

Interazioni digitali - Realtà aumentata

La realtà aumentata (*Augmented, Mixed* o anche *Extended Reality*, in breve AR) è la modalità che ci consente di vedere il mondo reale con l'aggiunta, posta a registro, di informazioni digitali.

Pertanto, il meccanismo complessivo consiste in: acquisire informazioni sullo spazio e sugli oggetti posti nello spazio -> elaborare le informazioni, riconoscendo oggetti e relazioni -> rappresentare in forma digitale oggetti e altre informazioni in combinazione con gli oggetti e lo spazio fisico.

Nel campo del suono, possiamo definire "realtà aumentata" l'udire contemporaneamente e in stretto abbinamento suoni reali e suoni prodotti da software.

Le prime e più significative parole chiave della realtà aumentata sono dunque: **unione, sovrapposizione**. Mentre la realtà virtuale (*Virtual Reality*, in breve VR) sostituisce completamente la realtà fisica percepita dai nostri occhi con una "realtà" digitale stereoscopica (e in alcuni casi interviene anche saturando altri sensi, come l'udito e il tatto), la realtà aumentata agisce sovrapponendo informazioni, non sostituendole. Mantenere il contatto diretto con la realtà sensibile, consente anche di poter collaborare più facilmente con altre persone, siano esse dotate o meno di strumenti utili per la realtà aumentata. Infatti, di norma, la **fruizione di contenuti in realtà virtuale rende difficoltoso interagire con altre persone**, pur se esistono alcuni device per la realtà virtuale dotati anche di due videocamere (vedi Intel Project Alloy, ora dismesso): in tal modo è possibile comporre assieme il modello digitale con le riprese reali. In questo caso l'interazione si potrebbe definire come Augmented Virtuality.

Possiamo identificare **tre modalità** di fruizione della realtà aumentata, **anche in combinazione tra loro**:

- Mediante **occhiali/visori** (ad esempio, Microsoft HoloLens, MagicLeap, ...) – Modalità immersiva – Rappresentazione legata al punto di vista dell'utente – Tipicamente monoutente.
- Mediante **smartphone o tablet** (usando App e ARKit per iOS e ARCore per Android o altri SDK) – Meno immersivo – Rappresentazione legata al punto di vista dell'utente – Mono/multi utente.
- Mediante **proiezione** (Spatial Augmented Reality, SAR) – Immersiva – Rappresentazione non legata al punto di vista dell'utente – Tipicamente multiutente.

Pertanto, una ulteriore suddivisione è costituita dalla necessità o meno di usare **strumenti invasivi** (non naturali), come occhiali o visori stereoscopici. In tal senso, l'uso di smartphone o tablet si pone in una posizione intermedia, tuttavia da considerare "naturale" e quasi "trasparente" (tuttavia scarsamente immersiva), essendo questi strumenti ormai di uso comune.

Tre principali categorie di sensori sono utilizzate nei sistemi AR:

1. Sensori utilizzati per il **tracciamento** (*tracking* di posizione e orientamento, assoluti e relativi).
2. Sensori per la raccolta di **informazioni ambientali**, ad esempio il riconoscimento di oggetti, suoni, colori.
3. Sensori per la raccolta dell'**input dell'utente**, tocco, gesture, direzione dello sguardo, comandi vocali, ...

Nel primo caso sono utilizzate videocamere, utili per la *computer vision*, eventualmente con l'aiuto di marker fiduciali non simmetrici, come ad esempio i QRcode. Sono anche utilizzati sensori GPS, giroscopi, accelerometri e altri sensori utili a determinare alcuni gradi di libertà (traslazioni e rotazioni).

Sono poi anche impiegati sensori per la scansione 3D dello spazio circostante (scanner 3D e restituzioni fotogrammetriche) ed altri sensori utili per raccogliere maggiori informazioni sull'ambiente (sensori di temperatura, CO₂, Ph, luce, velocità del vento, umidità, suono, ecc.).

