

Introduzione

Al giorno d'oggi la tecnologia ha raggiunto livelli tali da offrire un gran numero di modalità di input/output.

Nonostante questo molte interfacce continuano a interagire con devices “obsoleti” e non molto intuitivi quali mouse e tastiera e questo pregiudica il livello di interattività.

Per certi casi questi devices si rivelano ancora utili e intuitivi e non si sente il bisogno di interfacce di più alto livello, di più facile utilizzo e di maggior naturalezza ma per un'interazione più avanzata e naturale questi interfacce non si adattano bene alle esigenze dell'uomo.

Purtroppo queste tipologie di interfacce sono rare e per questo non si riescono a trovare standard da applicare così che ogni nuova interfaccia ad-hoc richiede uno studio del settore di applicazione affinché i gesti, i simboli e la voce trasmessa siano i più giusti e intuibili per un campo di applicazione generale.

L'utilizzo di queste interfacce ovviamente permette di avere input più variegati, più “corretti” e sicuramente più intuitivi.

In questo paper verranno analizzati tre tipi di possibili interfacce: gesture recognition, symbols recognition e voice recognition e si cercherà di definire un vocabolario che non pretende di essere universale ma per lo meno che abbia la possibilità di coprire un ampio numero di scenari.

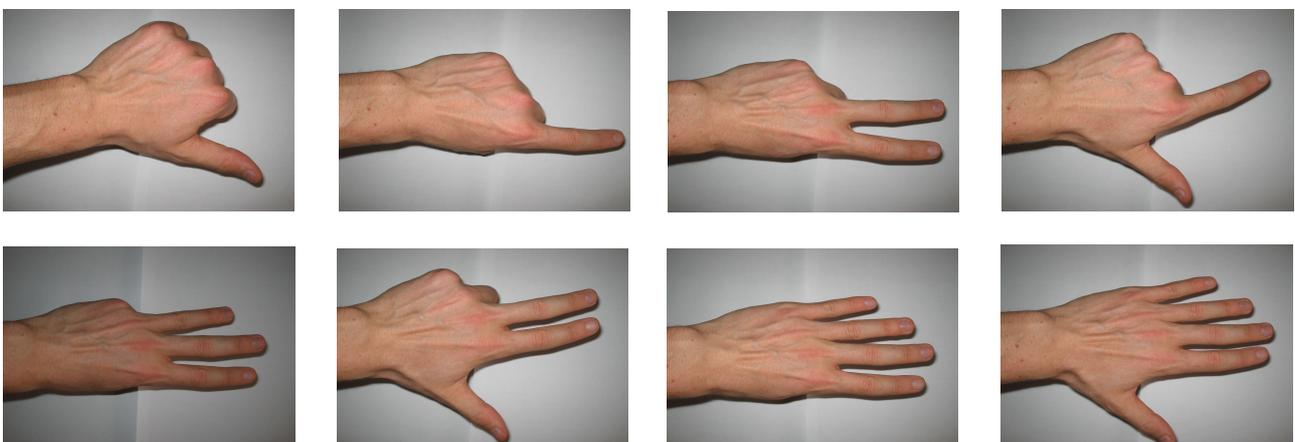
Il sistema “E-2?” usa una telecamera per il gesture recognition e il symbols recognition, un microfono/sintetizzatore vocale per il voice recognition.

Il tutto è realizzato utilizzando un insieme di movimenti, gesti del corpo significativi che sono riconosciuti dal sistema. Questi gesti sono poi interpretati come comandi o richieste che a loro volta scatenano determinate azioni.

