

Workstation meta-collaborative basate su sistemi di visione e algoritmi intelligenti

Cristina Nuzzi, Simone Pasinetti, Giovanna Sansoni

Università degli Studi di Brescia, Dipartimento di Ingegneria Meccanica e Industriale
c.nuzzi@unibs.it, simone.pasinetti@unibs.it, giovanna.sansoni@unibs.it

Abstract

In questo contributo si presentano le caratteristiche di una workstation industriale meta-collaborativa, ideata per garantire la collaborazione uomo-robot a prescindere dalla presenza o meno di barriere fisiche tra le parti, in ottica di Industria 4.0. Per garantire ciò, il sistema proposto, realizzato in ROS, si basa sul canale di comunicazione visivo e adotta una comunicazione per mezzo di comandi gestuali. Lo sviluppo del sistema è ancora in corso, pertanto si presentano risultati parziali relativi al riconoscimento e alla traduzione del gesto, realizzato tramite l'object detector R-FCN.

1 Introduzione

Il paradigma **Industria 4.0** introduce il concetto di "smart factory", ovvero l'interconnessione tra sistemi informatici e sistemi fisici (sistemi *cyberfisici*). In particolare, questo lavoro si focalizza sul concetto di "smart production", basato su una concreta collaborazione tra macchine, strumenti ed operatori umani per mezzo di supporti informatici [Hermann *et al.*, 2016].

Negli ambienti industriali classici, le workstation tipiche sono di due tipologie: **completamente manuale**, ad esempio per operazioni di assemblaggio di piccole parti, o **completamente automatizzata**, dove il robot esegue in modo ripetuto e ad alta velocità il compito per il quale è stato programmato [Villani *et al.*, 2018]. Queste due tipologie di workstation presentano velocità produttive diverse, quindi sono di difficile integrazione tra loro. Una prima soluzione al problema è rappresentata dalle **workstation collaborative**, che permettano all'operatore di collaborare in modo più stretto con le macchine. Esse prevedono l'utilizzo di robot Collaborativi [Colgate *et al.*, 1996], progettati per garantire la sicurezza dell'operatore, e che possono portare a termine solo task semplici (movimentazione di oggetti, assemblaggio).

Una seconda soluzione è rappresentata dalle **workstation meta-collaborative**, dove lo spazio di lavoro dell'operatore e del robot non è necessariamente condiviso e la cooperazione può avvenire a prescindere dalla presenza di barriere fisiche tra i due. Per cooperare è necessario utilizzare un opportuno canale di comunicazione tra le parti, che deve avere le seguenti caratteristiche:

1. **sicurezza**: il canale deve poter garantire la sicurezza sia dell'umano che del robot;
2. **naturalezza**: il canale deve poter permettere all'operatore umano di comunicare in modo naturale con il robot;
3. **velocità**: il canale deve garantire latenza bassa, caratteristica critica sia per garantire una buona reattività alle condizioni di sicurezza, sia per rendere la comunicazione rapida.

Un canale di comunicazione efficiente e che soddisfa i requisiti illustrati è rappresentato dalla visione. Un adeguato set-up di videocamere (RGB, ToF) permette al robot di conoscere costantemente l'ambiente che lo circonda: ciò permette sia di adottare adeguate strategie di intervento per assicurare la sicurezza di robot e umani [Pasinetti *et al.*, 2018] [Polverini *et al.*, 2017], sia di riconoscere i movimenti e i gesti del corpo dell'operatore, metodo di comunicazione considerato naturale e intuitivo [Liu e Wang, 2018].

2 Sistema proposto

Il sistema da noi sviluppato si basa sul canale visivo (rappresentato da opportuni sistemi di visione) e sull'utilizzo di comandi gestuali per comandare il robot (Figura 1).

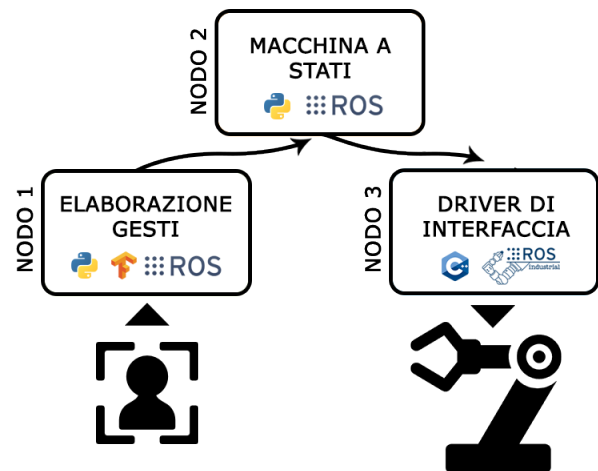


Figura 1: Schema a blocchi del funzionamento del sistema proposto.

La scena e in particolare l'operatore vengono ripresi da alcune telecamere. I frame acquisiti dalla telecamera preposta all'acquisizione dei gesti vengono analizzati dal **nodo di elaborazione gesti** in modo da estrarre il comando dato dall'operatore e tradurlo come un codice operativo per il sistema. Questo codice verrà inviato al **nodo della macchina a stati** che si occuperà di gestire a livello software l'operazione selezionata dall'utente e creerà il pacchetto dati operativo da inviare al **nodo del driver di interfaccia** tra ROS e il controllore del robot. Quest'ultimo è fondamentale per interfacciarsi al software proprietario del robot che si andrà ad utilizzare e, a seconda della casa produttrice della macchina, è necessario utilizzare un driver differente.