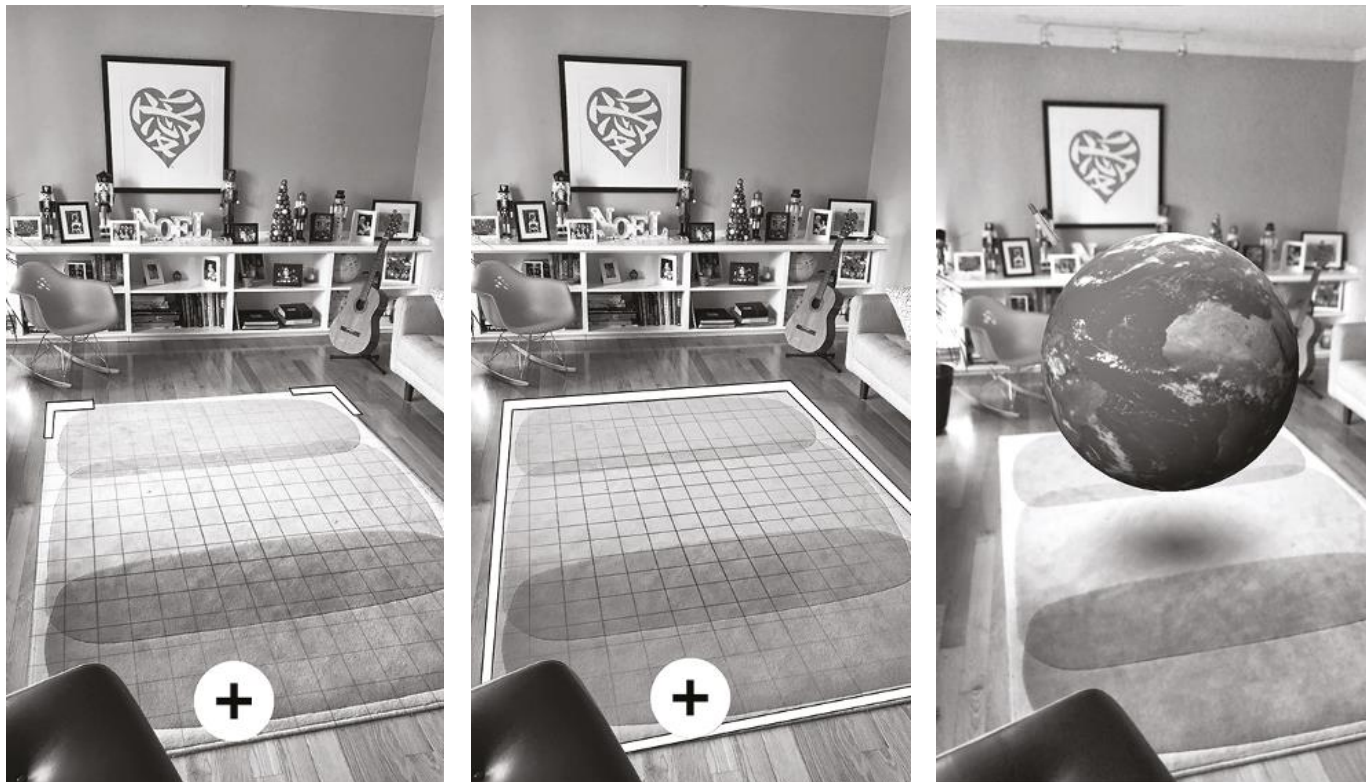
Infine, oltre ai sensori passivi visti fino ad ora (ad esclusione degli scanner 3D), sono anche utilizzati sensori attivi, utili per raccogliere informazioni relative all'input dell'utente (informazioni in parte ottenute anche mediante elaborazione). Di questa categoria fanno parte gli oggetti icona così come le superfici touch.

Infine, un aspetto essenziale e spesso trascurato nel progetto, è dato dal tempo di latenza, vale a dire il lasso di tempo necessario per acquisire, elaborare e inviare le informazioni.

Per ultimo, è anche da considerare la prossima larga diffusione della banda 5G.

Alcune modalità di interazione per la realtà aumentata

Tocco. Esempio: **tocco** del tappeto per identificarlo (a sinistra); **feedback** di identificazione avvenuta del piano sul quale si vuole operare (al centro); **trascinamento** dell'oggetto digitale sul piano individuato.



