

SLAM TECHNOLOGY FOR MIXED REALITY

OVERVIEW DELL'AUGMENTED REALITY E DELLA VIRTUAL REALITY

di Tiziana Primavera



Fig. 1 - Applicazione A.R. di tipo mobile.

Tra le tecnologie interattive di visualizzazione, la tecnologia informatica, derivante dalla *Computer Vision*, nota comunemente come Realtà Aumentata (acronimo AR dal termine inglese *Augmented Reality*), costituisce oramai uno dei *tools* più ottimali, agevoli e performanti al servizio dell'uomo in innumerevoli contesti applicativi. (Fig.1)

Fine unico e prevalente di questa tecnologia dal carattere disruptive è sostanzialmente quello di fornire dati visuali e di natura multimediale, per supportare processi di apprendimento o decisionali.

E' possibile asserire che la suddetta tecnologia informatica persegue un'unica finalità, quella di fornire *un'esperienza percettiva che risulti più esaustiva possibile*, consentendo la visualizzazione di dati di natura anche eterogenei, *ma sempre riferiti al contesto d'uso dell'applicazione*.

Spesso erroneamente confusa con la Realtà Virtuale, (acronimo VR dal termine inglese *Virtual Reality*), se ne differenzia per le finalità antitetichette perseguite, pur condividendo con essa gran parte del *know how* tecnologico.

La Realtà Aumentata studia sistemi proiettivi di contenuti digitali *da comporre ad un flusso di dati provenienti dal mondo reale*, ed è pertanto strettamente correlata allo spazio circostante in cui l'utente è posto per fruire attivamente dell'applicazione.; diversamente la Virtual Reality, prescinde totalmente dallo spazio reale, in quanto il suo obiettivo principale, in netto contrasto con la Realtà Au-

Una panoramica sulla tecnologia SLAM, gli SDK di ultima generazione, la Mixed Reality, l'Augmented Reality e la Virtual Reality.

mentata, è la creazione di un mondo interamente sintetico, la cui caratteristica prevalente è quella di una forte e preponderante immersività.

Nel caso della VR, la visione stereoscopica, l'altissima risoluzione degli attuali *headset* preposti alla sua fruizione, caratterizzati dal *field of view* piuttosto ampio, contribuiscono certamente ad accrescerne il fattore immersivo, consentendo di veicolare stimolazioni sensoriali altamente impattanti e verosimili, recepite così profondamente dal sistema sensoriale umano, da poter essere oramai impiegate in diversi protocolli riabilitativi in contesti medico curativi di numerose patologie psicologico- comportamentali. Per quanto attualmente impiegata in ambiti simulativi, preposti alla formazione professionale o al *training on the job*, *killer application* della VR, è stato e rimane certamente il mondo del *gaming*, potendo la tecnologia tranquillamente simulare mondi complessi fotorealistici tridimensionali, completamente immaginifici, con cui interagire generalmente attraverso dispositivi manuali denominati *controller* (Fig.2).

Allo stato dell'Arte, anche in merito all'*Augmented Reality*, il superamento sensoriale può non risultare limitato al solo senso visivo: in questi sistemi interattivi di ultima generazione, grazie alla loro maturità tecnologica ed ai considerevoli progressi conseguiti nella sensoristica, l'utente di un'applicazione può essere oramai posto nella condizione di poter interagire liberamente con i dati digitali (Fig.3), movimentandoli agevolmente nello spazio, secondo una codicistica gestuale predefinita al fine di impartire diversi *input* di comando (*spatial computing*).

Attualmente, proprio in relazione alla *user experience* delle attività modali di *input*, settori di ricerca specifici sono rivolti al perseguimento della naturalezza dei comandi di natura gestuale, mutuando ed integrando negli algoritmi, concetti derivanti dalle neuroscienze.

Secondo un'ottica comparativa tra le due scienze citate, quella che certamente si colloca con maggior rispondenza nella categoria delle tecnologie contemporanee di natura *Ubicomp*, è certamente quella della *Augmented Reality*, si tratta infatti di una tecnologia "trasparente", in grado di dissimulare l'apporto tecnologico che la contraddistingue, non evidenziandolo, ma trattandolo come una sorta di sfondo.

