sistema è usabile. Si rimanda al paragrafo 6.2.3 per una descrizione più dettagliata sull'interpretazione dei risultati.

Il SUS prevede 10 affermazioni: l'interessato valuta ogni asserzione in base all'esperienza personale relativa all'uso del sistema preso in esame. L'utente può esprimere un giudizio in una scala che va da 1 a 5 in cui 1 equivale a "fortemente in disaccordo" con la relativa affermazione, e 5 equivale a "fortemente d'accordo".

Le asserzioni proposte nel questionario sono le seguenti:

1. Penso che mi piacerebbe usare questo sistema frequentemente.
2. Ho trovato il sistema inutilmente complesso.
3. Penso che il sistema fosse facile da usare.
4. Penso che avrei bisogno del supporto di una tecnico per poter usare questo sistema.
5. Ho trovato che le varie funzioni di questo sistema erano ben integrate.
6. Penso che ci sia troppa incoerenza in questo sistema.
7. Immagino che la maggior parte delle persone imparerebbe a usare questo sistema molto velocemente.
8. Ho trovato il sistema molto complicato da usare.
9. Mi sentivo molto sicuro nell'utilizzare il sistema.
10. Ho avuto bisogno di imparare molte cose prima di poter utilizzare questo sistema.

Di seguito sono riportati i risultati per ogni singola risposta.

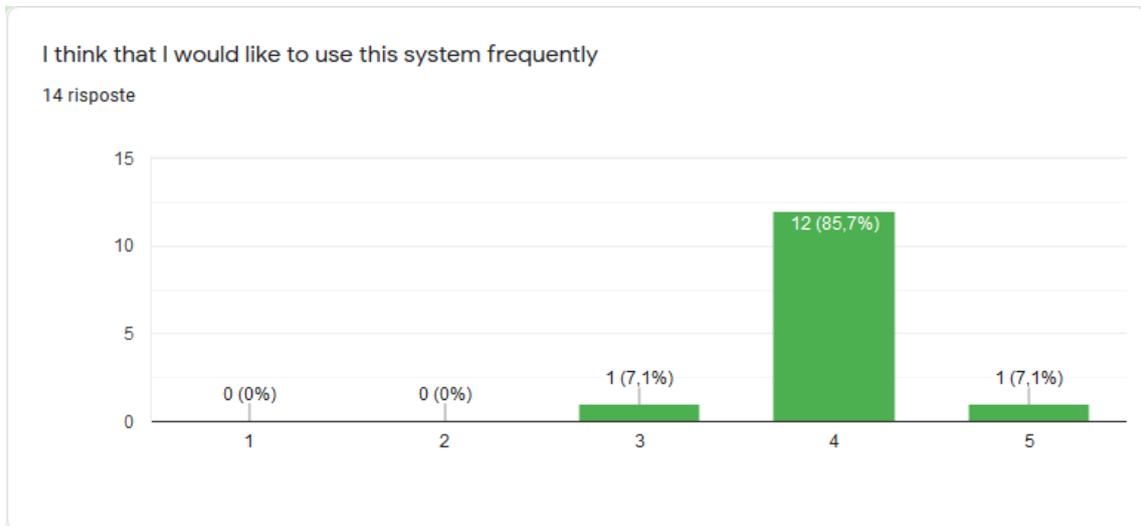


Figura 6.2. Asserzione 1: "mi piacerebbe utilizzare spesso questo sistema".

