

Telepresenza “intelligente”: la realtà virtuale nei robot avanzati

Riccardo Cassinis[°], Mardy Terceros Rojas^{°°}

Sommario

L'articolo discute le esigenze di un sistema di telepresenza utilizzabile, fra l'altro, per il controllo a distanza di robot operanti in ambienti inaccessibili per l'uomo. Particolare enfasi viene posta al problema dell'effetto presenza, inteso come la capacità di creare, attraverso opportune serie di attuatori e di sensori, la sensazione per l'operatore di trovarsi realmente immerso nell'ambiente in cui opera il robot, dandogli così la possibilità di interagire con quest'ultimo in maniera quanto più possibile efficace e realistica.

Vengono presentate alcune ipotesi di soluzione, unitamente ad alcuni risultati sperimentali già ottenuti.

1. - Introduzione

Il termine realtà virtuale si riferisce alla generazione realistica di un mondo tridimensionale nel quale l'operatore, dotato di mezzi opportuni, possa interagire con gli oggetti virtuali generati dal calcolatore come se si trovasse in un mondo reale, movimentando, afferrando e vedendo da diverse angolazioni.

Esiste tuttavia una certa confusione terminologica, che richiede alcune precisazioni. Anche se il termine “realtà virtuale” viene comunemente usato per indicare tutti quei sistemi in cui l'uomo percepisce informazioni sensoriali che non provengono direttamente dall'ambiente in cui è immerso, si danno due casi ben distinti: il primo in cui il mondo è completamente creato dal calcolatore, e non esiste quindi nella realtà (realtà virtuale vera e propria); il secondo in cui il mondo che l'operatore vede e in cui ha

[°] Dipartimento di Elettronica per l'Automazione, Università di Brescia.

^{°°} Progetto di Intelligenza Artificiale e Robotica, Dipartimento di Elettronica e Informazione, Politecnico di Milano.

l'impressione di muoversi e di agire esiste realmente, ma è in un luogo diverso da quello in cui si trova l'operatore stesso. Questo secondo caso viene più correttamente chiamato *telepresenza* [4]. Questo articolo tratta del secondo caso, perché è quello che riguarda più da vicino la telemanipolazione: si mostrerà comunque che, per ottenere un buon effetto di telepresenza, è necessario ricorrere anche a tecniche di creazione di mondi sintetici.

La telepresenza può avere un vasto campo di applicazioni industriali e non, per esempio nel campo della manipolazione di materiali pericolosi e/o in ambienti ostili; un altro interessante campo di applicazioni può essere quello dell'ausilio ai disabili: il disabile può pilotare un robot e può muoversi in un ambiente dove avverte l'illusione di essere realmente presente.

2. - Lo stato dell'arte

I concetti di telepresenza e di realtà virtuale sono entrati già da parecchio tempo nella letteratura riguardante la robotica, soprattutto per quanto riguarda la telemanipolazione. Gli esempi di applicazione della telepresenza alla robotica sono molteplici: saranno citati qui solo quelli ritenuti più significativi [1].

L'origine della telepresenza può essere fatta risalire al 1965, quando Sutherland propose *The Ultimate Display*, un visore in cui l'immagine generata dal calcolatore dovrebbe apparire come se fosse reale [2]. Sutherland riconobbe la necessità di un ingresso sensoriale il più completo possibile, mettendo in evidenza che la caratteristica più importante è il feedback cinetico, cioè l'interazione tra movimento dell'operatore e visualizzazione dell'ambiente reale e/o virtuale sui visori in modo tale da rendere l'illusione di un'osservazione diretta. La base del sistema di Sutherland è l'HMD (Head-Mounted Display), un sistema di visualizzazione che utilizza due tubi a raggi catodici solidali con la testa dell'operatore.

Nel 1983 Callahan dell'Architecture Machine Group (M.I.T.) ha costruito una versione perfezionata del visore di Sutherland. Questa macchina adopera l'inseguitore magnetico *Polhemus Navigation Sciences* dell'MIT, montato sull'HMD per rilevare la posizione e l'orientamento della testa dell'operatore. Altra innovazione è l'uso di un disco ottico per ottenere una buona qualità dell'immagine e una risposta veloce ai movimenti della testa.

Fino dal 1985 McGreevy e Fisher della NASA Ames Research Center adoperano le lenti LEEP (Large Expanse Extra Perspective) che permettono di espandere il campo visivo dell'HMD.

Nel campo della telemanipolazione è stato proposto dall'IVA (Intravehicular Activities) un sistema complesso di stazione spaziale [3]: l'operatore indossa un casco che riproduce visivamente una scena del luogo remoto, simulando la presenza della sua testa; le sue braccia sono inserite in strutture che trasmettono segnali a braccia artificiali situate nel luogo remoto, così che queste ultime possono assumere lo stesso orientamento delle braccia dell'operatore. Le mani dell'operatore sono dotate di guanti speciali con rilevatori di posizione, che permettono di afferrare oggetti nel campo di lavoro, restituendo nel contempo segnali tattili per creare l'illusione di toccare realmente gli oggetti.

I laboratori Ames della NASA hanno messo a punto un casco virtuale e interattivo per utenti multipli che permette il monitoraggio reciproco dei propri movimenti e l'esplorazione in prima persona di ambienti virtuali che simulano il luogo remoto.

In questa realizzazione l'operatore è dotato di casco che consente la visione stereoscopica su display, di guanti interattivi, e di una "tuta" che permette all'intero corpo di esplorare ed interagire con gli oggetti virtuali generati dal calcolatore.

