



INTRODUZIONE ALLA COMPUTER GRAFICA

Con il termine Computer Graphics, solitamente abbreviato **CG**, ci riferiamo ad una sezione dell'informatica che sviluppa tecniche capaci di simulare visivamente la realtà. La realtà è normalmente **tridimensionale** e, quindi, anche la CG deve rappresentare oggetti con coordinate spaziali a tre dimensioni.

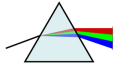
D'altro canto, con il termine Virtual Reality, abbreviato **VR**, facciamo riferimento ad una branca della CG che si occupa di creare ambienti e oggetti 3D reali, offrendo anche l'interazione con essi, fornendo all'utente la possibilità di interagire con un **mondo virtuale**.

Ci teniamo a specificare che i dispositivi di visualizzazione 3D sono normalmente **bidimensionali**, come schermi e monitor. Per questo tendiamo a parlare di **rappresentazione**, in 2D o 3D, e non di immagine stessa. Possiamo distinguere una rappresentazione **raster**, cioè quando l'immagine viene divisa in tanti pixel quadrati, in cui ognuno di loro possiede informazioni cromatiche, solitamente in formato RGB, e di dimensione. Al contrario, una rappresentazione **vettoriale** si basa su forme geometriche, quindi linee, punti e curve.

Una qualsiasi immagine, che sia in raster o vettoriale, è un array bidimensionale di informazioni cromatiche. L'array può avere una **risoluzione spaziale**, dimensioni, e una **risoluzione colore**, che ne definisce la profondità. Maggiore è la risoluzione, maggiori sono le dimensioni. Nella risoluzione è molto importante l'**Aspect Ratio**, cioè il rapporto tra componenti orizzontali e verticali.

Con la CG è fondamentale produrre immagini raster, partendo da una descrizione vettoriale, che viene detta **modello tridimensionale**, su coordinate X, Y e Z, rispetto ad un sistema di riferimento, detto **nuvola di punti**. I punti della nuvola sono strettamente connessi tra di loro e formano una **mesh**, una superficie poligonale, i cui componenti sono punti, lati e facce (**points, edges and faces**). Normalmente si usano poligoni basilari, come triangoli o altri poligoni quadrilateri. Con il crescere del numero dei poligoni si avrà una più fedele rappresentazione della realtà.

Una delle prime rappresentazioni visive in 3D venne detta **Wireframe**, cioè la descrizione di un oggetto tramite i lati dei poligoni che lo compongono.