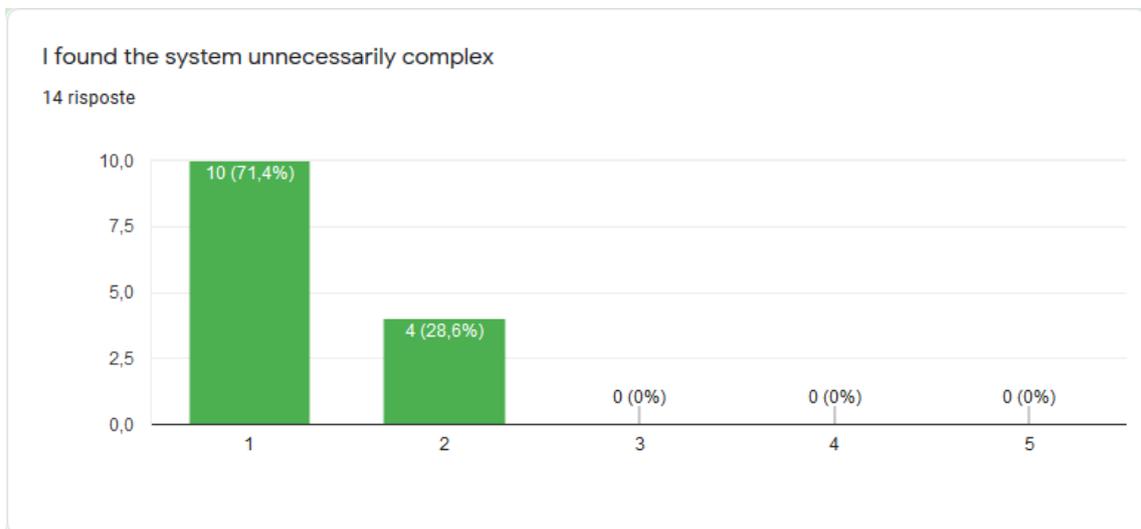


Figura 6.3. Asserzione 2: "ho trovato il sistema inutilmente complesso".

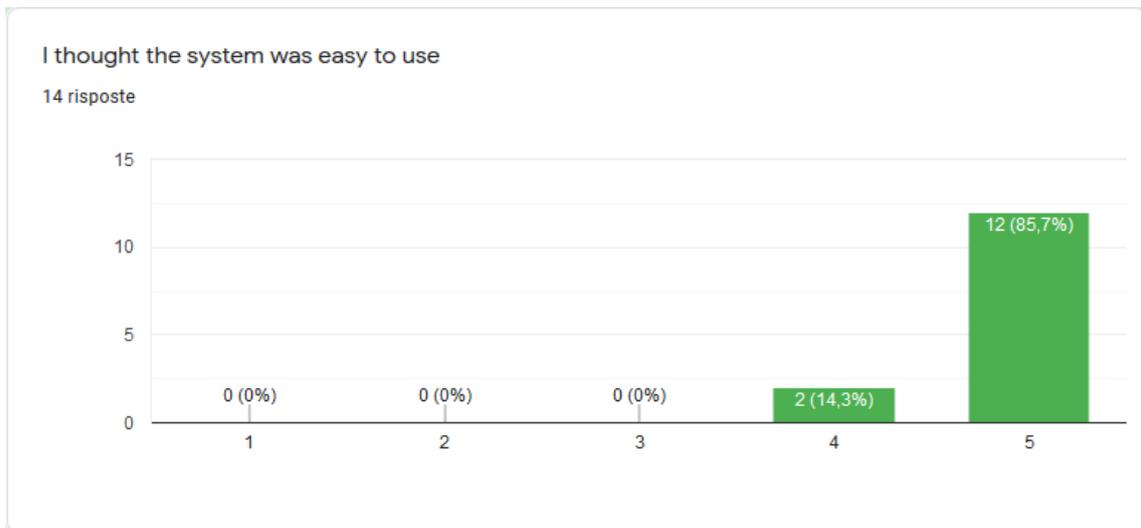


Figura 6.4. Asserzione 3: "penso che il sistema fosse facile da usare".

Dalle prime tre asserzioni si può evincere che il sistema risulta semplice da utilizzare e non crea particolare disagio per gli utenti.