Gesture. Tracciate sia da camere esterne sia da camere interne del device. In questo ultimo caso, spesso la difficoltà principale consiste nell'impossibilità di vedere le proprie mani inquadrare dalle camere interne (inoltre, nel caso di smartphone/tablet almeno una mano è occupata a sostenere il device).

Da considerare pertanto anche la possibilità di interpretare le gesture (o i movimenti) di altre persone presenti sulla scena, in funzione di avviare con esse una collaborazione attiva.

Click sulla gesture come metodo di selezione.



Direzione dello sguardo. Utile per identificare una direzione o una posizione nello spazio. Può essere rilevata sia dal device (visore o smartphone/tablet) sia da camere esterne. Nei visori/smartphone/tablet il controllo può essere facilitato da una griglia prospettica di riferimento, che individua il piano attivo (vedi esempio con ARCore). In altri casi una sorta di mirino aiuta a individuare singoli oggetti mediante una osservazione persistente: l'oggetto potrà poi essere evidenziato per una conferma. Una tecnologia in via di perfezionamento, l'*eye tracking*, consentirà una maggiore facilità, naturalezza e precisione nell'individuare singoli oggetti nello spazio.



Voce. Il riconoscimento vocale è una tecnologia già matura e in rapidissimo sviluppo (vedi ad esempio Amazon Alexa e le APP sia per iPhone sia per Android, oltre alle applicazioni Cortana e Siri). Unito alla sintesi vocale, costituisce un meccanismo fondamentale di interazione naturale – spesso in combinazione con altri – sia per la selezione/conferma di azioni sia per l'introduzione o la comprensione di testi.

Motion controller / oggetti icona. Alcuni visori comprendono semplici device (vedi, ad esempio, l'HoloLens Clicker a sinistra o il Magic Leap One Control, a destra) che consentono all'utente di effettuare **click**, **scorrimento**, **rotazione** ed altro. Tali dispositivi hanno lo scopo di colmare il divario tra "strumenti" del tutto naturali (come gesture, sguardo, voce) con strumenti non naturali come tastiera e mouse. Di norma, tali strumenti possono essere usati solo in accoppiamento con i relativi visori. Tuttavia, è senz'altro da considerare la possibilità di utilizzare oggetti icona per innescare forme di collaborazione attiva.