L'EQUAZIONE DELLA RADIANZA

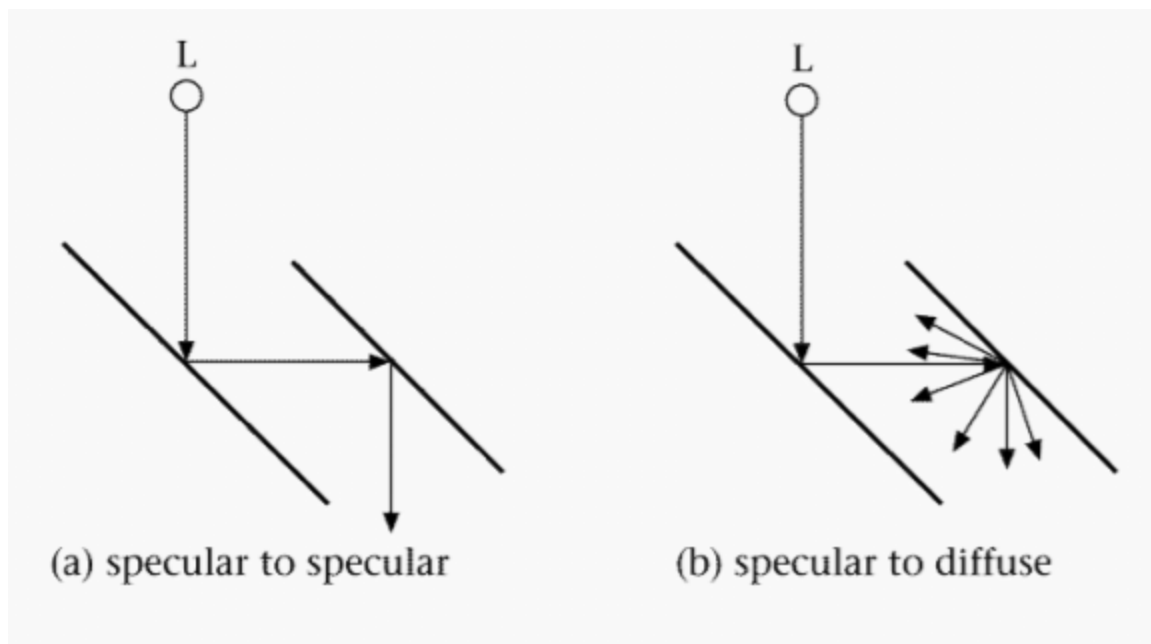
La **luce** è da tempo uno dei problemi principali della CG, nello specifico della trasformazione in real-time. Questo perché bisogna prevedere molte cose, come l'intensità della luce, le ombre che proietta, i riflessi che provoca su diversi tipi di superficie. E questo non è neanche il problema più grande. La luce visibile è una **radiazione elettromagnetica**, con una lunghezza d'onda che va da 400 a 700 nm, cioè nanometri. La luce può essere vista anche come una **particella**. Le particelle di luce sono dette **fotoni**, cioè pacchetti di energia che viaggiano nel vuoto a velocità costante. Il principale problema della luce è il **trasporto**, cioè come essa interagisce con una superficie, in un volume di spazio. Il **flusso**, simbolo Θ , è il tasso di energia che scorre attraverso una superficie, ogni unità di tempo. L'energia è proporzionale al flusso di fotoni che vengono trasportati. Si dice che il flusso totale in un volume si trova in **dinamico equilibrio**, con le particelle che fluiscono a velocità costante. Parliamo di **conservazione dell'energia** quando essa viene assorbita dalla materia all'interno del volume.

La luce ha un'equazione fondamentale. In input essa viene emessa nel volume, scorrendo dall'esterno. In output viene riflessa nella materia e poi assorbita. Quindi **INPUT = OUTPUT**. La formula fondamentale $\Theta(p, w)$ denota il flusso p nella direzione w . Conoscere bene queste diciture ci permette di mettere a fuoco e di risolvere il grosso problema relativo al **rendering**, cioè la conversione di un'immagine bidimensionale in un oggetto tridimensionale.

La **radianza** L è il flusso che lascia una superficie irradiata. Nella CG non parliamo in termini di fotoni, ma in termini di **raggi**. Ogni raggio ha la sua radianza. La **radiosità** è il flusso per area che irradia da una superficie. L'**irradianza**, invece, è il flusso per area che arriva in una superficie. L'unità di misura fondamentale per la riflettanza è la **BRDF**, acronimo di Bidirectional-Reflectance-Distribution-Function. Una volta fatte queste premesse, possiamo cimentarci sul vero oggetto di studio di questo paragrafo, ossia l'**equazione della radianza**. Una radianza $L(p, w)$, in un punto p e in una direzione w , è la somma della radianza emessa e della radianza totale riflessa. L'equazione della radianza incorpora tutte le proiezioni in 2D che serviranno poi per la trasformazione tridimensionale tramite il rendering. In parole

povere, bisogna trovare soluzioni approssimative a questa equazione per poter effettuare il rendering. In genere abbiamo:

- **metodi locali:** tengono conto solo degli effetti delle sorgenti luminose, trascurando le riflessioni tra gli oggetti.
- **metodi globali:** si occupano di quest'ultima parte. In particolare, questa procedura ingloba la tecnica del *ray-tracing*, riflessioni speculari, e del *radiosity*, riflessioni diffuse.



RAY-TRACING

RADIOSITY



VISIONE

Cosa intendiamo quando parliamo di visione? Un vocabolario direbbe che la visione è un processo di percezione degli stimoli luminosi, unito alla funzione e alla capacità di vedere. In termini tecnologici parliamo di **visualizzazione**, cioè la comunicazione di informazioni tramite rappresentazioni grafiche. Il concetto di **percezione** è davvero fondamentale, cioè è importante capire come l'uomo elabora gli stimoli visivi. Innanzitutto possiamo scendere in un ambito più anatomico, soffermandoci sui **fotorecettori**. Si tratta di neuroni in grado di tradurre la luce assorbita in impulsi elettrici che vengono poi inviati al cervello. Ci sono due tipi di fotorecettori.