Sotto la direzione del Dott. Thomas Furness, è stato sviluppato al Armstrong Aerospace Medical Research Laboratory della Wright-Patterson Air Force Base, un HMD sperimentale. L'uso di piccolissimi CRT e di un sistema ottico presenta al pilota una scena tridimensionale con un'apertura di 120°. Esso inoltre permette una comunicazione vocale e gestuale con il sistema di elaborazione, ed è dotato di un sistema tridimensionale di riproduzione del suono.

Il VERDEX (Virtual Environment Remote Driving EXperiment) del U.K. Advanced Robotics Research Centre's è un sistema di interfaccia che permette interazioni intuitive fra uomo e robot [4]. La prima fase del progetto è stata dedicata allo sviluppo di un "banco di prova" dell'abilità dell'operatore. In un secondo tempo, l'attenzione è stata rivolta allo sviluppo di un sistema integrato video-audio e di un nuovo guanto con retroazione tattile, chiamato TELEACT.

3. - I sensori per la telepresenza

3.1 - Generalità

Le esigenze della telepresenza impongono l'uso di svariati tipi di sensori. Di questi, alcuni sono gli stessi che vengono utilizzati per il funzionamento autonomo dei robot, mentre altri, vuoi per la loro natura, vuoi per il modo con cui vengono montati ed utilizzati, sono tipici di questa applicazione.

Anche se i metodi di misura e di trasferimento dei segnali sono oramai assestati, le esigenze della robotica hanno fatto sì che i criteri di utilizzazione dei dati sensoriali siano profondamente mutati negli ultimi anni. Infatti, mentre all'inizio i sensori erano usati essenzialmente come "misuratori" di quantità fisiche note, attualmente l'accento è posto soprattutto sull'uso che delle informazioni sensoriali si fa nel robot. Questo uso può essere profondamente diverso da quello per cui erano originariamente previsti i sensori impiegati: ad esempio, usare un misuratore di temperatura per rilevare la presenza di esseri umani in un ambiente non è certo un uso "classico" dell'informazione di temperatura.

Inoltre, va prendendo sempre più piede la tecnica della fusione multisensoriale (data fusion), in cui i dati provenienti da più sensori eterogenei vengono combinati, spesso in maniera variabile e dipendente dalle circostanze, per ottenere dati la cui natura è in alcuni casi completamente differente da quella che le informazioni provenienti dai singoli sensori potrebbero fornire. Ciononostante, le prestazioni dei sensori dei robot sono sempre piuttosto primitive se paragonate alle capacità umane.

E' opportuno citare brevemente i sensori che trovano applicazione nel problema in esame. Saranno quindi trascurati quei dispositivi che sono utilizzati esclusivamente per il funzionamento del telemanipolatore.

3.2 - Sensori visivi

I sensori più frequentemente impiegati nei sistemi di telepresenza sono ovviamente quelli visivi, in particolare le telecamere. Il segnale che essi forniscono è infatti direttamente trasmissibile all'occhio dell'operatore attraverso display a raggi catodici o, più modernamente, a cristalli liquidi. Generalmente si impiegano coppie di telecamere, i cui segnali sono distintamente inviati a due display, posti ciascuno davanti ad un occhio dell'operatore con l'interposizione di lenti apposite. Così facendo, è possibile ricreare l'effetto stereoscopico della visione binoculare, anche se, per renderlo realistico, è necessario fare in modo che le telecamere siano mobili, e ripetano con esattezza i movimenti della testa dell'operatore.

La visione stereoscopica non è infatti dovuta solo alla disponibilità di due immagini che il cervello elabora e fonde, ma anche allo spostamento relativo che gli oggetti apparentemente subiscono quando si cambia il punto di osservazione. Tale spostamento, che causa fra l'altro fenomeni di occlusione e di riapparizione di oggetti, spiega come l'uomo possa avere una buona percezione della posizione relativa degli oggetti anche se li osserva con un occhio solo.

La resa dell'effetto stereoscopico è molto problematica se manca questa possibilità di movimento, reale o simulato, in sincronia con quello della testa dell'operatore. Tale movimento è oramai comunemente impiegato in tutti i sistemi di telepresenza avanzati.

3.3 - Sensori di tipo non visivo

Le due altre categorie di sensori che possono trovare utili applicazioni nella telepresenza sono costituite dai sensori di distanza e da quelli di contatto. Genericamente parlando, i sensori di distanza misurano, servendosi di metodi diversi, la distanza degli ostacoli che si trovano davanti al sensore stesso. Tale misura può essere effettuata lungo l'asse del sensore, oppure in un'area di cui viene riportata, sotto forma di matrice, la distanza di ogni punto del sensore (range scan). Indipendentemente dal principio di funzionamento e dagli errori da cui possono essere affetti, vale per i dati forniti da questi sensori la considerazione generale che non esiste la possibilità di rappresentarli come si fa con i segnali visivi, perché l'uomo non è dotato di sensori equivalenti. Questo problema verrà analizzato più avanti.

I sensori di contatto dal canto loro possono fornire informazioni circa il loro contatto con una superficie solida. Questa informazione può essere binaria o analogica; nell'ultimo caso, esiste l'equivalente nel corpo umano, che è costituito dal senso del tatto. Sono stati costruiti a questo proposito sistemi in grado di replicare, in modo più o meno accurato e su superfici più o meno estese (la sola mano, il braccio, l'intero corpo) le sensazioni tattili. Va detto però che allo stato attuale della tecnologia questa riproduzione è sempre molto approssimativa, anche se per alcuni scopi può risultare sufficiente.

