

Lezione 4 13/03/2024

Tracking

Il **tracking** è l'operazione dove tramite un device, andiamo ad acquisire in maniera continuativa delle informazioni su un oggetto (mani, testa, sguardo...).



Il movimento di un oggetto rigido (non deformabile) può essere diviso in un displacement nello spazio e una rotazione su 3 assi perpendicolari. Vengono utilizzati 6 **degree of freedom (DOF)**, che 3 coordinate di posizione e 3 angolazioni che ne descrivono l'orientamento.

Con **3-DoF** posso fare per esempio la navigazione dei video a 360° (rotazione della testa). Per ottenere questo tipo di informazioni sono necessari dei sensori abbastanza semplici come accelerometri e giroscopi che sono già presenti su device come smartphone (o direttamente sugli headset)

Sui device più avanzati con **6-DoF** posso ottenere sia la rotazione che la **traslazione** (spostamento del corpo). Possiamo quindi sapere anche se l'utente si è mosso in avanti, indietro, lateralmente o verticalmente.

In AR c'è sempre il problema di come vado a registrare l'oggetto, ovvero come vado a posizionare l'oggetto virtuale nel mondo reale. Quindi avere i 6-DoF è meglio perché si possono tracciare i movimenti dell'utente meglio. In AR è anche necessario tracciare oggetti del mondo reale, non solo l'utente.

In AR con tracking si intende generalmente quello 3D.

Calibrazione

La **calibrazione** è un processo meccanico/fisico dove andiamo a confrontare le misure che arrivano da due device diversi, dove uno è quello di riferimento, andando quindi ad aggiustare il secondo device rispetto a quello di riferimento. Nell'AR e nel VR abbiamo bisogno di calibrare i componenti del sistema, specialmente quelli usati per il tracking.

La **registrazione** (registration) in AR si riferisce all'allineamento dei sistemi di coordinate tra gli oggetti reali e quelli virtuali.

Nello specifico, con i see-through display viene fatta la registrazione.

Questa è la

registrazione statica (static registration), fatta quando l'utente e la videocamera sono fermi. Serve per stabilire un sistema di coordinate comune tra il sistema virtuale e gli oggetti reali.

Invece la

registrazione dinamica è lo step successivo, quando l'utente si muove e quindi gli oggetti virtuali mantengono (tramite il tracking) la posizione corretta nel mondo reale.

La **calibrazione** è quindi fatta da fermi, fatta di solito una volta, o quando viene riavviato il sistema. A questo punto si può fare la **registration** (allineamento di proprietà spaziali), e infine tramite il tracking si fa la **dynamic registration**, mantenendo gli oggetti virtuali nella posizione tridimensionale corretta anche con il movimento.

La Calibrazione è

A. Un confronto fra due misure ottenute da due device

La Registrazione è

B. l'allineamento di due misurazioni provenienti da sistemi di coordinate

Il Tracking è

C. la misurazione dinamica di proprietà spaziali di un oggetto

Spatial Display Model

Ho un elemento di cui voglio fare il tracking, il software deve capire nel tempo dove deve essere piazzato, deve fare una serie di calcoli in cui trasforma, allinea e registra le coordinate.