- **coni**, che servono generalmente a distinguere i colori, utilizzando una tecnica di **tricromia**. L'occhio percepisce soltanto il verde, il rosso e il blu. Una combinazione lineare tra questi tre colori permette di distinguere tutti i colori della realtà. Questo processo è chiamato **sintesi additiva**.
- **bastoncelli**, che servono più che altro a distinguere e percepire i movimenti nel nostro campo visivo. Essi, inoltre, sono molto più sensibili alla luce e permettono di vedere in condizioni di scarsa visibilità.

Tornando a parlare in un ambito prettamente più informatico, diciamo che la luce, normalmente, è un'energia elettromagnetica con una lunghezza d'onda compresa tra 380 e 760 nanometri, chiamata **spettro visibile**. Il **colore**, in CG, è estremamente importante, essendo una componente fondamentale della realtà. Così, sono stati inventati dei **modelli di colore**. Si tratta fondamentalmente di modelli matematici che permettono di rappresentare i colori in forma numerica, cioè con la sottrazione o l'addizione di determinati componenti. Ne citiamo alcuni:

- **RGB** (*Red Green Blue*). Si tratta di un modello additivo che somma tre colori, appunto il rosso, il verde e il blu.
- **HSB** (*Hue Saturation Brightness*). Si tratta di un altro modello additivo, dove però vengono sommati non colori veri e propri ma alcune proprietà, appunto tonalità, saturazione e luminosità.

- **CMYK** (*Cyan Magenta Yellow Black*). Si tratta di un modello sottrattivo che mischia, appunto, ciano, che sarebbe un azzurro opaco, magenta, giallo e nero.
- **YUV** ($Y'UV$). Si tratta di un modello utilizzato nelle trasmissioni televisive ed è concepito per rispecchiare il comportamento della visione umana.

Parlando di modelli di colore, è giusto anche parlare del **gamut**. Il gamut è l'intera gamma di colori che può essere prodotta da un determinato modello. Ovviamente, un modello di colore non è in grado di riprodurre l'intera gamma di colori percepibili all'occhio umano. Definendo uno spazio-colore, chiamandolo **LAB**, dove L indica la luminosità e a e b indicano le coppie colore-opponente (rosso-verde, blu-giallo) e che, quindi, copre tutti i colori dello spettro visibile, possiamo dire che $LAB > RGB > CMYK$. Ossia RGB è in grado di coprire circa il 70%, mentre CMYK circa il 50%. Di solito, quando un colore è fuori dalla gamma di un modello, è sostituito con un altro che si trova, però, all'interno di esso. Possiamo anche combinare più gamut di più modelli di colore. Ovviamente, quando parliamo di anatomia umana e luce, ci troviamo di fronte a molte problematiche. Diverse sono le immagini sul web di illusioni ottiche e percezioni alquanto strane. Citiamo per esempio quel vestito da donna che alcune persone lo vedono blu e nero e altre oro e bianco. Non è magia, infatti è sempre il nostro cervello che decide ed interpreta quanta luce debba essere assorbita dalla retina. Questo a meno che non ci siano patologie indicate, come può essere il daltonismo, una percezione distorta dei colori, o l'acromatopsia, la totale incapacità di percepire i colori.



STEREOSCOPIA

Quando parliamo di stereoscopia ci riferiamo ad una tecnica di realizzazione e visione di immagini, che serve a dare un'**illusione di tridimensionalità**, analoga a quella generata dalla visione binoculare dell'occhio umano. Prima di entrare nel merito, scendiamo nel dettaglio definendo alcuni meccanismi anatomici che ci permettono di avere una visione stereoscopica della realtà.

- **accomodazione.** Si tratta di un meccanismo che regola la lunghezza focale degli occhi per mettere, appunto, a fuoco qualcosa. Anatomicamente si basa sul cambiamento dello spessore del cristallino, causato da tensione e rilassamento dei muscoli ciliari.
- **convergenza.** Si tratta di un meccanismo in grado di cambiare l'angolazione degli occhi, in modo da far convergere gli oggetti vicini e parallelizzare quelli lontani.