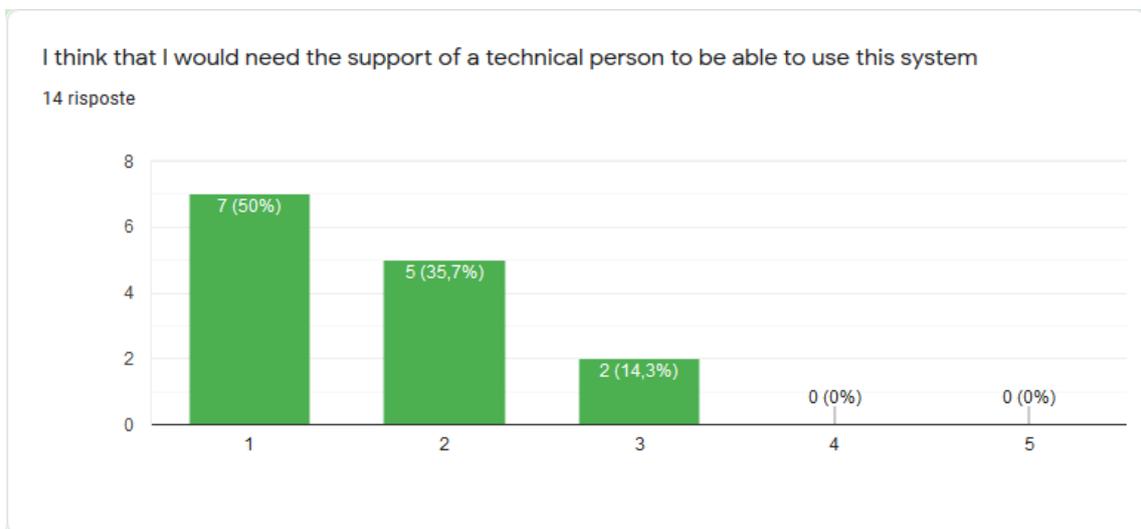


Figura 6.5. Asserzione 4: "penso che avrei bisogno del supporto di una tecnico per poter usare questo sistema".

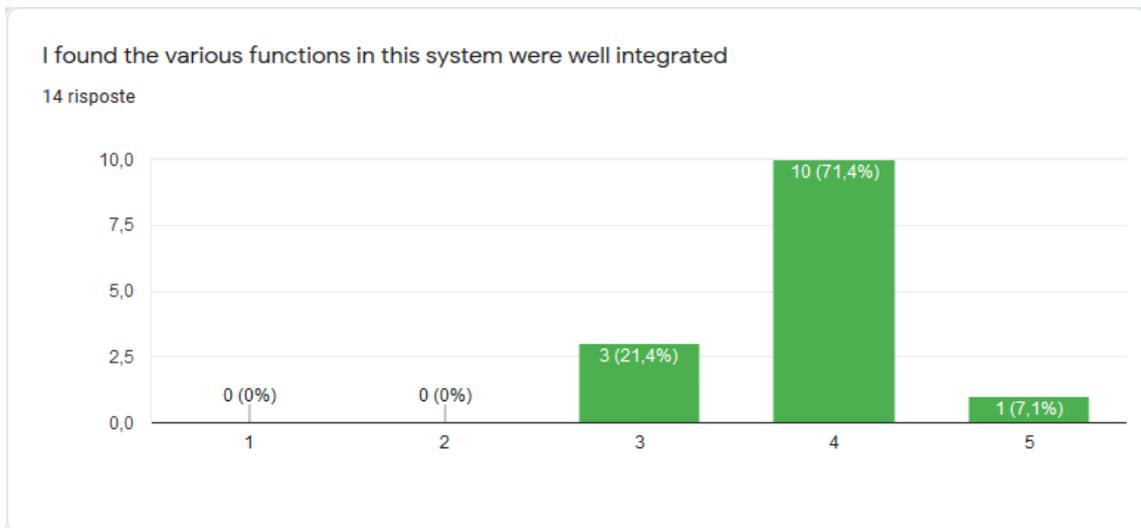


Figura 6.6. Asserzione 5: "ho trovato che le varie funzioni di questo sistema erano ben integrate".