4. - Problemi di ricostruzione e di rappresentazione dell'ambiente

Il problema della ricostruzione delle sensazioni sensoriali a distanza, proprio della telepresenza applicata alla telemanipolazione, è stato spesso trascurato o perlomeno non affrontato in modo adeguato. Infatti, perché l'operatore abbia la sensazione di trovarsi nel luogo in cui si trova il manipolatore, occorrerebbe una enorme quantità di sensori, alcuni dei quali non possono essere costruiti artificialmente. Si pensi ad esempio al fatto che il corpo umano è dotato, su tutta la sua superficie, di recettori tattili e termici che determinano una serie di sensazioni fondamentale per stabilire, ad esempio, la situazione di "contatto" del corpo con oggetti esterni. Inoltre, l'uomo ha la possibilità di rilevare, anche se in modo approssimato, la posizione delle varie parti del proprio corpo, e le forze che su di esse agiscono. E' chiaro che, se si limita la riproduzione di queste sensazioni a un loro sottoinsieme (ad esempio, eliminando le sensazioni di temperatura, fondamentali per determinare la natura degli oggetti che si toccano) o se ne riduce la quantità (come si fa quando si usano guanti o tute avvolgenti, che rendono sì alcune sensazioni tattili e di forza, ma in misura molto inferiore a quella della pelle) si perde irrimediabilmente la sensazione di "presenza". Se non si prendono poi particolari accorgimenti, si possono generare addirittura stati di malessere nell'operatore, dovuti alla mancanza di determinate sensazioni che nella pratica normale sono associate ad altre. Si pensi per esempio ad un simulatore di automobile poco sofisticato: se viene simulata una curva, ma non viene fornita all'operatore la forza centrifuga o un suo surrogato, si ha immediatamente una sensazione sgradevole dovuta al fatto che ciò che gli occhi vedono non corrisponde a ciò che il resto del corpo percepisce.

Un altro fattore importante da tenere in considerazione è che i telemanipolatori sono spesso dotati di sensori che il corpo umano non possiede, o possiede con caratteristiche diverse: possiamo citare ad esempio il sonar, utilissimo strumento per la misura delle distanze, di cui l'uomo non ha un equivalente, e che pure andrebbe in qualche modo sfruttato per integrare le informazioni trasmesse all'operatore.

4.1 - Livelli di telepresenza

Da quanto detto appare evidente che il problema della telepresenza non è risolvibile con una semplice remotizzazione di dati sensoriali, ma richiede già ad un primo livello una elaborazione complessa, secondo lo schema di figura 1.

Lo schema di figura 1 non può però essere considerato soddisfacente, perché la sensorialità umana è solo in parte passiva, ed impiega invece in modo massivo azioni, sia sugli stessi sensori (ad esempio spostamenti della testa per cambiare il punto di vista), sia sull'ambiente (toccare oggetti, spostarli, ecc.). Lo schema deve essere quindi completato come mostrato in figura 2.