Di solito, accomodazione e convergenza lavorano insieme, senza avere una particolare corrispondenza fisiologica, ma basandosi solo e soltanto sull'esperienza. Un altro concetto molto basilare della stereoscopia è la **disparità binoculare**. L'occhio sinistro non vede la stessa immagine dell'occhio destro. E poi c'è anche il concetto di **parallasse**, vale a dire l'angolo che si forma su un oggetto quando viene osservato da due punti di vista differenti. Ed è su questo che si sono basati gli studiosi durante la storia per poi definire la stereoscopia moderna. Contando solo sulla vista, ci sono due tecniche ben precise per vedere in stereoscopia un'immagine in 2D:

- **stereoscopia incrociata:** l'occhio destro guarda l'immagine sinistra e viceversa. Per provarla è sufficiente avere uno stereogramma ed osservarlo incrociando gli occhi. Così facendo, le immagini prima appariranno sdoppiate e poi si sovrappongono perfettamente.
- **stereoscopia parallela:** l'occhio destro guarda l'immagine destra e il sinistro quella sinistra. Si tratta di una stereoscopia molto difficile da dimostrare,

poiché non è consona alla posizione naturale degli occhi. Per vedere due immagini in parallelo è necessario avvicinarsi molto e poi muoversi lentamente indietro.

Al giorno d'oggi possiamo distinguere due tipi di stereoscopia.

- **stereoscopia passiva.** Due immagini vengono proiettate contemporaneamente sullo stesso schermo e un ipotetico osservatore indossa degli occhiali in modo che su ogni lente venga impressa quell'immagine che effettivamente gli compete. A seconda delle tecniche di stereoscopia passiva, si distinguono diversi tipi di occhiali. Tra le tecniche ci sono. **Anaglifo:** utilizza un'immagine che viene ottenuta tramite la sovrapposizione di due fotogrammi in uno stereogramma, ognuno che subisce un profilo colorifico diverso. Rosso per l'immagine sinistra, blu per quella destra. Si tratta di una tecnica molto economica, ma che non garantisce una perfetta visualizzazione cromatica dell'immagine. **Polarizzazione:** utilizza un doppio proiettore le cui lenti sono dotate di filtri polarizzatori, orientati ad angolo retto, in modo da formare due immagini differenti una dall'altra. Anche gli occhiali sono dotati di due lenti polarizzate. Pure questa è molto economica, ma limita i movimenti della testa ed è utile soltanto all'interno di ambienti in cui non c'è bisogno di muovere la testa (cinema). **Gamma di colore:** si usa sempre un proiettore, al cui interno c'è una ruota di colori alternata, rosso, blu, verde e giallo. Con questa tecnica si riesce a garantire sicuramente più libertà di movimento ed è possibile usare qualsiasi superficie per la proiezione. I costi, però, sono molto elevati.
- **stereoscopia attiva.** Due immagini vengono proiettate in maniera alternata su un medesimo schermo. Si utilizzano particolari tipi di occhiali, detti *shutter glasses*, che sono sincronizzati con il proiettore. Si ha un effetto stereoscopico molto realistico, con gli sfarfallamenti ridotti al minimo. I costi, però, sono molto elevati e gli occhiali molto delicati e poco durevoli. Inoltre, non è possibile usare tecniche di stereoscopia attiva quando si ha di fronte un grande numero di osservatori.
- **autostereoscopia.** A differenza delle due tecniche stereoscopiche appena definite, qui l'immagine tridimensionale può essere visualizzata senza l'ausilio di altri dispositivi (proiettore, occhiali), poiché all'interno del supporto ci sono già dei meccanismi che nascono ad un occhio l'immagine destinata ad un altro. Tra le tecniche autostereoscopiche abbiamo. **Barriera di parallasse:** un

filtro, una barriera, distribuisce i punti di vista agli occhi in maniera alternata.

Rete lenticolare: questa alternanza di punti di vista è ottenuta attraverso una rete di microlenti, collocata direttamente sulla superficie delle immagini.