Le risposte alle asserzioni 4 e 5 descrivono come il sistema sia abbastanza usufruibile senza il supporto di una persona esperta. Anche le funzioni del sistema risultano, nella media, efficienti.

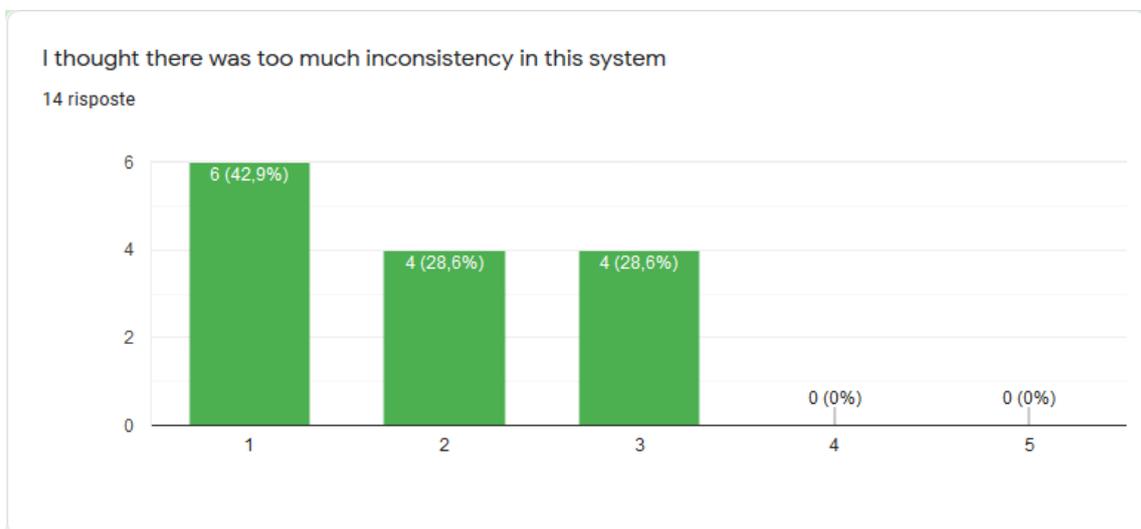


Figura 6.7. Asserzione 6: "penso che ci sia troppa incoerenza in questo sistema".

L'asserzione numero 6 rivela che il sistema sia consistente per la maggior parte dei tester.

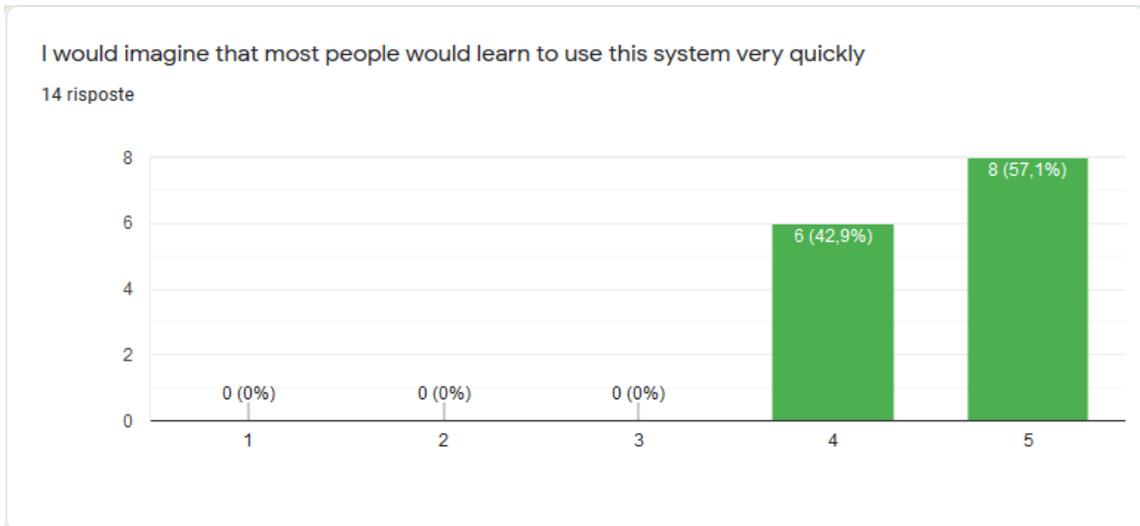


Figura 6.8. Asserzione 7: "immagino che la maggior parte delle persone imparerebbe a usare questo sistema molto velocemente".