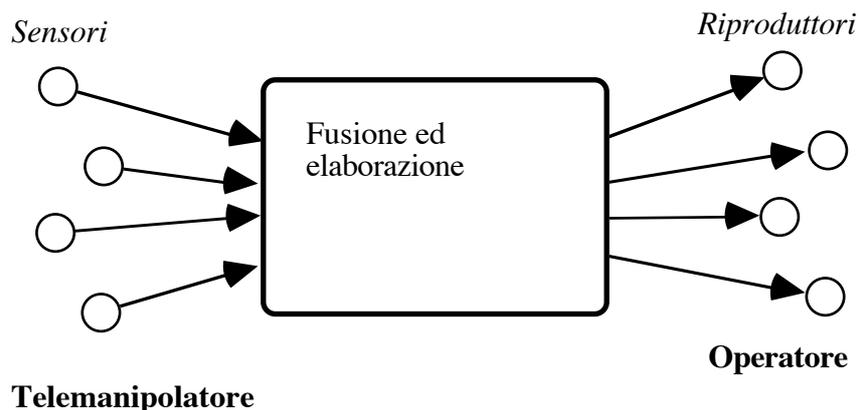


Fig. 1 - Primo livello di telepresenza

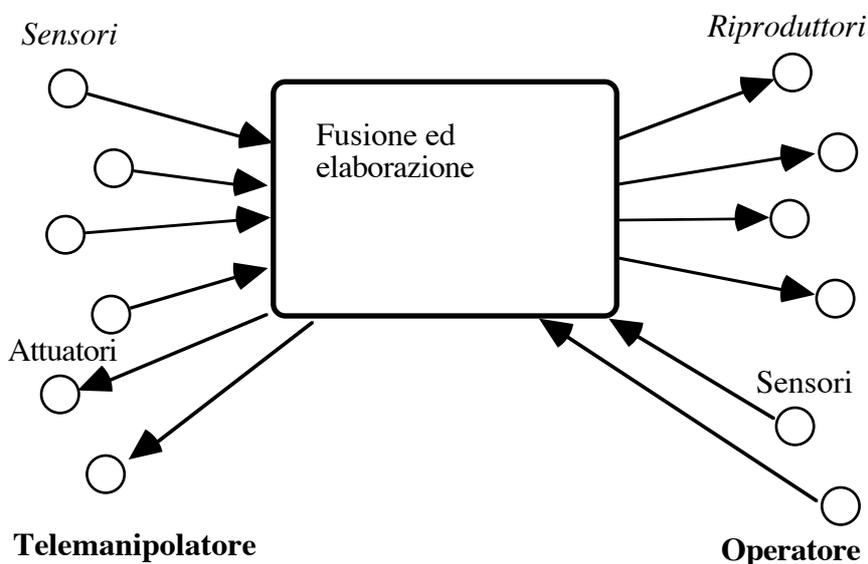


Fig. 2 - Secondo livello di telepresenza

Anche se il sistema di figura 2 può per certe applicazioni essere considerato soddisfacente [5, 6, 7, 8, 9], esso non è in grado di far fronte a diverse difficoltà, prima fra tutte l'impiego di sensori le cui informazioni non sono direttamente riproducibili e quindi comprensibili per l'uomo. Si pensi per esempio ad un dispositivo a laser in grado di fornire una mappa di distanze: è ovvia l'utilità di questa mappa, ad esempio per segnalare la vicinanza di oggetti pericolosi, ma è altrettanto ovvia l'impossibilità di rappresentare questa informazione in modo diretto. Una soluzione potrebbe essere quella di integrare l'informazione di "distanza pericolosa" con quella visiva, ad esempio utilizzando colori particolari per la rappresentazione degli oggetti interessati. Per rendere però possibile ciò, occorre che il sistema di elaborazione sia in grado di far corrispondere le informazioni di distanza con quelle visive. Esso deve quindi essere corredato di un modello del mondo, secondo quanto mostrato in figura 3.

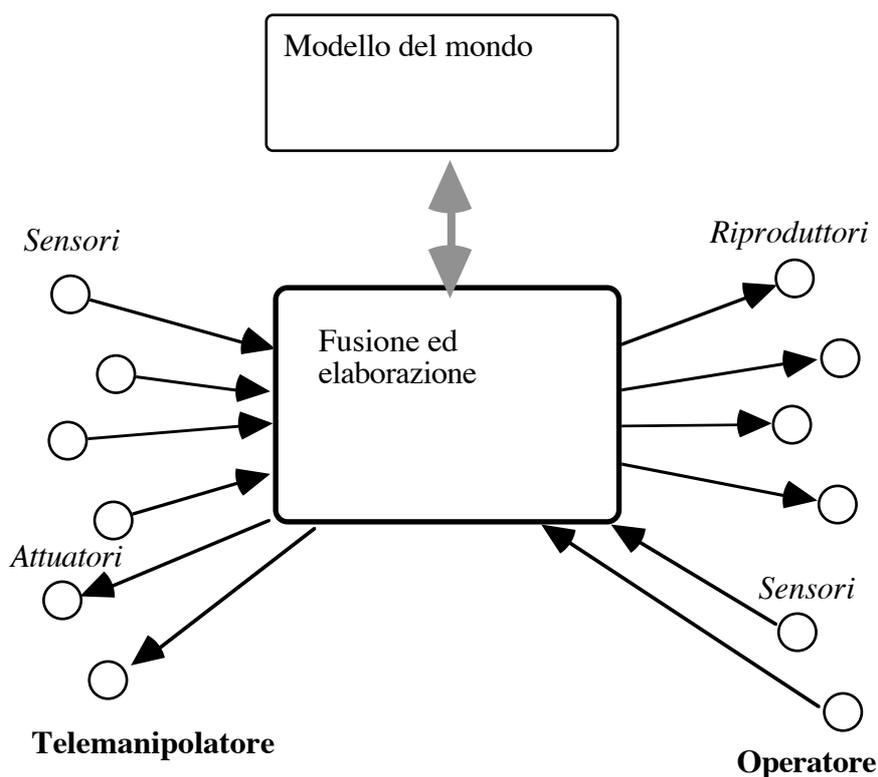


Fig. 3 - Terzo livello di telepresenza

Questo modello del mondo potrà essere, a seconda dei casi, precostituito utilizzando una descrizione formale dell'ambiente in cui il robot dovrà operare, o costruito in tempo reale durante il lavoro. E' evidente l'enorme complessità che sia l'una che l'altra soluzione implicano, e che costringono ad effettuare diverse scelte e semplificazioni. Va comunque sottolineato che lo scopo del presente lavoro è quello di fornire indicazioni perché la rappresentazione remota del mondo risulti accettabile ed utilizzabile, senza pretendere di fornire un completo effetto presenza, che d'altra parte non è ottenibile pienamente nemmeno in situazioni più semplici, quali ad esempio la riproduzione dei soli suoni.

4.2 - Fusione dei dati sensoriali

In genere, i progetti riguardanti la telepresenza non trattano il problema della fusione dei dati sensoriali, demandandolo al sistema sensoriale dell'uomo. In altre parole, le informazioni che provengono dai sensori del robot sono mandate singolarmente ai sensori dell'uomo: i dati visivi agli occhi, quelli tattili agli attuatori tattili (guanti, tute, ecc.), quelli sonori alle orecchie, ecc. Questa strategia, che è mostrata schematicamente in figura 4, può essere considerata corretta solo in un numero molto limitato di casi.

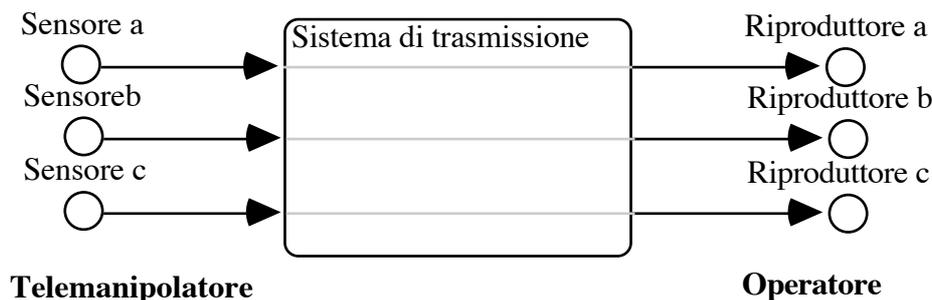


Fig. 4 - Corrispondenza “uno a uno” sensori - riproduttori

La proposta qui presentata invece ipotizza una separazione netta fra l’acquisizione di dati sensoriali e la loro riproduzione, sia per quanto riguarda il momento della loro acquisizione, sia per quanto riguarda le modalità di riproduzione. Una proposta in questo senso, a fronte di evidenti svantaggi dal punto di vista della complessità e quindi del costo del sistema, offre diversi vantaggi che possono essere così riassunti:

1. **Tempo di acquisizione dei dati.** Molti sensori hanno un tempo di acquisizione dei dati notevolmente lungo. Tuttavia, in un ambiente statico si può pensare che questi dati siano poco variabili nel tempo. Ciò vuol dire che dati acquisiti in un istante precedente possono essere utilizzati anche successivamente. A titolo di esempio, si consideri la possibilità di acquisire e memorizzare un’immagine video “panoramica”, ottenuta combinando insieme un certo numero di immagini video prese con diverse angolazioni della telecamera. Se l’ambiente è statico, da queste immagini è possibile ottenere, in qualunque istante, un’immagine realistica dell’ambiente visto dallo stesso punto sotto una qualunque angolazione.
2. **Complessità dell’elaborazione.** Anche la quantità di elaborazione necessaria (per esempio, per combinare un’immagine video monoscopica ed un “range scan” della stessa zona) richiede tempi piuttosto lunghi, ed è impensabile che l’operatore, dopo aver rivolto la propria attenzione verso una determinata zona, debba attendere diversi secondi prima di vedere gli oggetti in essa contenuti. Se però l’esecuzione e l’elaborazione sono state fatte “a priori” il problema non sussiste, in quanto il display può essere istantaneamente aggiornato, a patto di disporre di una sufficiente potenza di calcolo e di una adeguata capacità di memoria.
3. **Riduzione della banda.** In molti casi esiste una oggettiva limitazione alla larghezza di banda utilizzabile per il collegamento fra il robot e l’operatore. Una soluzione, almeno parziale, al problema può essere quella di utilizzare tutti i tempi morti per trasmettere informazioni e memorizzarle in prossimità dell’operatore.
4. **Tempo di comunicazione.** Questo problema, che è ben noto nel campo dell’esplorazione planetaria, comporta la necessità assoluta di utilizzare tecniche di pre-acquisizione dei dati e di display predittivi. Esso esula tuttavia dallo scopo della presente trattazione.

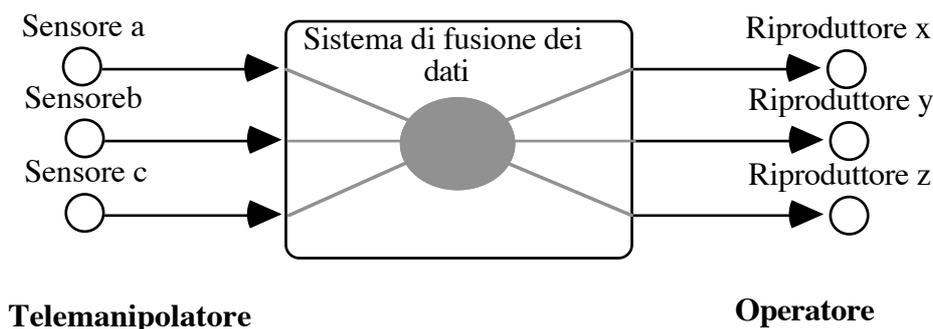


Fig. 5 - Introduzione della fusione dei dati

Passando ad analizzare i vari tipi di informazioni sensoriali, si può effettuare una divisione fra quelle informazioni che devono assolutamente essere acquisite e trasmesse in tempo reale¹ e quelle che possono essere acquisite anche in anticipo.

Nella prima categoria rientrano ovviamente le informazioni visive relative alla zona in cui il robot opera cambiamenti fisici (manipolando o spostando oggetti) e quelle sonore relative all'azione del robot (rumore degli oggetti spostati, di eventuali urti, ecc.).

Alla seconda possono essere invece fatti appartenere i dati che variano più lentamente: mappe simboliche e geometriche dell'ambiente, mappe di distanza ottenute con sensori sonar e laser, viste generali dell'ambiente².

4.3 - Costruzione del modello dell'ambiente

Da quanto detto precedentemente appare chiaro che la disponibilità di un modello CAD del mondo in cui si trova il telemanipolatore permette di utilizzare i dati forniti dai sensori in maniera più efficiente. In generale però il modello del mondo non è disponibile a priori, se non in forma inadatta ed approssimativa. Si consideri ad esempio il caso di un ambiente industriale: in genere di esso sono disponibili piante che, pur contenendo una descrizione abbastanza esatta delle dimensioni degli ambienti, non prendono assolutamente in considerazione le attrezzature. Queste ultime sono descritte nei disegni di layout degli impianti, ma sempre in maniera approssimativa, incompleta e, ciò che più conta, inadatta allo scopo della presente trattazione. D'altra parte, la costruzione manuale di un modello dell'ambiente è un'operazione estremamente dispendiosa. La soluzione risiede probabilmente nell'acquisizione semi-automatica del modello dell'ambiente, effettuata mediante un sistema basato sulla fotogrammetria dei vicini.

Un possibile modo di procedere è il seguente:

¹A meno di non utilizzare tecniche di predizione, che però esulano dallo scopo di questo lavoro.

²Le condizioni sotto cui queste assunzioni sono valide dipendono ovviamente dalle caratteristiche dell'ambiente: una strada con traffico presenta una variabilità estremamente maggiore di un edificio disabitato, ma la mappa generale della strada non varia, mentre quella dell'edificio, nel caso per esempio di un incendio, può subire cambiamenti improvvisi ad esempio in conseguenza di crolli.