Olografia: si usa un HOE (*Holographical Optical Element*), che è posto davanti allo schermo di visualizzazione. **Autostereogramma:** si usa uno stereogramma a singola immagine, che provoca un'illusione ottica tridimensionale partendo da un'immagine bidimensionale



REALTÀ AUMENTATA

All'inizio del corso abbiamo parlato di realtà virtuale. Con questo termine ci riferiamo ad una realtà simulata, dove però non si è riusciti a raggiungere un alto livello di realismo. La realtà aumentata, invece, propone delle interfacce basate sulla realtà. Si basa sul posizionamento di oggetti virtuali in un ambiente reale, occupandosi del rendering real-time e del loro allineamento. La realtà aumentata è usata in molti ambiti, sia quello industriale che quello medico, turistico e militare. Normalmente, abbiamo tre tecniche principali di AR:

- **visori video see-through.** Sono i dispositivi più vicini alla VR, poiché vengono usate due telecamere con le quali si acquisisce l'immagine reale e poi viene fusa con quella di sintesi, prima di essere inviata all'occhio tramite un display. Sono dispositivi molto precisi, ma richiedono sistemi di messa a fuoco automatici, i cui tempi di risposta sono molto lunghi.
- **visori optical see-through.** Si vede il mondo reale così com'è, con l'aggiunta di oggetti virtuali aumentati attraverso l'impressione su lenti trasparenti. La qualità di questi oggetti, però, non è eccelsa.
- **display di proiezione.** Usano sistemi di proiezione per illuminare oggetti reali con immagini generate a computer. Sono dispositivi molto sensibili alla luce e limitano le informazioni visualizzate.

Per far funzionare al meglio la AR è necessario che, tra immagini reali e contenuti virtuali, ci sia un rapporto di **co-registrazione**, cioè che siano sincronizzati tra di loro. Uno dei problemi principali resta il calcolo del movimento, partendo dal punto di vista dell'utente. Di solito viene realizzato da sistemi di **Motion Tracking**. In base alla tecnologia di misurazione ci sono diversi sistemi di MT, che però prendono in considerazione alcuni parametri di riferimento:

- **volume lavoro**, lo spazio in cui il sistema deve funzionare.
- **frequenza di campionamento**, la frequenza con cui le variabili vengono aggiornate.

- **risoluzione**, variazioni di posizione.
- **latenza**, rapidità della risposta.
- **precisione**, entità di un eventuale errore.

Soffermiamoci, adesso, sui principali sistemi di MT:

- **sistemi meccanici**. Utilizzano potenziometri o encoder ottici per misurare la rotazione dei perni di vincolo delle aste di collegamento, calcolando la posizione degli oggetti. Si tratta di un sistema semplice ed economico, che però limita i movimenti.
- **sistemi elettromagnetici**. Utilizzano un trasmettitore e un ricevitore. Il trasmettitore genera un campo magnetico, che viene rilevato dal ricevitore attraverso delle spire. Sono dispositivi comodi, ma sensibili alle interferenze elettromagnetiche.
- **sistemi ottici**. Utilizzano sorgenti luminose disposte sull'oggetto da tracciare. Si tratta di un sistema molto preciso, ma estremamente costoso.
- **sistemi acustici**. Utilizzano un emettitore che genera un segnale sonoro e un microfono che lo raccoglie, calcolando anche il tempo trascorso. Sono dispositivi molto economici, ma molto sensibili a delle influenze esterne.
- **sistemi inerziali**. Utilizzano giroscopi per rilevare i cambiamenti di rotazioni attorno ad uno o più assi.

Individuare un corretto sistema di MT è molto complicato e i sistemi descritti fino ad ora sono tutti facilmente influenzabili da fattori esterni. Invece, utilizzando un approccio **image-based**, si cambia radicalmente il concetto di MT. Il tracciamento non è più collegato a complicati sistemi di trasmissione o ricezione, ma si basa sul riconoscimento di figure dette **marker**. Da essi si riceve un punto di riferimento nella scena in cui inserire contenuti aumentati. **ARToolkit** è una libreria in C/C++, open-source, orientata allo sviluppo di realtà aumentata marker-based. Un marker è fondamentalmente una figura stampata su carta, che viene riconosciuta dal software. Una volta individuato il marker, il software è in grado perfettamente di orientarsi nella scena. Ovviamente, l'approccio marker-based ha alcune limitazioni, infatti il marker deve per forza essere inquadrato dalla camera e questo richiede una focale molto ampia. Possiamo introdurre anche l'approccio **multimarker**, in cui il tracciamento avviene dall'individuazione di una costellazione di marker e non di