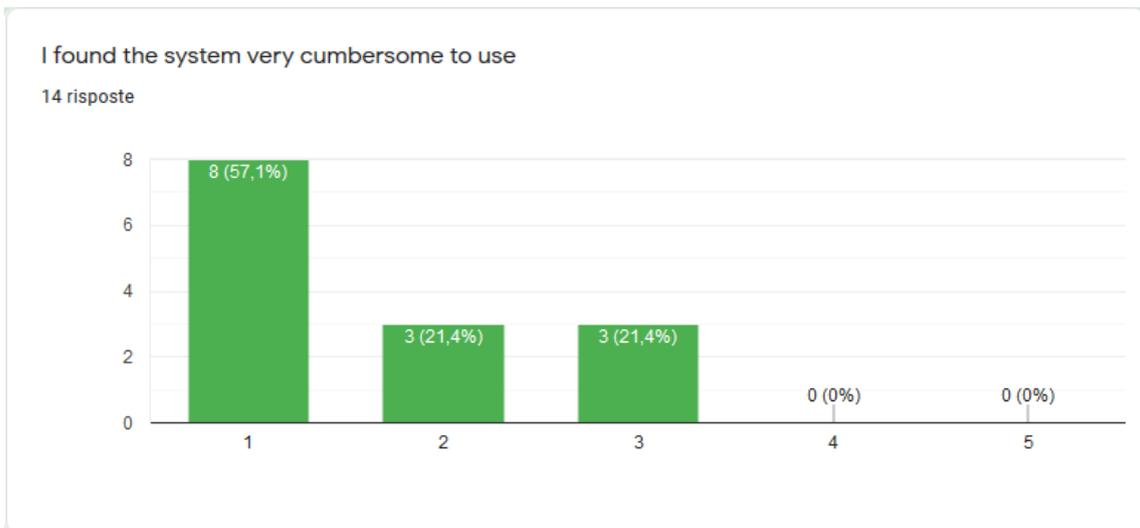


Figura 6.9. Asserzione 8: "Ho trovato il sistema molto complicato da usare".