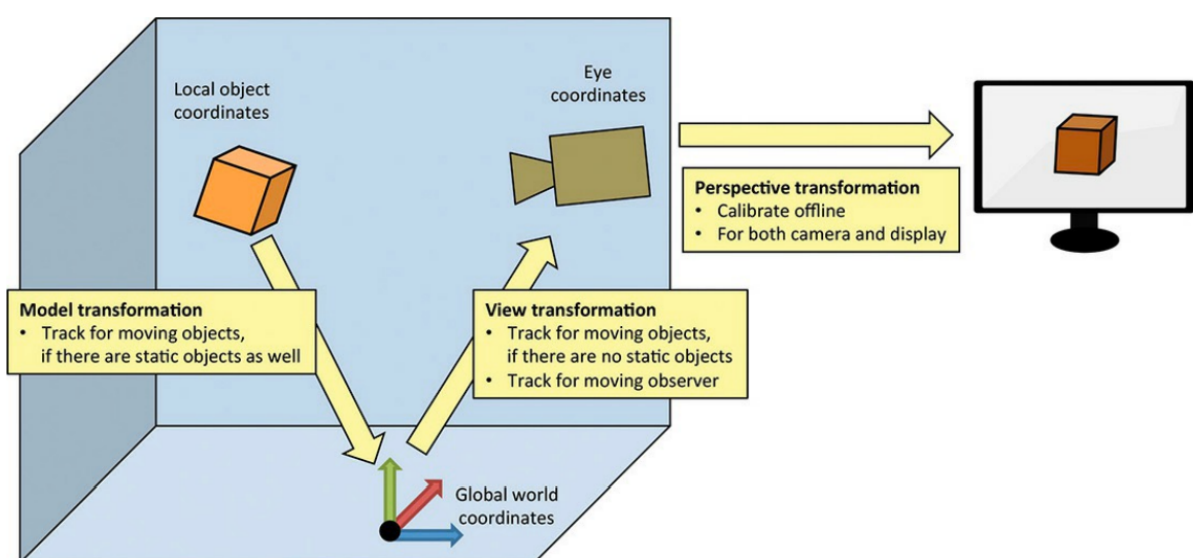
Per farlo si usano 3 modelli: la trasformazione del modello, la trasformazione della view e la trasformazione della prospettiva/proiezione.

La **model transformation** va a descrivere dove un oggetto, dato un ambiente statico, si sta andando a posizionare. Quindi come cambia il modello, quando ho un oggetto che si sta spostando. Quando parlo di model transformation mi sto occupando della relazione tra le coordinate locali di un oggetto in relazione con le coordinate globali del mondo.

La **view transformation** si occupa della camera o del sensore che fa il tracking, di come si sta muovendo. La view transformation invece prende le coordinate del mondo 3D e le coordinate dell'observer (il soggetto, la camera), allineando come la camera si sta posizionando.

La **perspective/projective transformation** gestisce le coordinate dell'occhio e quelle dello schermo. Prende le coordinate della view dello schermo e cerca di allinearle le coordinate con quello che sto tracciando e con il movimento dell'utente, facendo il rendering.

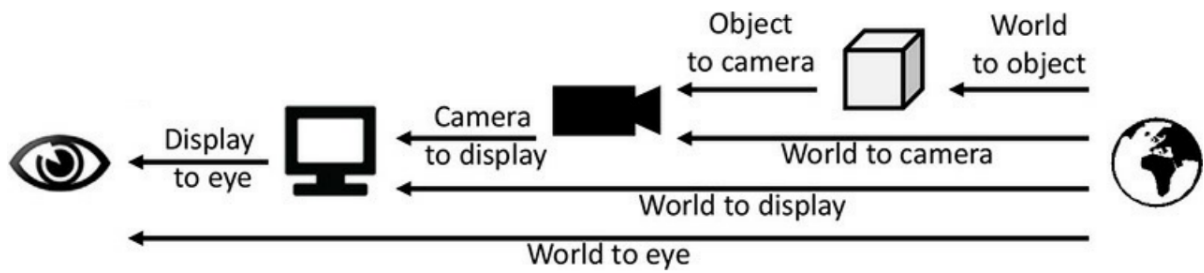
Alcuni parametri possono essere calibrati offline, mentre altri cambiano e quindi necessitano di tracking.



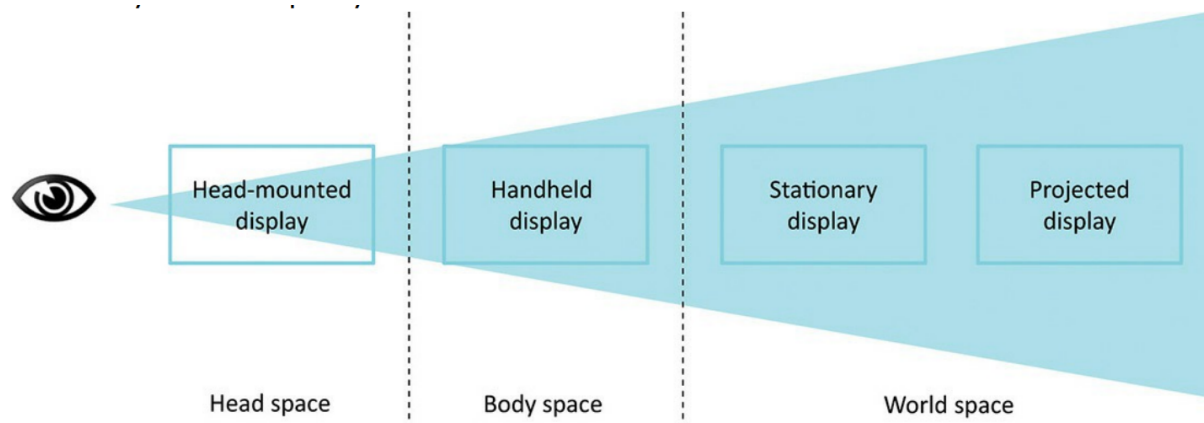
Si può parlare di sistemi di **coordinate locali**, ovvero posizionando lo 0,0 della stanza dove sono per esempio, sono coordinate su una scala più limitata (oppure la posizione dell'utente relativamente ad un angolo della stanza), mentre il sistema di **coordinate globale** è un sistema assoluto, quindi per esempio le coordinate di latitudine e longitudine precise.