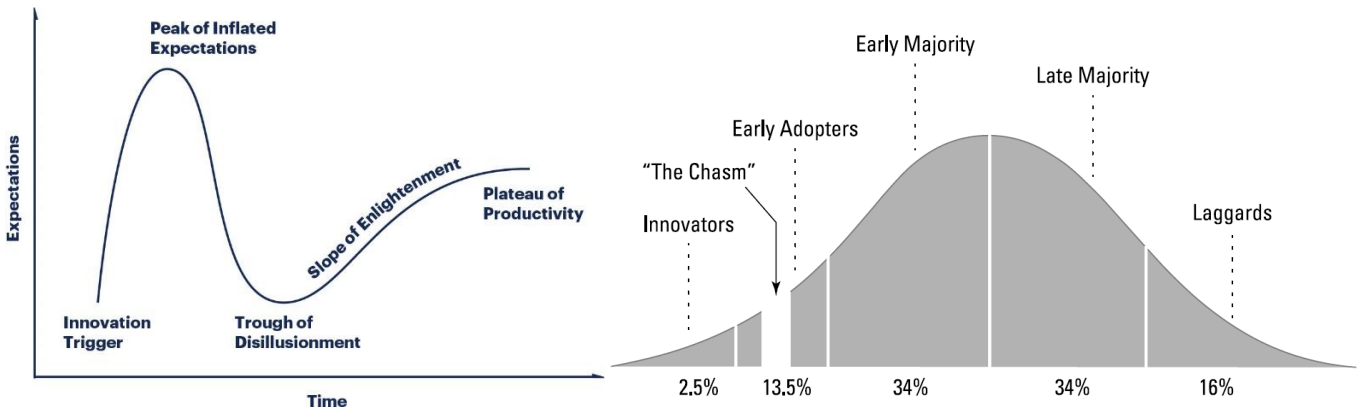
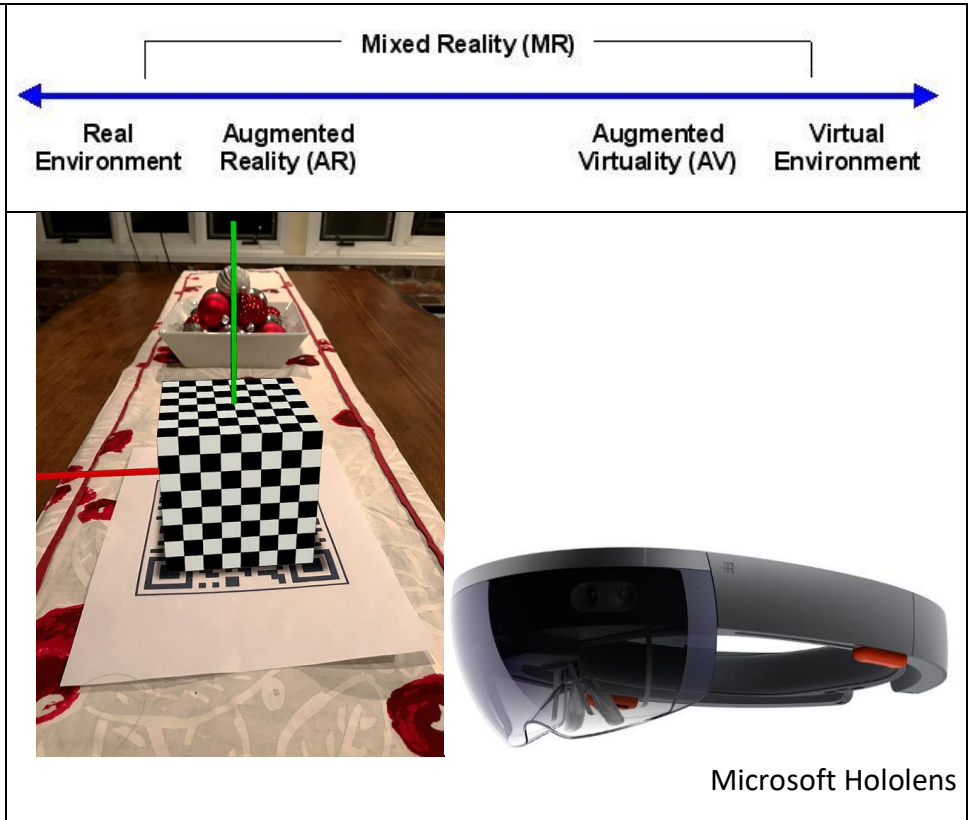
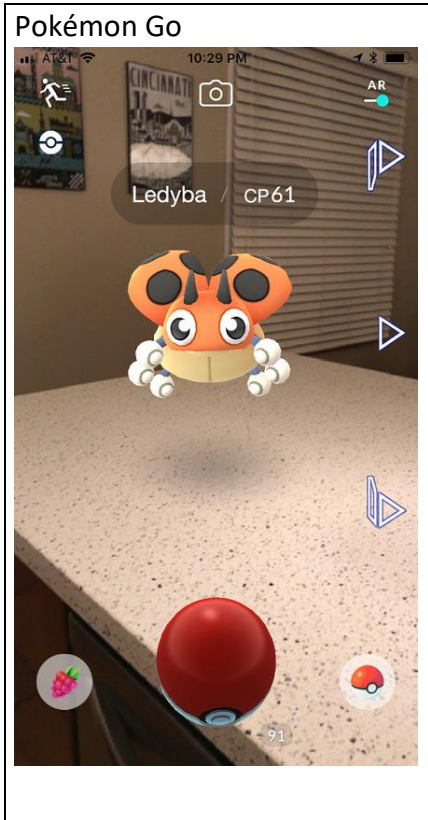
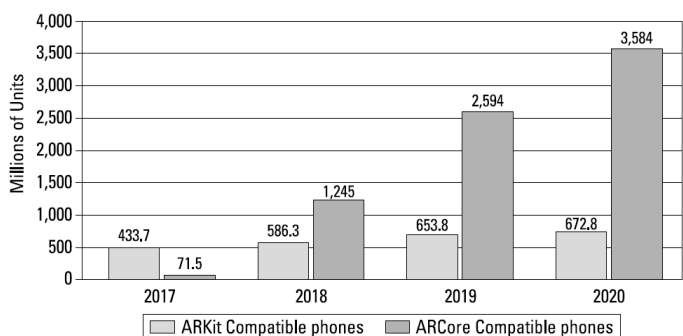