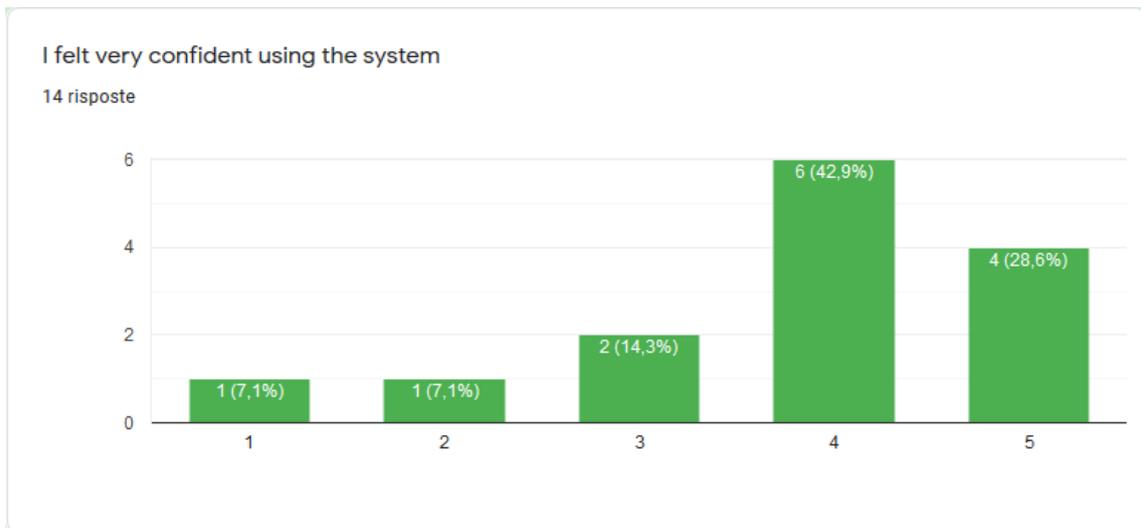


Figura 6.10. Asserzione 9: "mi sentivo molto sicuro nell'utilizzare il sistema".

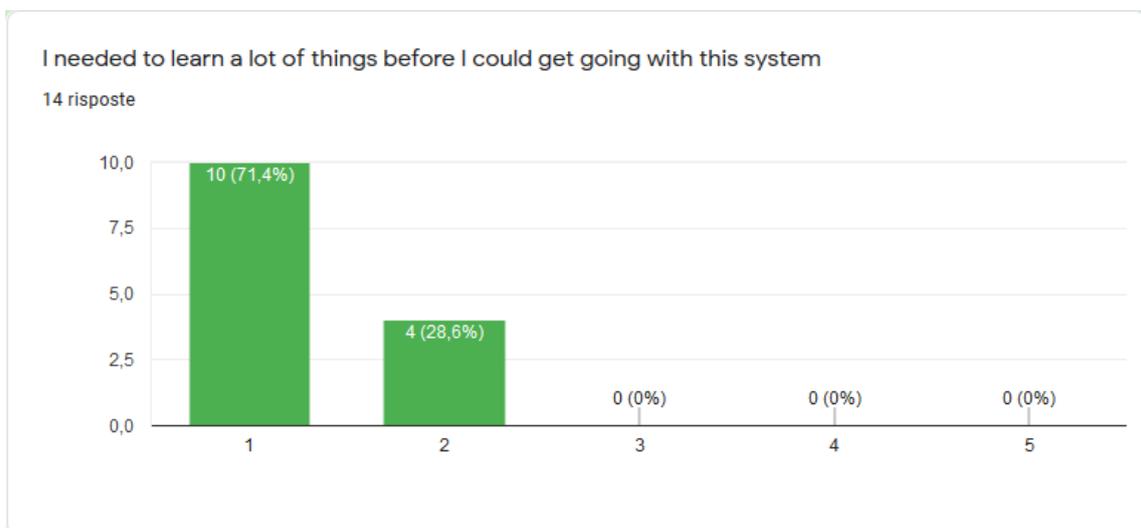


Figura 6.11. Asserzione 10: "ho avuto bisogno di imparare molte cose prima di poter iniziare con questo sistema".

Le risposte alle asserzioni 7, 8, 9 e 10 rivelano l'elevato grado di semplicità del sistema e come non sia necessaria nessuna o una minima conoscenza pregressa per poter utilizzare l'applicazione.

### 6.2.3 Valutazione SUS

Per ognuno dei 14 questionari è stato calcolato il rispettivo punteggio. Essi vanno da un minimo di 60 ad un massimo di 92.5. La media calcolata sul totale dei punteggi risulta essere 81. È possibile quindi affermare che il prototipo possiede i requisiti di usabilità richiesti dal test di valutazione SUS.

Punteggio SUS	Grado	Valutazione
>80.3	A	Eccellente
68 - 80.3	B	Buono
68	C	Sufficiente
51 - 68	D	Scarso
<51	F	Inutilizzabile

Tabella 6.1. SUS: interpretazione dei punteggi.

## Capitolo 7

# Conclusioni e sviluppi futuri

In tale documento è stato descritto lo sviluppo e le caratteristiche di un'applicazione di realtà aumentata che ha come scopo il supporto e l'addestramento degli operatori in compiti di assemblaggio e disassemblaggio in ambito industriale. L'intento dell'applicazione è quello di rendere più efficiente il processo industriale grazie ad un'ottimizzazione delle procedure di assemblaggio, migliorandone la velocità, diminuendone l'errore e minimizzando il supporto da parte di personale più specializzato.