Calcolare la **view transformation** può essere complesso, bisogna considerare molti componenti come il tracking della testa, degli occhi, del display (movimento), e se un sistema VST è usato anche il tracking della camera.

Il **modello spaziale** dei display AR può essere definito come la relazione tra fino a 5 componenti: gli occhi, il display, la camera (o il sensore di tracking), l'oggetto virtuale e il mondo reale.

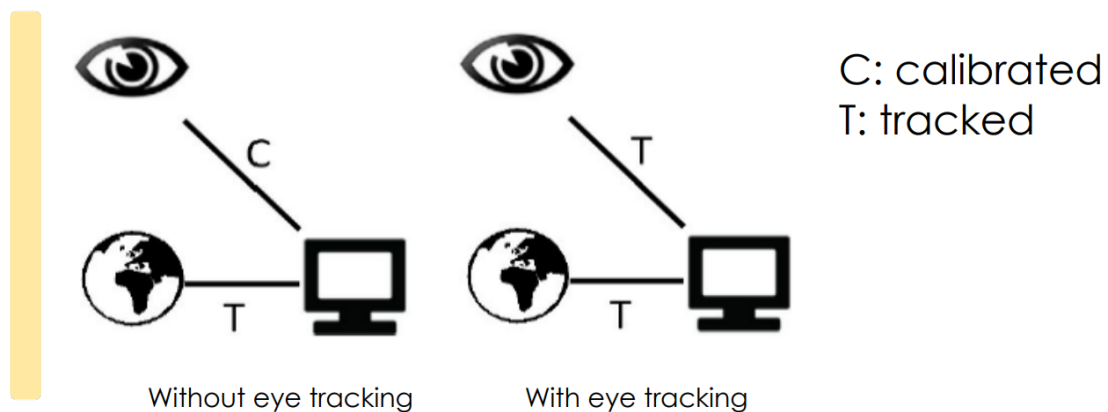


A seconda della distanza del display dall'occhio, i **display AR** vengono **classificati** in questo modo:



OST HMD

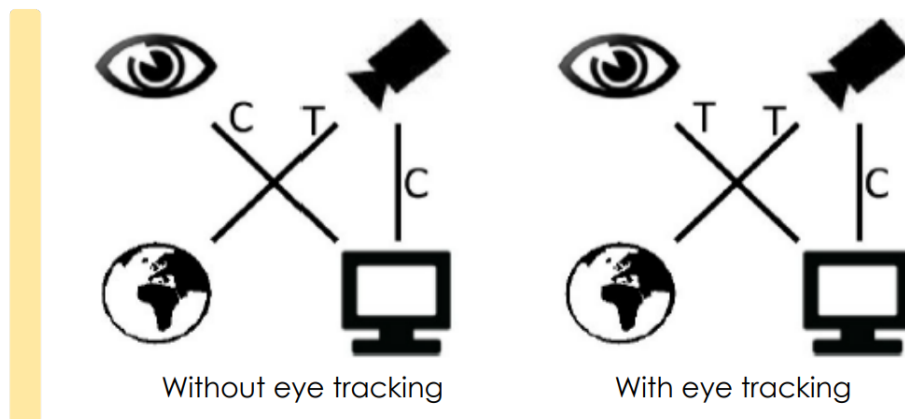
Su questo device lavoro con gli occhi, il display e il mondo.



