

Interfaccia gestuale per il controllo di robot

Alessandro Carfi, Fulvio Mastrogiovanni

Dipartimento di Informatica, Bioingegneria, Robotica e Ingegneria dei Sistemi (DIBRIS), Università degli Studi di Genova, Via Opera Pia 13, 16145, GE

alessandro.carfi@dibris.unige.it, fulvio.mastrogiovanni@unige.it

Abstract

L'evoluzione dei sistemi di produzione per l'industria 4.0, prevede una stretta interazione tra gli operatori umani e i robot. Tuttavia, oggi giorno, le interfacce a disposizione dell'operatore per comunicare con la macchina sono limitate e non sfruttano i canali comunicativi naturali di una persona. L'articolo introduce un'interfaccia uomo-robot che sfrutta la naturale propensione delle persone per la comunicazione gestuale. L'interfaccia gestuale, che sfrutta la tecnologia indossabile per la percezione dell'operatore, permette un'interazione naturale con il robot e punta ad aumentare la sicurezza della cooperazione uomo-macchina e a ridurre lo sforzo cognitivo dell'operatore.

1 Introduzione

Il progetto *Industria 4.0* del governo federale tedesco prevede uno stretto rapporto tra produttori di beni e consumatori finali, questo richiede un'alta dinamicità delle fabbriche del futuro [Kagermann *et al.*, 2013]. Lo stretto rapporto tra consumatori finali e produttori implica ampia varietà di prodotti e brevi cicli di produzione. Per soddisfare la richiesta di personalizzazione dei prodotti, limitando i costi di produzione, sono richiesti alti livelli di flessibilità nella linea di produzione nella catena di distribuzione. Le *Smart factories* affrontano questi requisiti adattando e modificando sia il processo produttivo sia quello della filiera di distribuzione alle necessità di un mercato in evoluzione [Lucke *et al.*, 2008], e i robot sono un componente fondamentale di questa innovazione. Nelle industrie del futuro, gli operatori umani dovranno lavorare a stretto contatto con i robot che dovranno essere spesso riconfigurati per adattarsi ai nuovi cicli produttivi. Il paradigma dell'Industria 4.0 richiede dunque, la disponibilità di tecniche per permettere una facile e veloce riconfigurazione dei robot. La Programming by Demonstration (PbD) risolve questi problemi permettendo di insegnare ad un robot una nuova azione intuitivamente. La PbD prevede due fasi: insegnamento, in cui l'operatore mostra uno o più esempi dell'azione desiderata e apprendimento, in cui gli esempi vengono processati per determinare il comportamento del robot. Diverse soluzioni sono state esplorate per la fase d'insegnamento dall'utilizzo di un semplice teach pendant per mostrare al robot come

muoversi, fino all'utilizzo di tecniche di computer vision che permettono al robot di imparare dal movimento dell'insegnante umano. Il Kinesthetic Teaching (KT) è una tecnica di insegnamento, ampiamente esplorata sia nel mondo accademico che come soluzione industriale, per la quale un operatore è in grado di insegnare al robot un particolare compito muovendolo fisicamente. Il KT rende particolarmente veloce ed intuitivo riconfigurare un robot per un nuovo compito ma, per quanto il processo di insegnamento possa essere considerato naturale, presenta alcune criticità che abbiamo analizzato in lavori precedenti [Carfi *et al.*, 2019].