I risultati ottenuti dai test oggettivi sono soddisfacenti: nonostante gli utenti abbiano commesso qualche errore nello svolgere la procedura, questa è stata comunque portata a termine da tutti i tester. Gli utenti senza esperienza pregressa hanno avuto bisogno dell'aiuto di terzi nel task riguardante l'inserimento della ram (descrizione dell'errore nel paragrafo 2.1, capitolo 6). Questo fatto mette in evidenza, ancora una volta, l'importanza di fornire informazioni il più possibile dettagliate.

Tale progetto ha dimostrato come la projection mapping abbia grandi potenzialità per essere sfruttata non solo nel campo dell'intrattenimento e dell'arte, ma anche in ambito industriale. Il suo punto di forza maggiore è sicuramente il fatto che gli utenti non debbano indossare alcun tipo di caschetto, occhiali, tenere in mano apparecchi come smartphone o tablet. Questo fatto rende la fruizione delle informazioni totalmente non invasiva per gli operatori, che in certi contesti, risulta essere un requisito fondamentale.

Il sistema descritto è ancora in fase prototipale. Molti sono i limiti e molte possono essere le migliorie che possono essere apportate a tale applicazione.

Il sistema potrebbe essere integrato con un modulo di image processing atto al riconoscimento dei vari componenti da montare, in modo da poterli individuare qualunque sia la loro posizione ed orientamento sul piano di lavoro. Non solo, il sistema dovrebbe essere in grado di capire, sempre grazie alla computer vision, se il componente è stato sistemato correttamente. Una terza miglioria potrebbe essere quella di correlare le informazioni della camera di profondità con le informazioni provenienti dalla camera RGB in modo da identificare le mani dell'operatore ovunque si trovino. In tal modo, si potrebbe capire se l'operatore sta prendendo il componente sbagliato. Infine, L'interpretazione dei gesti delle mani potrebbe essere un'altra funzionalità interessante: si potrebbe creare un set di gesti riconoscibile dal sistema utile in casi in cui non sia possibile l'interazione tramite la proiezione di bottoni virtuali per mancanza di una superficie adatta.

Le potenzialità della realtà aumentata sono considerevoli, se ne prevede una rapida ascesa in un futuro molto prossimo. Tale tecnologia grazie anche ai numerosi investimenti in merito, progredisce rapidamente. Si prevede che nei prossimi anni entrerà stabilmente a far parte delle vite dei non addetti ai lavori e delle persone comuni. Tale progresso occuperà una posizione di primaria importanza nella rivoluzione tecnologica 4.0

# Ringraziamenti

Dedico questo lavoro a tutte le persone che mi hanno accompagnato durante tutto il percorso universitario: familiari, amici, professori, colleghi e tutte quelle persone che sono state solo di "passaggio" durante questi anni. In particolare, ringrazio il Prof. Andrea Sanna per il costante appoggio, la grande disponibilità, la professionalità e gli ottimi consigli. Un grazie a Giuseppe Donvito e tutto il personale di Illogic per avermi sempre fatto sentire parte di loro fin dal primo giorno. Un grazie speciale anche ai colleghi Andrea C., Andrea P., Andrea S., Dario, Fabio, Giovanni e Lorenzo per la loro amicizia, il loro supporto e la loro pazienza! Ringrazio tutti i compagni di studio in particolare Anna, Antonio e Damiano. Grazie anche agli amici della mia cara Sardegna, soprattutto ai miei compaesani e i miei ormai ex-colleghi conosciuti durante gli studi della triennale che sono e saranno sempre un punto di riferimento per me. Ringrazio tutti gli amici e le persone conosciute a Torino in questi anni. Un grazie particolare va a quelli che io chiamo "gli amici del parco" di cui fanno parte tutti quei cari amici conosciuti in Piazza D'armi con cui ci siamo allenati duramente d'estate e d'inverno e che mi hanno accompagnato per tutta la mia permanenza a Torino. Un grazie anche al Politecnico e alla città di Torino per tutto ciò che mi hanno dato. Il grazie più importante va alla mia famiglia. Ai miei genitori che mi hanno sempre supportato e, che grazie ai loro sacrifici, ho avuto la fortuna di scegliere il mio futuro, e vivere questa esperienza. Spero di avervi reso orgogliosi con i miei risultati e di continuare a farlo anche in ambito lavorativo. I miei cari fratelli Fabrizio, Paolo e Nicola che ci sono sempre stati per qualsiasi cosa. La mia ragazza Jessica che mi è sempre stata vicino.