La percezione dell'interazione con gli elaboratori tende infatti, nel caso dell'AR, alla massima trasparenza, integrando nell'ambiente elementi digitali, ma evitando che l'utente percepisca le modalità con le quali si verificano tali cambiamenti, secondo concetti di sviluppo ed implementazione, che contraddistinguono *l'ubiquitous computing*.

SLAM - SIMULTANEOUS LOCALIZATION AND MAPPING

L'*Augmented Reality* inoltre una tecnologia definita *context aware*, ovvero *consapevole del contesto*, e questa definizione trova il suo fondamento con la recente integrazione, nelle sue architetture informatiche costituenti, di una tecnologia articolata, denominata SLAM, acronimo per definire i termini esplicativi "*Simultaneous Localization and Mapping*" ovvero "Localizzazione *real-time* e Mappatura dell'ambiente" (Fig.4).

Grazie all'integrazione dello SLAM, termine generico, che può contemplare un insieme diversificato di procedure informatiche, preposte per eseguire specifici obiettivi di tracciamento dello spazio reale circostante, *hardware* e *software* sono predisposti per monitorare il movimento del dispositivo nello spazio e al contempo una mappatura ambientale, tradotta con una *mesh* restitutiva delle geometrie spaziali.

Sono numerosi gli ambiti applicativi dello SLAM, esso costituisce uno dei fattori chiave che sottendono la progettazione dei veicoli privi di pilota, dei droni, delle auto di nuova generazione a guida autonoma, della robotica ed anche delle applicazioni di Realtà Aumentata. Lo studio di questa tecnologia di orientamento e consapevolezza spaziale, che attualmente contraddistingue le caratteristiche di alcuni *hardware top level* di *Mixed Reality*, ha avuto origine nei settori afferenti il mondo della robotica.

"WHERE I'M":

Nell'ambito della robotica era il quesito cardine, cui dare risposta, per consentire la movimentazione degli automi: ogni robot per potersi muovere all'interno di uno spazio sconosciuto, necessitava di una serie di dati informativi, che consentivano alla macchina di assumere autonomamente consapevolezza della propria posizione in esso.

Suddetti dati erano riassumibili sinteticamente in una mappa da aggiornare (*mapping*) ed una mappa da consultare (*localization*), proprio lo SLAM è l'algoritmo che richiede la soluzione di entrambi i quesiti suindicati.

Le implicazioni derivanti dalla sua adozione, nel settore dell'AR sono certamente enormi, con esso si è finalmente giunti ad un'accuratezza davvero considerevole nel delicato processo di collimazione dei dati reali con quelli digitali (*i.e. dati tridimensionali, derivanti dalla diagnostica per immagini sovrapposti al corpo del paziente durante interventi chirurgici*).

Numerosi e diversi tipi di sensori, GPS, giroscopi, telecamere, accelerometri etc concorrono ad individuare molteplici dati, che vengono successivamente elaborati, grazie ad una serie di complessi calcoli e algoritmi, per poter costruire una mappa tridimensionale.

Tale mappa ambientale consta di una *mesh*, ottenuta da una nuvola di punti piuttosto accurata (*dense point cloud*) e viene contemporaneamente impiegata per identificare la posizione del *device* di fruizione, collocato al suo interno. Indubbiamente le implicazioni derivanti processi SLAM nell'ambito di applicazioni di *Augmented Reality*, sono state enormi, al punto da definire nuovi *standard* e tassonomie contemporanee.



Fig. 2 - Lancio della playstation VR (presentazione Sony - <http://fortune.com/2016/01/05/virtual-reality-game-industry-to-generate-billions/>).

MIXED REALITY - SLAM, OCCLUSION E COLLISION

Proprio la consapevolezza dei dati geometrici dell'ambiente (che contraddistingue un'applicazione evoluta AR contemporanea), consente infatti al sistema di poter trattare la *mesh*, definendo quello specifico effetto visuale percettivo di perfetta integrazione del dato digitale nel reale, chiamato *occlusion*.