2.1 Riconoscimento gesti

A valle di uno studio sul linguaggio dei gesti per creare un linguaggio intuitivo e veloce da utilizzare per gli esseri umani, sono stati scelti 15 gesti statici per comporre le istruzioni del sistema. Ogni gesto a una mano può essere eseguito sia con la mano destra che con la mano sinistra mostrando il palmo alla telecamera, per un totale di 2 varianti. Il gesto "Horiz" è l'unico che presenta 4 varianti: con il palmo rivolto verso la telecamera e con il dorso rivolto verso la telecamera, sia per la mano destra che per la mano sinistra (Figura 2).

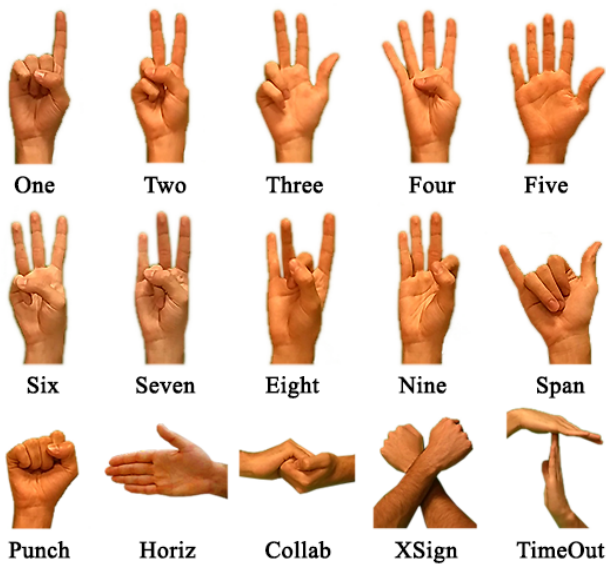


Figura 2: Gesti utilizzati, qui rappresentati eseguiti con la mano destra in posizione frontale.

Il dataset così composto è stato acquisito sperimentalmente, per un totale di 7219 immagini per il dataset di training (circa 250 immagini per gesto, 5 attori distinti) e 2804 immagini per il dataset di test (circa 100 immagini per gesto, 2 operatori distinti diversi dai precedenti). Ogni immagine contiene da uno a due gesti contemporaneamente. Il modello utilizzato è il R-FCN Object Detector [Dai *et al.*, 2016] pre-allenato, sul quale è stato eseguito un fine-tuning utilizzando i nostri dati. Le prestazioni del modello così ottenuto sono riassunte nella prima riga in Tabella 1 e in Tabella 2. Si noti che le occorrenze e le metriche sono calcolate sul numero

di gesti e non sul numero di immagini. I risultati mostrano come i gesti proposti per rappresentare le cifre oltre il 5 siano particolarmente difficili da riconoscere per il modello. Un secondo test allenato sugli stessi dati ma eliminando i gesti rappresentanti le cifre da 6 a 9 presenta invece prestazioni molto migliori, come evidenziato dai risultati nella seconda riga in Tabella 1 e in Tabella 2.

# Gesti	TP	TN	FP	FN
15	1818	0	556	700
11	3464	0	456	640

Tabella 1: Tabella delle occorrenze.

# Gesti	Recall	Precision	F1-Score	Accuracy
15	76.58%	72.20%	74.33%	59.14%
11	88.37%	84.41%	86.34%	75.96%

Tabella 2: Tabella delle metriche.

Riferimenti bibliografici

- [Colgate *et al.*, 1996] J. Edward Colgate, Witaya Wannasuphprasit, e Michael A. Peshkin. Cobots: Robots for collaboration with human operators. 1996.
- [Dai *et al.*, 2016] Jifeng Dai, Yi Li, Kaiming He, e Jian Sun. R-fcn: Object detection via region-based fully convolutional networks. *CoRR*, abs/1605.06409, 2016.
- [Hermann *et al.*, 2016] M. Hermann, T. Pentek, e B. Otto. Design principles for industrie 4.0 scenarios. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 3928–3937, Jan 2016.
- [Liu e Wang, 2018] Hongyi Liu e Lihui Wang. Gesture recognition for human-robot collaboration: A review. *International Journal of Industrial Ergonomics*, 68:355 – 367, 2018.
- [Pasinetti *et al.*, 2018] Simone Pasinetti, Cristina Nuzzi, Matteo Lancini, Giovanna Sansoni, Franco Docchio, e Alberto Fornaser. Development and characterization of a safety system for robotic cells based on multiple time of flight (tof) cameras and point cloud analysis. In *2018 Workshop on Metrology for Industry 4.0 and IoT*, pages 1–6, April 2018.
- [Polverini *et al.*, 2017] Matteo Parigi Polverini, Andrea Maria Zanchettin, e Paolo Rocco. A computationally efficient safety assessment for collaborative robotics applications. *Robotics and Computer-Integrated Manufacturing*, 46:25 – 37, 2017.
- [Villani *et al.*, 2018] Valeria Villani, Fabio Pini, Francesco Leali, e Cristian Secchi. Survey on human-robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics*, 55:248 – 266, 2018.