Figure 3-1 shows the growth rate for ARKit- and ARCore-enabled devices over the next few years, as predicted by ARTillery Intelligence. ARTillery predicts nearly 4.2 billion handheld AR devices will be in consumers' pockets by 2020, a massive market and an incredible rate of growth. Apple is predicted to lead the charge at first of mobile AR, with Google and Android catching up and surpassing iOS as replacement cycles for Android devices kick in a few years out.

TABLE 2-1 Virtual Reality Desktop Headset Comparison

	HTC Vive	Oculus Rift	Windows Mixed Reality	PlayStation VR
Platform	Windows or Mac	Windows	Windows	PlayStation 4
Experience	Stationary, room-scale	Stationary, room-scale	Stationary, room-scale	Stationary
Field of view	110 degrees	110 degrees	Varies (100 degrees)	100 degrees
Resolution per eye	1,080 x 1,200 OLED	1,080 x 1,200 OLED	Varies (1,440 x 1,440 LCD)	1,080 x 960 OLED
Headset weight	1.2 pounds	1.4 pounds	Varies (0.375 pound)	1.3 pounds
Refresh rate	90 Hz	90 Hz	Varies (60-90 Hz)	90-120 Hz
Controllers	Dual motion wand controllers	Dual motion controllers	Dual motion controllers, inside-out tracking	Dual PlayStation move controllers



Motion tracking (<https://www.librasoft.cloud/news/arkit-vs-arcore-cosa-scegliere-per-sviluppare-esperienze-di-augmented-reality/>)

La realtà aumentata necessita che il dispositivo sia in grado di tracciare la propria posizione e l'orientamento all'interno del mondo reale. Sia ARKit che ARCore ottengono questo risultato grazie alla **VIO: Visual Inertial Odometry**.

La VIO si basa sui dati rilevati dai sensori di movimento (giroscopio, bussola, GPS, ...) e alla fotocamera per identificare il movimento del dispositivo nei 6 gradi di libertà (**degrees of freedom**, i cosiddetti **6 DOF**: 3 di traslazione: sinistra/destra, sopra/sotto, avanti/indietro, e 3 di rotazione: beccheggio, rollio, imbardata). La corretta stima della posizione reale del dispositivo è la chiave fondamentale per valutare le coordinate spaziali 3D dell'esperienza AR all'interno della quale posizionare e tenere traccia degli elementi virtuali (3D e 2D), e come essi dovranno essere mostrati sullo schermo del device.

Plane detection

Una volta acquisito il sistema di coordinate spaziali, l'esperienza di AR "realistica" necessita che il dispositivo sia in grado di capire dove posizionare gli oggetti virtuali rispetto agli elementi del mondo reale: "sopra" quale piano? "Attraverso" quale parete? Rotolando o spostandosi su quale oggetto? ...

Sia ARCore che ARKit fanno uso di tecniche di **riconoscimento ambientale**, e in particolare di *plane detection*, per permettere agli sviluppatori di lavorare direttamente e con modalità semplici, al netto di sofisticati algoritmi basati su gradienti di luci e calcolo della profondità.

Il risultato è quindi la possibilità di posizionare gli oggetti virtuali su piani virtuali a loro volta basati sulle superfici reali precedentemente identificate.

Lighting estimation

Dopo aver capito dove posizionare gli oggetti virtuali, e rispetto a quali oggetti reali, il tocco finale che rende "credibile" la realtà aumentata è dato dal rendering degli oggetti, con particolare riferimento all'illuminazione delle superfici, che deve essere del tutto compatibile con l'illuminazione reale.