Inoltre sempre grazie alla suddetta consapevolezza spaziale del sistema, è consentito definire interazioni avanzate di natura anche fisica intercorrenti tra dati digitali e dati geometrici del reale, denominati "collision"

Entrambe le due interazioni indicate (*occlusion* e *collision*), contraddistinguono potenzialità specifiche derivanti dalla perfetta fusione dei dati di natura reale e digitale, e costituiscono le principali peculiarità dei sistemi moderni di *Augmented Reality*, denominati con il nuovo termine di *Mixed Reality*.

La *Mixed Reality*, che incorpora la tecnologia SLAM ha assunto pertanto un preciso significato, da non confondere con quello indicato dal medesimo termine nell'originaria tassonomia di Milgram¹ del 1994 (Fig. 5).

All'interno di tale tassonomia, di rilevante importanza, il termine di *Mixed Reality* stava ad indicare semplicemente quell'insieme di sistemi interposti tra dati puramente reali e quelli puramente virtuali, ma nel biennio a cavallo del 2015, la propaganda pubblicitaria promossa da Microsoft per il lancio dei propri *hardware*, ha definito nuovi paradigmi identificativi, influenzando considerevolmente il settore.



Fig. 3 - Codici gestuali (<https://docs.microsoft.com/en-us/windows/mixed-reality/gestures>)

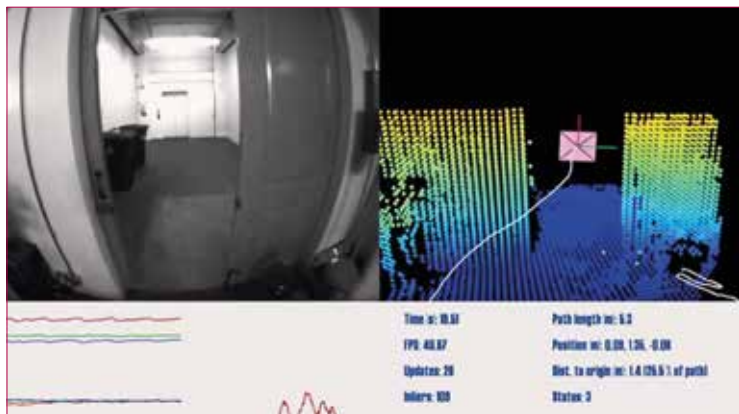


Fig. 4 - Project Tango. Progetto del 2014, conclusosi nel 2018. Il programma pioniere nel settore, consisteva nel realizzare (in collaborazione) con produttori scelti una serie di smartphone dotati di sensori speciali e strumentazioni preposte per la tecnologia SLAM (<https://www.slashgear.com/google-project-tango-will-3d-map-the-whole-world-with-hidof-20317590/>).

Così, ancora oggi, con il termine di *Mixed Reality* si vuole generalmente far riferimento a quel livello massimo di interattività tra dato digitale e dato reale, (consentito grazie ad una perfetta collimazione di natura geometrica reale/ virtuale), mentre il termine *Augmented Reality* intende un'illusione percettiva di elementi sintetici sovrapposti alla scena osservata, percepiti prospetticamente ed apparentemente integrati in essa, nell'ambito però di un sistema, che non ha assoluta consapevolezza spaziale-geometrica del contesto.

Con l'avvento dell'integrazione di sistemi SLAM nelle applicazioni AR, si sono aperti innumerevoli nuovi scenari applicativi, estremamente evoluti, in una parabola crescente di utilità, anche per ambienti di grandi dimensioni ed all'aperto, che spaziano dalle VIA alla prefigurazione di progetti BIM etc, consentendo di impiegare modelli 3D complessi in scene interamente tracciate e garantendone la visualizzazione completa e il miglior posizionamento nell'ambiente.

SDK DI NUOVA GENERAZIONE

E' possibile riconoscere che per le tecnologie AR/MR, il 2017 sia stato un anno nevralgico, per via della disponibilità di due SDK estremamente affidabili, l'ARKit per il sistema operativo iOS e l'ARCore, per l'ecosistema Android. Si è infatti assistito ad una adozione *mainstream* tra gli sviluppatori, che hanno accolto i due *tools* con particolare entusiasmo, per via delle procedure semplificate di comprensione - implementazione e del *tracking* estremamente robusto.