- Vengono acquisite diverse immagini dell'ambiente di lavoro del robot, o da un operatore, o attraverso le telecamere dello stesso robot manovrato opportunamente. In ogni immagine devono essere individuati alcuni *punti di appoggio*, le cui coordinate spaziali devono essere note.
- Ogni immagine digitalizzata è correlata ad altre immagini, riprese da altri punti di vista, aventi come riferimenti, in tutto od in parte, gli stessi punti d'appoggio.
- Il risultato dell'operazione è un'immagine tridimensionale di notevole accuratezza, che, unitamente agli opportuni attributi fisici degli oggetti, introdotti direttamente dall'operatore, costituisce il modello dell'ambiente a cui fare riferimento durante le operazioni del robot.

E' evidente che il metodo di acquisizione del modello del mondo deve essere per quanto possibile efficiente in termini di accuratezza e di tempi di lavoro. Per ottenere ciò, si può fare ricorso ad elementi innovativi per agevolare il compito degli operatori, come ad esempio le schede PhOX della KOH-I-NOOR. Tali schede, basate su micro-processore, sono inseribili in PC di tipo IBM AT sul quale vengono generate le immagini di sintesi, ad esempio con AutoCAD 11.

La stazione di lavoro per l'acquisizione del modello del mondo può quindi essere costituita da un PC AT, dalla scheda PhOX, da una tavoletta grafica per la definizione delle coordinate dei punti significativi, e dall'interfaccia per l'acquisizione delle immagini dell'ambiente reale.

La fotogrammetria dei vicini usata nelle schede PhOX si basa, come già detto, sulla ripresa fotografica dell'ambiente sul quale vengono applicate mire o punti di appoggio che debbono essere visibili sul fotogramma e dei quali debbono essere note le coordinate rispetto al sistema riferimento terreno.

Il metodo di lavoro è basato sulla trasformazione lineare diretta che effettua la correlazione tra le coordinate tridimensionali dei punti di uno spazio a tre dimensioni e le coordinate piane dei corrispondenti punti giacenti su di un piano, il piano proiettivo.

In pratica, l'operatore CAD digita sulla tavoletta grafica i punti corrispondenti nelle diverse fotografie (ad esempio, quattro fotogrammi); ciascun punto, tramite la scheda PhOX, determina le coordinate spaziali del punto stesso che, attraverso il CAD, viene restituito sullo schermo. Il reticolo dei punti restituiti consente la precisa determinazione delle caratteristiche tridimensionali dell'ambiente restituito.

4.4 - Utilizzazione del modello dell'ambiente

Per poter sfruttare convenientemente i dati contenuti nel sistema, che sono in parte reali (acquisiti in tempo reale o differito) e in parte sintetici o semi-sintetici, occorre integrarli utilizzando un sistema grafico di caratteristiche opportune. Allo stato attuale della tecnologia, è possibile realizzare sistemi grafici con ottime caratteristiche utilizzando come base sistemi CAD operanti su calcolatori di tipo PC della fascia alta, integrandoli dal punto di vista software con pacchetti per il miglioramento delle caratteristiche grafiche (animators, ray tracing, ecc.), e dal punto di vista hardware con transputer per aumentarne le capacità di elaborazione. E' così possibile arrivare ad ottenere immagini sintetiche sufficientemente realistiche ad un ritmo di almeno 25 immagini/s.

La necessità di utilizzare un sistema CAD deriva principalmente dal fatto che ogni serie di dati sensoriali viene acquisita da un determinato punto di vista. Se si vogliono utilizzare gli stessi dati per rappresentare la situazione vista da una posizione diversa, è necessario in molti casi che la prima acquisizione sia stata “compresa” e modellizzata opportunamente. Si consideri ad esempio il caso di una classica “scansione sonar”, cioè di un diagramma che fornisce, per ogni direzione in un piano orizzontale, la distanza dal sensore dell’oggetto più vicino. In questo caso, spostare il punto di misura richiede solo una trasformazione geometrica per ottenere una rappresentazione, plausibile, delle distanze dal nuovo punto di vista. Se invece pensiamo ad una immagine stereoscopica ripresa con due telecamere, ci rendiamo conto che, per poter ottenere l’immagine di ciò che si vedrebbe spostando le telecamere, occorre interpretare l’immagine, almeno nel senso di effettuare un matching per calcolare i dati di profondità. Quindi, la serie di passaggi segmentazione – rappresentazione iconica – modellizzazione, fatta in maniera automatica o semi-automatica, diventa essenziale. Ricordiamo inoltre che in genere si vuole associare ad ogni oggetto l’informazione proveniente da sensori eterogenei.

Una volta costruito il modello del mondo, esso può essere usato per rappresentare informazioni visive e non in maniera relativamente semplice: ad esempio, simulare la visione degli oggetti da un punto di vista diverso non è più un problema, così come associare ad ogni oggetto particolari caratteristiche (colore, brillantezza, tessitura della superficie, ecc.) associabili a specifici parametri (distanza, pericolosità, calore, ecc.).

5. - Un esempio di realizzazione pratica

In questo paragrafo descriveremo i punti salienti di una realizzazione sperimentale, che è stata effettuata allo scopo di verificare la fattibilità di quanto precedentemente descritto.

Per poter supervisionare, con interazione sensoriale e motoria, un ambiente remoto ed interagire all’interno di esso (telepresenza o tele-esistenza) è stato assemblato un *casco interattivo* e sono stati realizzati un sistema sensoriale montato su un robot di tipo industriale e l’opportuno software di interfacciamento e di controllo, secondo lo schema generale mostrato in figura 6.