Nel caso dove c'è la calibrazione, questa poi è statica. Eventualmente poi si può fare anche il tracking dell'occhio, oltre a quello del mondo esterno.

VST HMD

In questo caso in più ho anche la camera.



Qui calibro l'occhio con il display, ma anche le camere con il display. A quel punto la camera va a fare il tracking del mondo reale. Eventualmente posso fare il tracking anche degli occhi.

La view deve essere il quanto più possibile allineata con quella reale dell'utente.

Handheld display

In questo caso viene usato un oggetto come un tablet o uno smartphone che viene tenuto in mano, contiene sia il display che la videocamera.

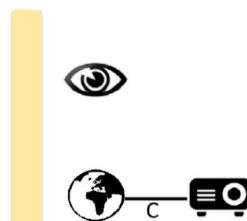
In questi device il tracking viene fatto sfruttando la telecamera stessa del device, ma si possono usare anche altri metodi.

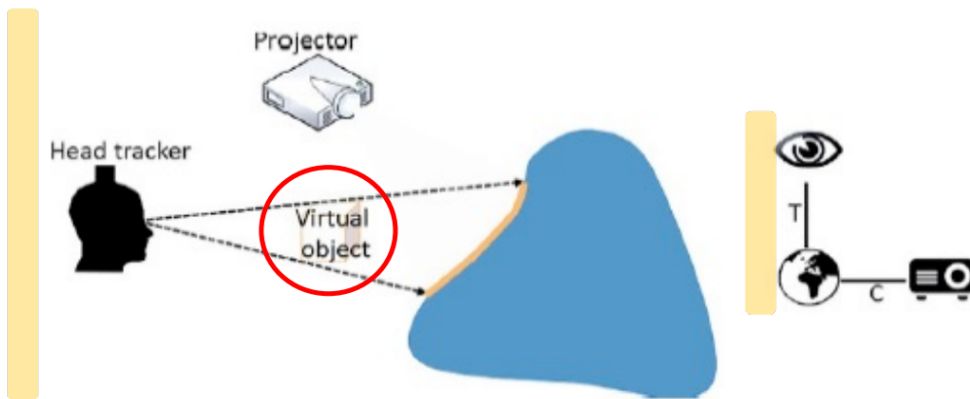


Projected display

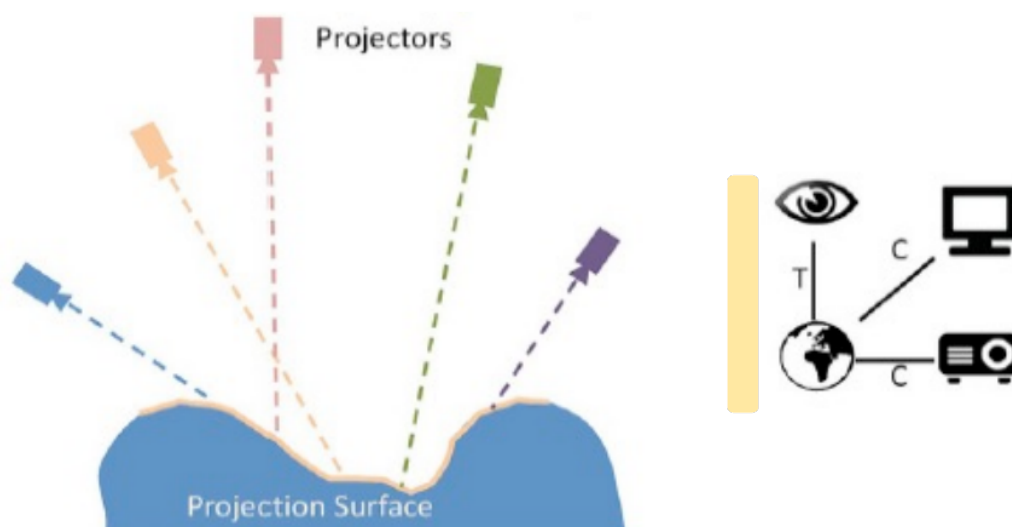
Qui in più abbiamo anche i proiettori, che vanno calibrati con il mondo fisico, mentre non vengono calibrati gli occhi.