Entrambi gli SDK interpretano i gradienti di luce del mondo reale per capire quali sono le sorgenti di luce reale, per poi illuminare in maniera analoga gli oggetti virtuali, proprio come se fossero disposti nel mondo reale. Ovviamente, entrambi i *framework* permettono l'aggiunta e l'alterazione delle diverse sorgenti di luce virtuale, anche in modo dinamico.

Altre caratteristiche comuni

Con approcci diversi, sia ARKit che ARCore propongono **interessanti caratteristiche secondarie**:

- il tracciamento di oggetti 2D e 3D in movimento
- l'utilizzo di oggetti 2D e 3D come *trigger* in grado di fare partire una interazione o una nuova esperienza
- l'utilizzo di *anchors* e *feature points*, ossia punti specifici del mondo reale mappati nel mondo virtuale
- il supporto per *framework* di terze parti (es. Unity) e di hardware specializzati (es. Oculus)
- il supporto a tecniche avanzate di *image detection* (es. per *face recognition* o lettura qr code)
- l'integrazione con gli altri *framework* nativi delle due piattaforme (geolocalizzazione, *cloud storage*, ...)

Differenze: quantità vs qualità

Le differenze tecniche sono minimali: bisognerebbe scendere in dettagli implementativi per capire eventuali pro e contro delle scelte adottate da Apple (ARKit) e Google (ARCore) nei rispettivi algoritmi di *motion tracking*, *plane detection* e *lighting estimation*.

Tuttavia, a livello più macroscopico si può sperimentare che **ARCore riesce a mappare mondi reali più vasti** all'interno di una stessa scena virtuale, senza bisogno di riavviare la sessione AR ex novo, identificando più *feature point* nel mondo reale, e facendolo più rapidamente.

D'altro canto, **ARKit è solitamente più accurato** nella scelta dei *feature point*. Inoltre, dal momento che ARKit è sul mercato da più tempo ed espone un'API (*Application Programming Interface*) più matura, è il *framework* preferito da alcuni dei player non hi-tech che vogliono cimentarsi nella realtà aumentata (es. IKEA).

Un terzo SDK, **Vuforia**, pur possedendo meno funzioni rispetto ai primi due *framework*, utilizzando le funzioni sia di ARKit sia di ARCore, è compatibile con entrambe le piattaforme.

Link:

<https://socioherald.com/global-head-mounted-display-hmd-market-global-industry-analysis-and-forecast-2018-2026/290807/>

<https://www.wevolver.com/article/new.augmented.reality.head.mounted.display.offers.unrivalled.viewing.experience>

https://xr.plus/?gclid=EAIaIQobChMIq6mcq8Gh5gIVhs13Ch2isAxFEAAYAiAAEgKlofD_BwE

<https://medium.com/libertyit/the-2019-a-to-z-of-augmented-reality-474f595b687a>

<https://medevel.com/16-virtual-reality-vr-frameworks/>

<https://www.macworld.com/article/3446839/apple-ar-glasses-rumors.html>

<https://www.magicleap.com/>

<https://next.reality.news/news/meta-rises-from-ashes-under-new-ownership-along-with-new-ceo-0198056/>

<https://www.microsoft.com/it-it/hololens>

<https://www.producthunt.com/alternatives/intel-vaunt>

<https://www.bynorth.com/>

<https://www.manomotion.com/>

<https://www.mirareality.com/>

<http://realmaxinc.com/>

<https://avegant.com/>

<https://wayray.com/>

<https://daqri.com/worksense/>

<https://www.vuzix.com/products/blade-smart-glasses>

<https://mobidev.biz/blog/augmented-reality-future-trends-2018-2020>

<https://www.magicleap.com/news/product-updates/connect-with-friends-with-avatar-chat>

<https://techspo.co/2020-augmented-reality-events/>

<https://developer.vuforia.com/>

<https://bluewhaleapps.com/blog/comparing-arkit-vs-arcore-vs-vuforia-the-best-augmented-reality-toolkit>