# Bibliografia

- [1] T.P. Caudell; D.W. Mizell. *Augmented reality: An application of heads-up display technology to manual manufacturing processes*, 1992. URL: <https://ieeexplore.ieee.org/document/183317>.
- [2] Steven Feiner; Blair Macintyre; Dorée Duncan Seligmann. *KARMA: knowledge based augmented reality for maintenance assistance*, 1993. URL: <https://dl.acm.org/doi/10.1145/159544.159587>.
- [3] J. Rekimoto, Y. Ayatsuka. *CyberCode: designing augmented reality environments with visual tags*, 1996. URL: <https://dl.acm.org/doi/10.1145/354666.354667>.
- [4] A.C. Boud; D.J. Haniff; C. Baber; S.J. Steiner. *Virtual reality and augmented reality as a training tool for assembly tasks*, 1999. URL: <https://ieeexplore.ieee.org/document/781532>.
- [5] J. Molineros; R. Sharma. *Computer vision for guiding manual assembly*, 2001. URL: <https://ieeexplore.ieee.org/document/929052>.
- [6] A. Tang; C. Owen; F. Biocca; Weimin Mou. *Experimental evaluation of augmented reality in object assembly task*, 2002. URL: <https://ieeexplore.ieee.org/document/1115105>.
- [7] J. Fruend; M. Grafe; C. Matyszczok; A. Vienenkoetter. *AR-based training and support of assembly workers in automobile industry*, 2002. URL: <https://ieeexplore.ieee.org/document/1106989>.
- [8] S. Wiedenmaier; O. Oehme; L. Schmidt; H. Luczak. *Augmented reality (AR) for assembly processes - an experimental evaluation*, 2002. URL: <https://ieeexplore.ieee.org/document/1106989>.

<https://ieeexplore.ieee.org/document/970534?arnumber=970534>.

- [9] S. Wiedenmaier; O. Oehme; L. Schmidt; H. LuczakN. Pathomaree; S. Charoenseang. *Augmented reality for skill transfer in assembly task*, 2005. URL: <https://ieeexplore.ieee.org/document/1513829>.
- [10] Hugo Álvarez; Iker Aguinaga; Diego Borro. *Providing guidance for maintenance operations using automatic markerless Augmented Reality system*, 2011. URL: <https://ieeexplore.ieee.org/document/6162887>.
- [11] A. Sanna; F. Manuri; F. Lamberti; G. Paravati; P. Pezzolla. *Using handheld devices to support augmented reality-based maintenance and assembly tasks*, 2015. URL: <https://ieeexplore.ieee.org/document/7066370>.
- [12] Xuyue Yin, Xiumin Fan, Jiajie Wang, Rui Liu, Qiang Wang. *An Automatic Interaction Method Using Part Recognition Based on Deep Network for Augmented Reality Assembly Guidance*, 2018. URL: <https://doi.org/10.1115/DETC2018-85810>.
- [13] J. Brooke. *SUS: a "quick and dirty" usability scale*, 1986. URL: <https://www.usabilitest.com/>.
- [14] P. W. Jordan; B. Thomas; B. A. Weerdmeester; A. L. McClelland. *Usability Evaluation in Industry*, London: Taylor and Francis, 1996.