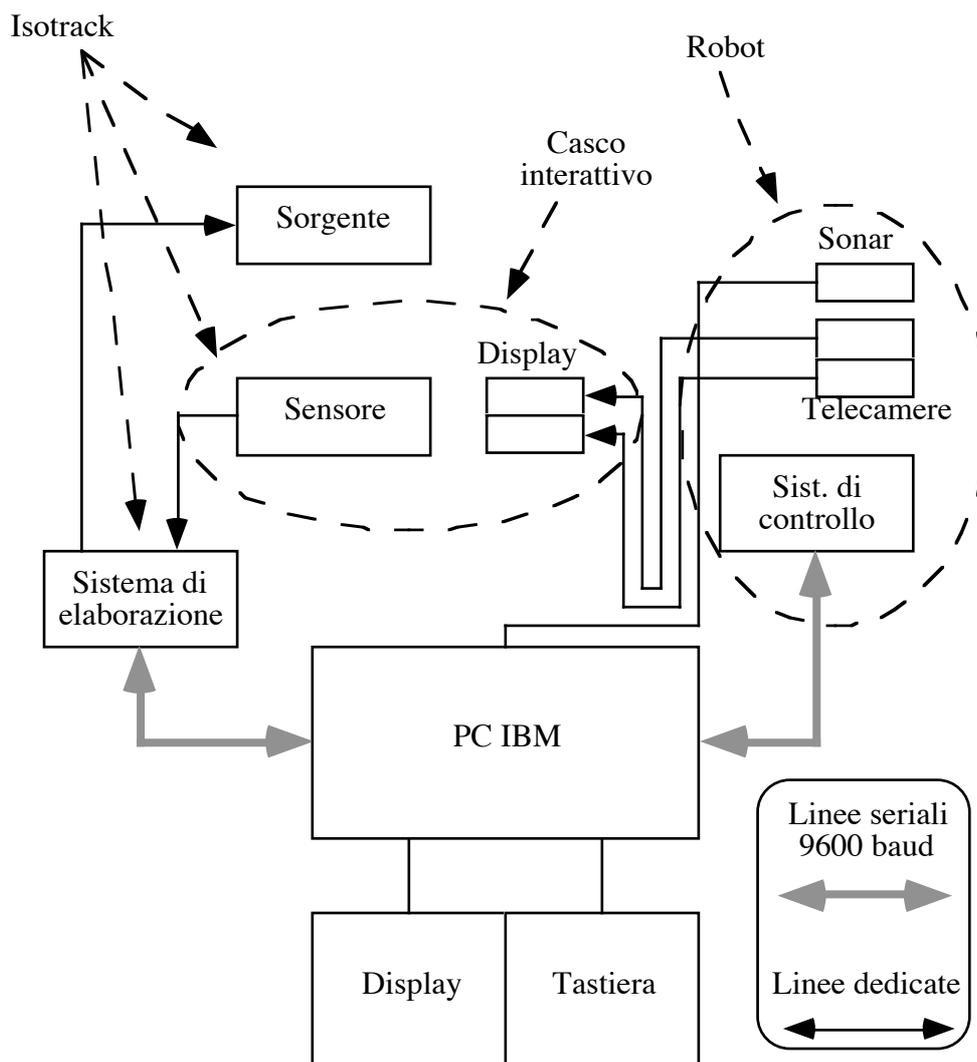


Fig. 6 - Rappresentazione schematica del sistema.

Il flusso di dati fra questi sottosistemi è suddiviso in tre canali:

- 1) Ricezione delle immagini e visualizzazione sul casco interattivo.
- 2) Rilevazione della posizione della testa dell'operatore, elaborazione del segnale risultante, e corrispondente movimentazione del robot.
- 3) Lettura del sonar, elaborazione del segnale di distanza e visualizzazione di quest'ultimo.

5.1 - Il casco interattivo

E' stato costruito un casco (figura 7) dotato di due display a cristalli liquidi a colori, da 3 pollici, posti frontalmente agli occhi dell'operatore, e di due lenti speciali, frapposte tra i display e gli occhi. Le lenti permettono all'occhio di mettere a fuoco i display nonostante la piccola distanza, e soprattutto di dilatare l'immagine (effetto grandangolo) in modo da dare una sensazione di reale presenza nell'ambiente in cui si trovano le telecamere e quindi il robot.



Fig. 7 - Il casco interattivo.

Sul casco è stato montato un sensore 3SPACE-Isotrak allo scopo di rilevare in tempo reale la posizione e l'orientamento della testa dell'operatore. Un opportuno software si occupa di elaborare il segnale dell'Isotrak e quindi di movimentare le telecamere in modo da ripetere il più fedelmente possibile il movimento della testa dell'operatore.

Il sistema 3SPACE ISOTRAK é composto da una sorgente, un sensore e un sistema elettronico di elaborazione. Il sensore utilizza un campo magnetico a bassa frequenza (generato dalla sorgente), per determinare la posizione e l'orientamento, fornendone le sei coordinate (tre cartesiane e tre angolari). Questa informazione viene trasmessa al computer attraverso una linea seriale in ASCII o in codice binario. Il sistema fornisce i dati di circa 20 punti/s.

La precisione dipende dalla distanza del sensore dalla sorgente, e varia fra 3 e 6 mm RMS. La precisione angolare è inferiore ad 1° RMS, indipendentemente dalla distanza. La distanza massima fra sorgente e sensore è di 750 mm.

5.2 - Il sistema remoto

Sul robot sono state installate due telecamere (MICAM COLOR 650 della Digital Vision) ad alta sensibilità (2.5 lux) con obiettivi grandangolari, in modo da coprire un campo visivo di circa 90° per ciascuna telecamera. Le telecamere sono posizionate in modo da replicare la struttura fisica del sistema visivo umano, con gli assi degli obiettivi paralleli fra loro e distanti 70 mm (figura 8).