La validità dell'Arkit risiede in diverse ragioni, tra di esse una è certamente l'impiego di un sistema ibrido di tracciamento.

Esso si avvale infatti, simultaneamente di un tracciamento ottico ed uno di natura inerziale (IMU), che collaborano per una corretta definizione del punto di vista dell'osservatore: l'individuazione di *natural features* di superfici planari, definisce una *sparse point cloud* (Fig.6), da cui il sistema, semplificando considerevolmente il *workflow*, estrapola tre soli punti per la ricostruzione della superficie osservata, con numerosi vantaggi in merito ai processi computazionali; tale piano così intercettato, successivamente costituisce il riferimento per il posizionamento del contenuto digitale.

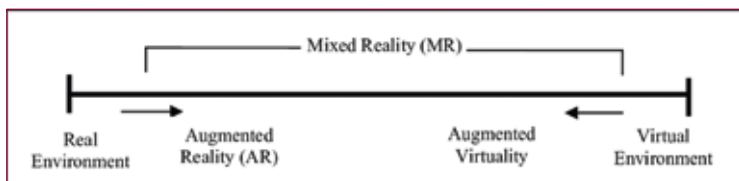


Fig. 5 - Milgram Reality - Virtuality Continuum.

I sistemi di tracciamento visivo e inerziale si basano su sistemi di misurazione completamente diversi senza interdipendenza; il fattore positivo è che i punti di forza di ciascun sistema compensano i punti deboli dell'altro.

Ciò significa che la fotocamera potrebbe anche essere oscurata, risultando l'applicazione piuttosto stabile anche in condizioni di luminosità estremamente variabile e altresì potrebbe visualizzare una scena con poche caratteristiche naturali (come un muro bianco).

WHAT'S NEXT..

Gli scenari di sviluppo delle tecnologie interattive di visualizzazione sono estremamente magmatici e gran parte della ricerca è concentrata sulla segmentazione dei dati geometrici derivanti dalla tecnologia SLAM, grazie all'integrazione nei loro sistemi dell'intelligenza artificiale (A.I.) per l'interpretazione semantica dei dati, ed al contempo è in corso una ricerca attiva per sostenere la Ricostruzione 3D fotorealistica (*depth perception*) in tempo reale, non impiegando telecamere di profondità, ma mediante l'utilizzo semplificato di una singola telecamera RGB.

Tutto ciò costituisce solo un accenno al panorama di ricerca inerente il settore, e non c'è da stupirsi di tanto fermento, in quanto oramai i grandi colossi economici mondiali concorrono alacremente per ottenere la *leadership* dell'innovativo mercato XR, un mercato di una tecnologia destinata ad essere annoverata come *l'ottavo media sociale*.



Fig. 6 - Project Tango. Analisi comparativa dei due SDK ArCore ed ARKit di prima generazione (<https://www.youtube.com/watch?v=dNXBvDKRg1M>).

NOTE

1 Milgram Reality - Virtuality Continuum - in quegli anni di ricerche di natura pionieristica, una classificazione così dettagliatamente articolata, soddisfaceva al contempo sia le esigenze di un potenziale sviluppatore, che quelle di un eventuale fruitore, consentendo al primo di individuare con una certa chiarezza il paradigma idoneo di visualizzazione per l'implementazione del proprio contributo ed al secondo di conoscere aprioristicamente le caratteristiche e le peculiarità inerenti uno specifico sistema.

ABSTRACT

An overview on SLAM technology, the latest SDK generation, the Mixed Reality, the Augmented Reality and the Virtual Reality.

PAROLE CHIAVE

COMPUTER VISION ; TECNOLOGIA SLAM; MIXED REALITY; REALTÀ AUMENTATA; REALTÀ VIRTUALE; SDK;

AUTORE

TIZIANA PRIMAVERA
TIZIANA.PRIMAVERA@UNIROMA1.IT
DOCENTE UNIVERSITARIO, AUGMENTED REALITY PHD - SAPIENZA

Soluzioni e Tecnologie
Geografiche per
la Trasformazione
Digitale

**THE
SCIENCE
OF
WHERE**

www.esriitalia.it



esri Italia
THE SCIENCE OF WHERE