uno soltanto. Ovviamente, restano le limitazioni dell'approccio marker-based normale. Per questo è possibile anche utilizzare una tecnologia **markerless**, cioè senza marker, che può essere **model-based**, cioè il riconoscimento tramite modelli tridimensionali, contorni e proiezioni, o **feature-based**, cioè l'individuazione delle caratteristiche salienti degli oggetti.



INTERAZIONE APTICA

Dopo aver parlato di VR e AR, unito alla visione stereoscopica, manca l'ultimo tassello, ossia l'utilizzo di interfacce aptiche, che aggiungono percezioni tattili al mondo virtuale. L'**aptica** è la scienza che studia il tatto e le interazioni con l'ambiente, tramite esso. Il tatto è una perturbazione meccanica della pelle, prodotta dal contatto fisico con un oggetto, stabilendo un canale bidirezionale che permette di avere informazioni sulla forma, la dimensione, la texture, la temperatura ecc. Il tatto fornisce informazioni al cervello tramite due livelli:

- **livello cinestetico**, che si riferisce al senso di posizione e movimento di parti del corpo, in associazione alle forze esercitate da un oggetto. Le informazioni vengono prodotte da alcuni recettori nei muscoli e nei tendini. I sistemi cinestetici più diffusi sono quello basati sull'interazione con parti del corpo, come *finger-based*, *point-based*, *hand-based* ed *exoskeleton*.
- **livello tattile**, che si riferisce alla cute e permette di ottenere le informazioni sul tipo di contatto e sulle proprietà fisiche dell'oggetto. Le informazioni vengono prodotte da alcuni recettori sulla pelle. I sistemi tattili più diffusi sono generalmente dispositivi termici o vibrotattili.

Unendo informazioni cinestetiche e tattili, definiamo informazioni aptiche, che a loro volta possono utilizzare sistemi aptici come *matrix-display* e altri dispositivi basati su sensori.

Un'**interfaccia aptica (HI)** è un dispositivo elettromagnetico, il cui end-effector collega l'arto dell'utente (dotato di altri dispositivi sensoriali) e un sistema di attuatori. I movimenti sono rilevati dai sensori sul braccio, mentre alcune forze di feedback sono applicate dagli attuatori, generalmente motori elettrici. Esse servono ad orientare l'utente sulla posizione e sulla natura degli oggetti con cui interagisce sullo spazio virtuale. Le interfacce aptiche sono **multimodali**, perché cooperano con altre interfacce, generalmente visive. Possiamo dividerle prima in due categorie, cioè le **interfacce desktop**, utilizzabili in base alla lunghezza del dispositivo, o **interfacce mobili**, con molti più gradi di libertà e talvolta anche indossabili. Poi, invece, individuiamo le **interfacce a impedenza**, quelle che

generano una forza in base alla posizione e alla velocità del nostro arto, o quelle ad **ammettenza**, che generano una forza in base alla forza percepita nel raggiungere una posizione o una velocità.

Un'interfaccia aptica ha alcuni parametri fondamentali:

- **numero di gradi di libertà**, cioè la possibilità di misurare movimenti lineari e spaziali, rotazione ecc.
- **dimensione**, cioè più è grande, più è complicata da indossare.
- **trasparenza**, cioè nascondere tutte le funzioni secondarie per far percepire all'utente soltanto le collisioni.

Quando parliamo di **rendering aptico (HR)**, ci riferiamo all'operazione con la quale forniamo la forza all'utente per interagire con un oggetto virtuale, gestendo anche l'aggiornamento delle informazioni aptiche (finalità HR). L'HR ha, però, varie problematiche. Risulta difficile replicare perfettamente le sensazioni tattili, poiché il tatto è un senso molto sensibile. Per fare un esempio, il tatto aggiorna una sensazione circa mille volte al secondo, in modo da fornire un'esperienza tattile convincente. Invece, in un ambiente virtuale, una scena 3D viene aggiornata venticinque o trenta volte al secondo.