Fig. 8 - Le telecamere montate sul braccio

Questa disposizione permette di ricostruire un'immagine stereoscopica molto simile a quella rilevata dagli occhi dell'uomo.

Le telecamere, inoltre, sono movimentate dal robot sulla base delle informazioni ricevute dall'Isotrak, in modo tale da replicare il movimento della testa dell'operatore.

Sul robot è stato montato inoltre un sensore ad ultrasuoni (POLAROID) che fornisce al PC, il valore della distanza rilevata, lungo l'asse mediano delle telecamere.

5.3 - Il robot GT-6A

Il GT-6A è un braccio articolato, prodotto dalla ditta francese ROBOSOFT, dotato di sei gradi di libertà rotazionali, che può essere montato su una base mobile (fornita dalla stessa ditta). Tale robot è dotato di un controllore basato su un bus VME e del sistema operativo in tempo reale "ALBATROS", progettato specificamente per il controllo di più assi.

5.4 - Il sistema di calcolo

Nella configurazione attuale è stato impiegato un calcolatore di tipo PC, che funge da elaboratore dei dati e da gestore della comunicazione fra i moduli.

Il compito principale del software che è stato sviluppato consiste nella trasformazione dei dati ricevuti dall'Isotrak in comandi di movimento per il robot. La particolare natura del sistema di controllo del robot ha reso necessario il calcolo delle trasformazioni cinematiche inverse necessarie a pilotare direttamente i sei motori del robot. Il software effettua anche tutte le necessarie conversioni di coordinate.

Durante lo sviluppo del programma si è data prioritaria importanza alla velocità di esecuzione per poter avere una elevata frequenza nella successione di movimenti del robot.

Il flusso di dati verso l'operatore consiste in dati numerici e/o brevi messaggi sovrapposti alle immagini reali visualizzate sul casco interattivo e su un monitor ausiliario (quello del PC); ciò consente all'operatore di avere un feedback di dati dall'ambiente remoto senza distogliere l'attenzione dal controllo del robot³.

La figura 9 mostra una veduta d'insieme del sistema.



Fig. 9 - Vista generale del sistema.

³Nella versione attuale i dati provenienti dal calcolatore non vengono visualizzati sui display del casco interattivo.

6. - Conclusioni e sviluppi futuri

Anche se il sistema sperimentale qui descritto non si presta all'implementazione del modello del mondo cui si è fatto cenno, per gli evidenti limiti della sua capacità di calcolo, alcuni risultati ottenuti sono stati determinanti per valutare le possibilità di realizzazione di un sistema più complesso [10]. Inoltre, il sistema può già essere utilizzato in alcuni casi, migliorandone le prestazioni con semplici interventi.

Per quanto riguarda la proposta di integrazione e modellizzazione delle informazioni sensoriali, è evidentemente necessaria una approfondita sperimentazione, tesa soprattutto a verificare le modalità di *matching* fra i dati sensoriali acquisiti in tempo reale e quelli generati artificialmente.

Le possibilità di realizzare in modo completo il sistema proposto, dato il costante sviluppo dei sistemi di calcolo e delle tecniche di generazione e di rappresentazione della realtà virtuale sono comunque ottime. La disponibilità di innumerevoli pacchetti software per il trattamento delle immagini fornite dai sistemi CAD permettono inoltre di realizzare il software occorrente in tempi ridotti.

7. - Bibliografia

- [1] Chung, J. C., et al.: *Exploring Virtual Worlds with Head-Mounted Displays*, Proc. SPIE, Vol. 1083, Los Angeles, Ca., 1989.
- [2] Sutherland, I.E.: *The Ultimate Display*, Proc. IFIP Congress 2, 1965.
- [3] Engelberger, J.: *Robotics in service*, Kogan Page Ltd, London, 1989.
- [4] Stone, R. J.: "*The Best Of Both Worlds*": *A Combined Virtual-Real Human-Computer Interface For Telepresence And Remote Driving*, proc. of '91 ISART - International Symposium On Advanced Robotics Technology, Tokyo, 1991.
- [5] Stone, R. J.: *The UK VERDEX Project, 2 Years On: Virtual Reality, Telepresence and the Human Factor*, Proc. Imagina '92, Monte Carlo, 1992.
- [6] McGreevy M. W., Stoker C. R.: *Telepresence for Planetary Exploration*, SPIE The International Society for Optical Engineering, Cooperative Intelligent Robotics in Space, R. J. de Figueiredo and W. E. Stoney Editors, vol. 1387 pp. 110-123, Boston, Massachusetts, November, 1990.
- [7] Noseworthy J. R., Gerhardt L. A.: *Three dimensional vision-requirements and applications in a space environment*, SPIE The International Society for Optical Engineering, Cooperative Intelligent Robotics in Space, R. J. de Figueiredo and W. E. Stoney Editors, vol. 1387 pp. 26-37, Boston, Massachusetts, November, 1990.
- [8] McDonnell J. R., Solorzano M. R., Martin Stephen W. M., Umeda A. Y.: *A Head Coupled Sensor Platform for Teleoperated Ground Vehicles*, Unmanned Systems, Vol. 8, N. 4, pp. 33-38, Fall Editore S.p.A., 1990.

- [9] Dimitriadis B. Papanikolau K.: *Stereoscopic Vision for AGV Guidance*, NATO Asi Series., vol 171, Expert Systems and Robotics, Springer-Verlag, 1991.
- [10] Terceros, M.: *Telepresenz intelligente*, Tesi di laurea, Università degli Stdi di Milano, 1992.

Ringraziamenti

La parte sperimentale del presente lavoro è stata svolta con il contributo del CESI SpA - Milano, a cui va il ringraziamento degli autori.