GESTURE RECOGNITION

Cercando di pensare alle possibili interazioni che ci potrebbero essere tra un uomo e un robot, evitando di entrare in uno scenario particolare e cercando di stare sul generale abbiamo identificato questo insieme di possibili gesti/movimenti significativi:

- 1) i numeri da 1 a 5 andando a riconoscere la disposizione delle dita di una mano e i numeri da 6 a 10 utilizzando entrambi le mani come riportato in figura:



- 2) gesti direzionali con il braccio ben disteso per indicare i movimenti che il robot deve fare (sia con braccio sinistro che braccio destro) come riportato in figura:



- 3) movimenti circolari con l'indice della mano (per simulare il disegno di un cerchio) con la mano sinistra, con la destra o con entrambe come riportato in figura:



- 4) movimenti di alzare e abbassare le braccia (contemporaneamente o una per volta) evitando di tendere il braccio come riportato in figura:



5) saluto (waving) come riportato in figura:



6) “vieni qui” con il dorso della mano verso il robot avvicinando l'avambraccio al proprio petto come mostrato in figura (o con la mano sinistra o con la destra, evitare con entrambi perché una possibile implementazione potrebbe rilevare la sola posizione del pollice) come mostrato in figura:



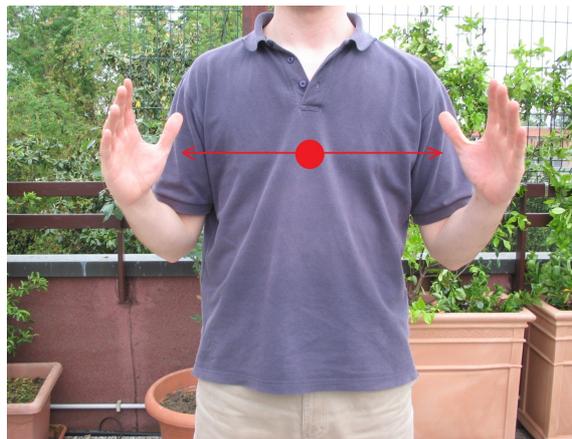
7) “vai via” con il palmo della mano verso il robot allontanando l'avambraccio dal proprio petto come mostrato in figura (o con la mano sinistra o con la destra, evitare con entrambi perché una possibile implementazione potrebbe rilevare la sola posizione del pollice) come mostrato in figura:



8) entrambi i pollici e indici delle mani in posizione ortogonale che formano gli angoli di una cornice per scattare una foto come mostrato in figura:



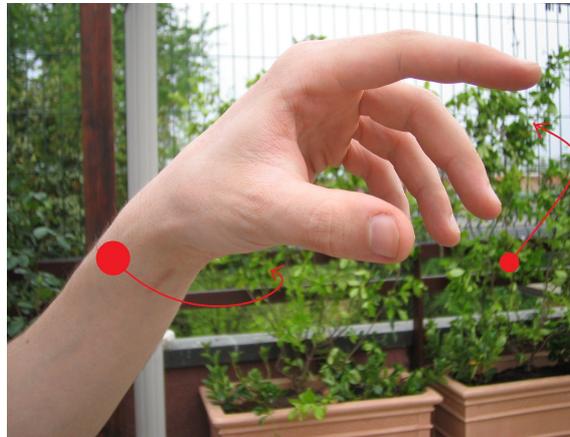
9) Da mani quasi congiunte che vanno ad allargarsi per fare lo zoom-in della telecamera come mostrato in figura:



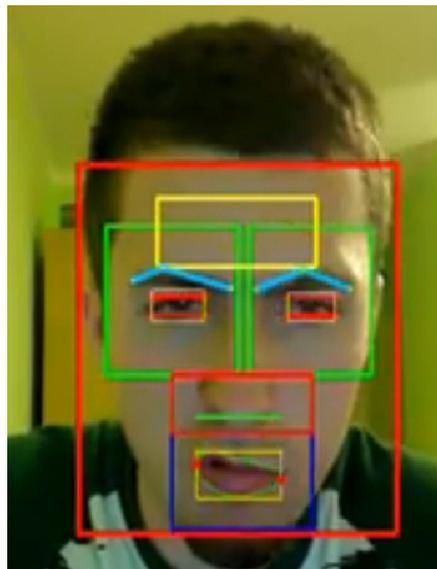
10) da mani distanti a mani quasi congiunte per fare zoom-out della telecamera come mostrato in figura:



11) una volta fatto zoom-in posso fare gesti più precisi e percepire per esempio attraverso il movimento congiunto di polso e dita per la torsione del collo o delle braccia del robot (gesto specifico ancora da determinare) come mostrato in figura:



12) una volta fatto lo zoom-in si possono anche intercettare le caratteristiche del volto (fronte, sopracciglia, occhi, bocca, viso, naso) al fine di “interpretare” lo stato d'animo della persona. (<http://www.youtube.com/watch?v=V7UdYzCMKvw>) come mostrato in figura: (OpenCV --><http://opencv.willowgarage.com/wiki/>)



di seguito vengono riportate alcune foto raffiguranti possibili gesti riconoscibili dal robot per alcuni giochi (in questo caso per esempio alcuni animali).

13) gallina



14) elefante



15) serpente



16) scimpanzé



17) uccello



18) leone (o quadrupede in generale)



19) fenicottero



20) coccodrillo



21) canguro



infine per l'implementazione di alcuni giochi d'avventura potrebbe essere utile l'utilizzo di una pistola



SYMBOLS RECOGNITION

Facendo un'analisi sui possibili simboli impiegabili per la comunicazione col robot “E-2?” si è subito arrivati alla conclusione che sia meglio differenziare le forme degli stessi invece che i colori in quanto una particolare o diversa intensità luminosa potrebbero distorcere la vera natura del colore compromettendo la corretta ricezione e lo scatenamento delle azioni ad esso correlate.

Inoltre non si ritiene di porre vincoli sul mezzo su cui vadano posti/disegnati i simboli. Per questo quando si parla di simbolo si intende esclusivamente la forma e non il supporto su cui è disegnato.

Per questo ci vengono in aiuto librerie già testate e affidabili come ArToolkit che appunto rimuovono questo vincolo. Unico requisito è che il simbolo sia il più definito possibile al fine di evitare ritardi eccessivi nell'elaborazione o addirittura “fraitendimenti”.

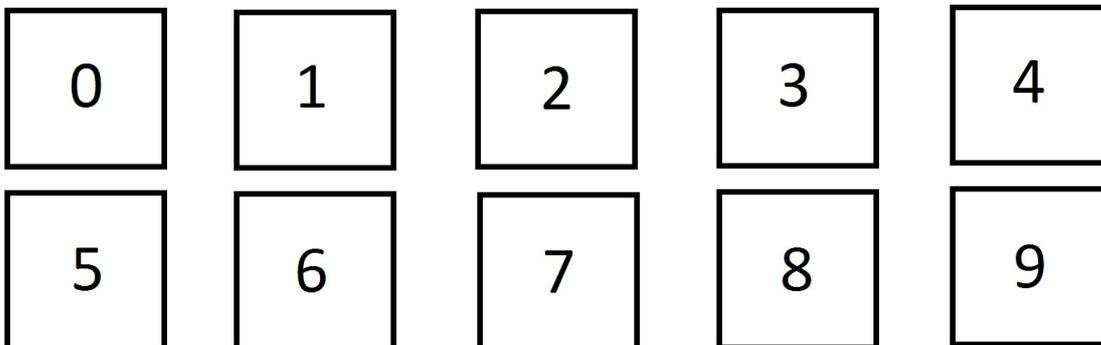
(<http://www.youtube.com/watch?v=dbm0TSthp5o&feature=related>)

(<http://www.youtube.com/watch?v=nQxzumHN4ps&feature=related>)

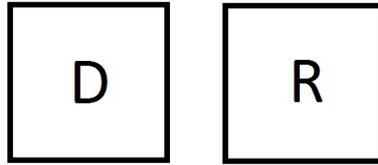
(ArToolkit--> <http://artoolkit.sourceforge.net/>).