Mentre invece **con proiettori più moderni dove vengono fatte proiezioni 3D, è necessario anche fare il tracking del punto di vista dell'utente.**





Infine, se abbiamo una superficie più complessa con molti proiettori che proiettano immagini che si possono sovrapporre, questo viene generalizzato e visto come un display. Va quindi calibrato il display con il mondo, ovvero quella parte di superficie dove viene proiettato, oltre alla calibrazione dei singoli proiettori vengono quindi anche calibrati insieme.



Come fare il tracking

Per fare tracking dobbiamo avere un sensore che va a percepire qualche tipo di segnale. Stiamo misurando qualche fenomeno fisico. Concettualmente possiamo usare molte cose diverse:

- **fenomeni fisici:** radiazione elettromagnetica (luce, infrarossi, laser, segnali radio), suoni, gravità, inerzia, con sensori specializzati.
- **Principio di misurazione:** sono misurati tramite la forza del segnale, la direzione, il time of flight... viene effettuata attraverso **triangolazione** (determina la posizione da due o più angoli misurati, assumendo almeno

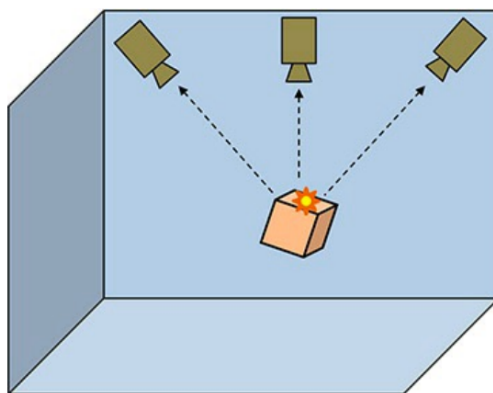
una distanza conosciuta) o **trilaterazione** (è un metodo geometrico per determinare la posizione da almeno 3 distanze misurate).

Un approccio comune è quello di usare multipli sensori di input contemporaneamente, sincronizzandoli (per assicurarci dell'acquisizione allo stesso momento) e **fondendoli** (combinazione di più sensori per avere una misura più accurata).

A seconda del tipo di sensore, **la fonte** può essere:

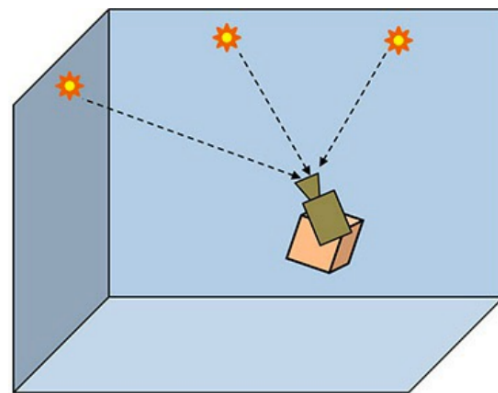
- **sorgenti passive**: quelle sorgenti che emettono il segnale in maniera naturale, non sono create ad hoc: per esempio il campo magnetico della terra oppure il sole. In questo caso si parla di "**sourceless sensing**"
- **sorgenti attive**: sono le sorgenti create da qualcosa di elettronico prodotto dall'uomo, messo apposta oppure anche per un altro motivo e sfruttata. Questi oggetti producono segnali acustici, ottici, onde radio, etc. Tendenzialmente soffrono di interferenza perché se tra l'emettitore e il sensore ci sono ostacoli il segnale potrebbe essere distorto.
 - **sorgenti attive dirette**: sono quelle direttamente montate sull'oggetto di cui voglio fare tracking (un oggetto esterno tipo una bottiglietta)
 - **sorgenti attive indirette**: l'oggetto deve essere in grado di riflettere il segnale dalla fonte da cui arriva

Direct Active Source vs



Outside-in tracking: sensors are mounted stationary in the environment and observe a moving target, such as a head-worn display.

Indirect Active Source



Inside-out tracking: sensors move with the tracked object and observe stationary references in the environment.

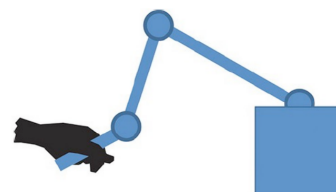
Tipi di tracking

- **Stazionaria:**
 - Tracking meccanico
 - Tracking elettromagnetico
 - Tracking con ultrasuoni
 - (non molto popolare nell'AR oggi)
- **Mobile**
 - Sensori di inerzia
 - GPS
 - Wireless
- **Optical** tracking

Mechanical Tracking

Abbiamo strumenti che fanno sia input che output (di solito). Per esempio un braccio meccanico mosso dall'utente, dove il movimento del braccio meccanico viene misurato e tracciato. Le angolazioni dei punti di giuntura sono tracciate da sensori.