L'operatore di una *Smart Factory*, oltre a necessitare di strumenti per riconfigurare velocemente il robot, deve poter controllare il robot direttamente. Tradizionalmente i teach pendant sono sempre stati utilizzati dagli operatori per interfacciarsi con i macchinari e, solo recentemente, è stato introdotto anche nel mondo industriale l'utilizzo di interfacce basate sull'utilizzo di schermi touch. Questo tipo di interfaccia prevede che l'operatore abbia le mani libere; la dotazione antinfortunistica (p.es. guanti da lavoro) può impedirne il corretto funzionamento. Queste limitazioni possono condizionare l'utilizzo delle interfacce tradizionali specialmente in un contesto collaborativo in cui l'operatore deve svolgere congiuntamente con il robot un'operazione ma contemporaneamente ha bisogno di comunicare con lo stesso, sia per ragioni di sicurezza che per completare l'operazione. È stato studiato l'utilizzo di interfacce vocali per applicazioni industriali in cui l'operatore deve collaborare con il robot ma il loro corretto funzionamento è disturbato da un ambiente di lavoro spesso rumoroso. In alcuni nostri lavori precedenti abbiamo esplorato la possibilità di riconoscere gesti utilizzando sensori indossabili come smartwatch o smartband [Carfi *et al.*, 2018]. L'utilizzo di sensori indossabili rispetto a telecamere ha il vantaggio di non richiedere un ambiente strutturato e l'operatore può comunicare con il robot anche in condizioni di luce non ottimali o nel caso avvengano occlusioni.

L'obiettivo del nostro lavoro è utilizzare tecnologia indossabile per rendere più accessibile la collaborazione tra uomo e robot in un ambiente industriale. Per farlo abbiamo progettato un'interfaccia che permette all'operatore di interagire con il robot tramite l'utilizzo di gesti. Tra le funzionalità presenti nella nostra interfaccia abbiamo incluso una versione di KT modificata per rendere l'interazione più naturale ed efficace.

2 Interfaccia Gestuale

La piattaforma robotica considerata nel nostro lavoro è Baxter, un robot collaborativo con due braccia. L'operatore per interagire con il sistema deve indossare uno smartwatch Android che comunica con il robot attraverso il Robot Operating System (ROS). Su Baxter è installata un'ancora che emette un segnale bluetooth, ricevuto dallo smartwatch, la cui intensità viene utilizzata per determinare la prossimità dell'operatore al robot. Questo permette al sistema di attivare l'interfaccia gestuale solo quando necessario e nel caso di un sistema con più di un robot, di controllare quello più vicino.

Baxter ha un monitor ad altezza uomo, che utilizziamo per visualizzare un'interfaccia grafica composta di menu e sotto menu che può essere navigata tramite l'utilizzo di gesti del braccio tracciato dallo smartwatch. Il set di gesti è preconfigurato ma può essere aggiornato per adattarsi alle esigenze dell'operatore, il riconoscimento avviene tramite l'utilizzo di SLOTH [Carfi *et al.*, 2018]. Il sistema riceve i dati dell'accelerazione lineare del polso e li processa utilizzando una rete neurale Long Short Term Memory (LSTM) che effettua una classificazione probabilistica. Il risultato del classificatore probabilistico è continuo e necessita di essere filtrato per evitare riconoscimenti multipli dello stesso gesto o falsi positivi. L'interfaccia fornisce all'operatore due funzionalità legate alla PbD: *Insegnamento*, grazie alla quale l'operatore può mostrare il nuovo compito, e *Riproduzione*, in cui il robot riproduce il comportamento che gli è stato insegnato. Per entrambe le funzionalità l'utente ha a disposizione una serie di comandi gestuali per controllare Baxter come: stop, pausa, torna indietro, rallenta, apri e chiudi l'end-effector. L'operatore può scegliere tra due varianti della funzionalità *Insegnamento*: *Autonomous KT* e *Classic KT*. L'ultima variante usa una versione classica del KT in cui il controllo dell'end-effector viene gestito dall'operatore tramite l'utilizzo di pulsanti presenti sul polso di Baxter o tramite comandi gestuali. *Autonomous KT* invece prevede che Baxter sia in grado di capire quando aprire e chiudere la pinza.