Di seguito possibili simboli:

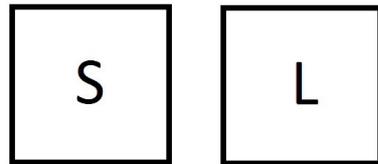
- numeri da 0 a 9 e loro composizione per la rappresentazione dei numeri come mostrato in figura (riferimento gesto 1)



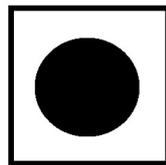
- "D"/"R" per indicare il movimento a destra come mostrato in figura (riferimento gesto 2)



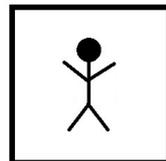
- "S"/"L" per indicare il movimento a sinistra come mostrato in figura (riferimento gesto 2)



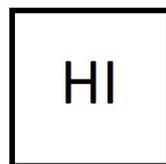
- cerchio come mostrato in figura (riferimento gesto 3)



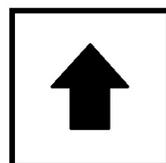
- omino stilizzato con braccia alzate come mostrato in figura (riferimento gesto 4)



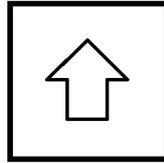
- "Hi" come mostrato in figura (riferimento gesto 5)



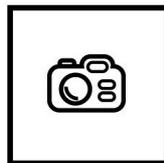
- freccia piena per indicare l'avvicinamento come mostrato in figura (riferimento gesto 6)



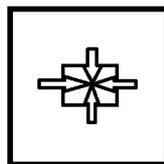
- freccia vuota per indicare l'allontanamento come mostrato in figura (riferimento gesto 7)



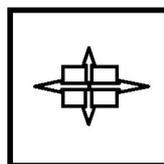
- macchina fotografica stilizzata per lo scatto di una foto come mostrato in figura (riferimento gesto 8)



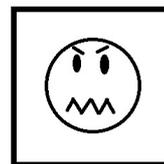
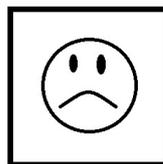
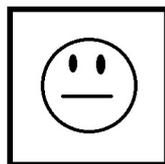
- quadrato con frecce entranti per indicare lo zoom-in come mostrato in figura (riferimento gesto 9)



- quadrato con frecce uscenti per indicare lo zoom-out come mostrato in figura (riferimento gesto 10)



- emoticons che identificano felicità, indifferenza, tristezza, rabbia come mostrato in figura (riferimento gesto 12)



- altri simboli/loghi specifici per indicare i vari punti chiave all'interno della fiera affinché il robot riesca ad orientarsi (un esempio è mostrato in figura)



tutti i simboli presentati sopra e quelli che eventualmente si aggiungeranno in seguito non devono necessariamente essere orientati poiché potrebbe risultare scomodo e in alcuni casi difficile. Dato che una qualsiasi rotazione è accettabile, è necessario però che i simboli siano tutti diversi l'uno dall'altro per evitare “equivoci” nel riconoscimento e quindi dello scatenarsi dell'azione relativa.

VOICE RECOGNITION

Per ottenere la massima interazione con il robot l'interfaccia migliore sarebbe senza dubbio quella vocale. Specifiche parole all'interno di frasi complesse potrebbero essere interpretate dal robot e attuate secondo il loro specifico significato. Questo studio ovviamente richiede un vocabolario più articolato di quello mostrato nei punti precedenti per ovvi motivi quindi di seguito si riportano solo le parole la cui composizione forma il significato dei punti precedenti:

- vieni
- vai
- via
- fermo
- destra
- sinistra
- ciao
- triste
- felice
- arrabbiato
- foto
- scatta
- zoom-in
- zoom-out

queste le voci fondamentali a qui andrebbero poi associate quelle specifiche per ogni scenario.