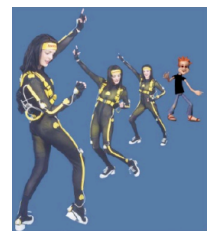
Questi sono estremamente precisi, possono avere anche feedback di forza per esempio. Però sono molto limitati dalla struttura meccanica. In AR è quindi complicato perché il braccio rimane nell'FOV e questo è indesiderabile perché si trova nella stessa zona dove vengono inseriti gli oggetti virtuali. Inoltre sono costosi.





Electromagnetic Tracking

Per fare tracking elettromagnetico devo avere delle sorgenti stazionarie che producono 3 campi magnetici ortogonali. A quel punto vado a misurare dei sensori (tramite la forza e la direzione del campo magnetico, usando piccoli sensori che hanno 3 coils ortogonali, si calcola la differenza tra il trasmettitore e il ricevitore), riuscendo a capire la posizione e la rotazione dell'oggetto dove ci sono i sensori.



Questi sensori possono avere problemi di interferenza con oggetti magnetici o metallici (però non necessita di essere nel campo visivo), e funzionano a 1-3 metri di distanza, con 6 DOF.

Questo tipo di tracking è usato per esempio tramite delle tute che tracciano i movimenti della persona. Necessitano di alcuni cavi che vanno in uno zaino e sono costosi. Approcci magnetici sono poco utilizzati.



Nella seconda immagine l'oggetto centrale crea il campo magnetico, che viene usato per trovare la posizione degli altri due oggetti.

Tracking Acustico

Vengono usati gli ultrasuoni in quanto non sono udibili dalle persone. Utilizzano le differenze nel time of flight (TOF) o nella fase delle onde sonore.

Il trasmettitore o il ricevitore possono essere messe sull'oggetto da tracciare, anche qui le distanze sono basse. Si possono ottenere sia 3 (un trasmettitore e tre ricevitori) che 6 DOF (3 trasmettitori e 3 ricevitori).

La posizione dell'oggetto da tracciare è limitata ad una superficie sferica intorno al trasmettitore.

Sono economici, ma soffrono di problemi per colpa dei cambi di temperatura e pressione (vanno ricalibrati) e quindi all'esterno sono poco pratici.

Sensori mobile

A differenza dei tracking systems stazionari che richiedono che l'utente non si muova molto, questi sono mobili.

Sono emettitori e sensori che di solito si trovano sugli smartphone. Dato che l'utente non può fare affidamento a servizi wireless sempre di qualità, la computazione per il tracking deve essere fatta localmente sul dispositivo, senza l'aiuto di infrastrutture nell'ambiente.

- **compasso inerziale:** fornisce una misura assoluta dell'orientamento rispetto al nord, tramite il campo magnetico terrestre.
- **sensori d'inerzia:** accelerometri e giroscopi, misurano l'accelerazione sui tre assi, e sono in grado di misurare la rotazione, l'orientamento (con il giroscopio, che misura la velocità di rotazione). Possono quindi anche calcolare il movimento a livello di posizione. Sono wireless, poco costosi e piccoli, danno informazioni con un'alta frequenza. Subiscono di drift (deriva) nel tempo.
- **gps:** utilizza i satelliti, dato che questi hanno una posizione precisa e conosciuta, se il ricevitore può agganciare almeno 4 satelliti diversi (misurando la differenza nel time of flight), allora può triangolare la sua posizione (in base alla distanza da ciascun satellite, con il time of flight). L'elevazione e le misurazioni indoor non sono attendibili. Non misura l'orientamento.

Associate le definizioni ai termini relative al tracking

La Calibrazione è

① ✓ 91% A

Un confronto fra due misure ottenute da due device

La Registrazione è

② ✓ 82% B

l'allineamento di due misurazioni provenienti da sistemi di coordinate differenti

Il Tracking è

③ ✓ 88% C

la misurazione dinamica di proprietà spaziali di un oggetto