3 Autonomous KT

In [Carfi *et al.*, 2019] è stato evidenziato come la necessità di controllare la pinza nell'insegnare un'azione di pick-and-place ad un robot, può causare inefficienze nelle traiettorie generate. Inoltre, l'osservazione del KT tra umani ha evidenziato che le persone durante l'apprendimento non sono passive ma gestiscono autonomamente la propria mano cercando di prevedere cosa l'insegnante si aspetta. Questo ha ispirato uno studio successivo che ha portato a valutare, tramite l'utilizzo della tecnica sperimentale Wizard of Oz, se un comportamento autonomo, per quanto riguarda il controllo dell'end-effector, possa rendere più naturale la fase di insegnamento e minimizzare gli errori dell'operatore. Similmente all'interno della nostra interfaccia gestuale è stato implementato un comportamento bio-inspirato che permette a Baxter di essere attivo durante la fase di apprendimento aprendo e chiudendo la pinza. Una volta attivata l'*Autonomous KT* la modalità di Baxter di compensazione della gravità permetterà all'operatore di muoverlo manualmente esercitando una pressione sul polso. Il sistema acquisisce le informazioni di posizione e ve-

locità dei giunti insieme all'informazione inerziale del polso dell'operatore, fornita dallo smartwatch, e le processa utilizzando una rete neurale LSTM. La rete neurale è stata addestrata a predire quando aprire e chiudere la pinza utilizzando diverse ripetizioni di un task di pick-and-place. Per questo motivo l'*Autonomous KT* può essere utilizzata per insegnare nuove azioni di pick-and-place ma la sua generalizzazione ad azioni diverse richiede ulteriori studi.

4 Conclusioni

La nostra ricerca non punta a sostituire tecniche ben consolidate e ampiamente utilizzate anche in contesto industriale per il controllo e il monitoraggio della catena di produzione, ma punta a fornire interfacce alternative e complementari. L'introduzione di sensori indossabili nell'industria sembra un trend destinato a crescere soprattutto per tematiche riguardanti la sicurezza. Il nostro obiettivo è di integrare questa tecnologia con i sistemi standard già presenti a livello industriale per fornire all'operatore uno strumento in più per interagire e collaborare con il robot. Allo stesso tempo questa nuova tecnologia cerca di sfruttare la naturale propensione dell'uomo alla comunicazione gestuale in modo tale da aumentarne l'intuitività e non richiedere un addestramento specifico dell'operatore. Il nostro secondo obiettivo è appunto quello di rendere più facile ed intuitivo l'accesso a queste tecnologie. L'utilizzo di interfacce gestuali riduce il livello di specializzazione richiesto all'operatore per utilizzare il robot ma ulteriori studi devono essere condotti sull'ergonomia di questa interfaccia in modo tale da selezionare set di gesti facili da eseguire, da ricordare e che allo stesso tempo non affatichino troppo l'operatore. Infine la stessa introduzione dell'*Autonomous KT* promette di rendere l'addestramento dei robot più facile e più robusto all'errore umano, anche se ulteriori studi devono essere condotti per confermare questa ipotesi ed estendere questa tecnologia in altri contesti applicativi.

Riferimenti bibliografici

- [Carfi *et al.*, 2018] Alessandro Carfi, Carola Motolese, Barbara Bruno, e Fulvio Mastrogiovanni. Online human gesture recognition using recurrent neural networks and wearable sensors. Nanjing, China, August 2018.
- [Carfi *et al.*, 2019] Alessandro Carfi, Jessica Villalobos, Enrique Coronado, Barbara Bruno, e Fulvio Mastrogiovanni. Can human-inspired learning behaviour facilitate human-robot interaction? *submitted to the International Journal of Social Robotics*, 2019.
- [Kagermann *et al.*, 2013] Henning Kagermann, Wolfgang Wahlster, e Johannes Helbig. Recommendations for implementing the strategic initiative Industrie 4.0: Final report of the Industrie 4.0 working group. *Produktion, Automatisierung und Logistik*, 2013.
- [Lucke *et al.*, 2008] Dominik Lucke, Carmen Constantinescu, e Engelbert Westkämper. Smart Factory - A Step towards the Next Generation of Manufacturing. In *Manufacturing Systems and Technologies for the New Frontier: The 41st CIRP Conference on Manufacturing Systems*. Tokyo, Japan, May 2008.