Università
degli Studi
della Campania
Luigi Vanvitelli

Dipartimento di Ingegneria

Dottorato di Ricerca in

Ingegneria Industriale e dell'Informazione

XXXVI Ciclo

Ph.D. Thesis

A Multimodal Workspace Monitoring and Control System for Human-Robot Collaborative Workcells

S.S.D. ING-INF/04 - Automatica

A.A. 2022/2023

Author
Gaetano Lettera
Matr. D129000084

Coordinator
Prof. Oronzio Manca

Advisor
Prof. Ciro Natale

The work presented in this thesis has been developed within the LABOR project, that has received funding from the European Commission under the Clean Sky's Horizon 2020 Research and Innovation Programme (grant agreement N° 785419 - LABOR).

A Multimodal Workspace Monitoring and Control System for Human-Robot Collaborative Workcells
by Gaetano Lettera

Sommario

Le celle di lavoro robotizzate autonome svolgono da molti anni un ruolo cruciale nella produzione industriale. Questi robot vengono adoperati per trasportare carichi pesanti ed eseguire una serie di compiti che richiedono estrema precisione, accuratezza e ripetibilità. Tuttavia, non sono in grado di adattarsi rapidamente ai cambiamenti del processo di lavoro oppure alle caratteristiche ambientali. Inoltre, secondo la normativa, i robot industriali devono operare all'interno di recinzioni di sicurezza volte ad impedire l'accesso dell'operatore umano alle aree pericolose. Negli ultimi anni i processi produttivi stanno richiedendo sempre più fortemente un cambiamento di paradigma, come discusso nei principi di progettazione incentrata sull'uomo dell'Industria 5.0. I cicli di vita brevi del prodotto e una forte personalizzazione del processo produttivo, infatti, necessitano di tecniche di produzione più flessibili: il risultato è che sono nati robot commerciali leggeri che consentono agli operatori umani di lavorare nel loro spazio di lavoro condiviso, combinando così i vantaggi di un robot con la flessibilità e la destrezza degli esseri umani, nonchè la loro capacità decisionale che sicuramente non può essere demandata ad una macchina. Tuttavia, oggigiorno all'interno delle aziende le applicazioni collaborative stanno avendo non poche difficoltà a sostituire i processi tradizionali, a scapito di potenziali ottimizzazioni qualitative di processo e di tempo ciclo, a causa delle normative di sicurezza in vigore che, essendo molto restrittive, comportano una bassa efficienza e tempi ciclo eccessivamente lunghi.

L'obiettivo di questo lavoro di tesi è presentare un nuovo algoritmo di collaborazione uomo-robot per applicazioni industriali che massimizzi le prestazioni del processo preservando la sicurezza umana. A tale scopo vengono analizzati i luoghi di lavoro di assemblaggio ed elaborati i vantaggi delle celle di lavoro collaborative. Passando in rassegna le attuali normative di sicurezza, si analizzano l'evoluzione e le sfide scientifiche proposte allo stato dell'arte e tecnologie all'avanguardia di nuovi sensori per lo sviluppo di algoritmi di collaborazione sicura uomo-robot. Particolare attenzione è stata posta sulla valutazione del rischio prima di presentare il nuovo metodo collaborativo che combina un efficace algoritmo di rilevamento del'operatore umano nel campo di lavoro condiviso con una strategia di controllo della veloctà del robot industriale che viene modulata in base ai risultati dell'analisi

del rischio a cui può essere sottoposto l'operatore, valutata in tempo reale. La tecnologia è conforme all'attuale standard di sicurezza ISO 10218-1/2.

La strategia di visione proposta per il monitoraggio dello spazio di lavoro del robot si basa su un sistema di percezione multimodale composto da una telecamera di profondità integrata con una telecamera termografica. Il sistema adotta algoritmi di intelligenza artificiale per identificare l'operatore all'interno della cella di lavoro. In particolare, una innovativa rete neurale convoluzionale è in grado di riconoscere i soli esseri umani presenti nella scena, dopo essere stata addestrata su immagini combinate con dati spaziali e dati termici. Questo approccio riduce al minimo i falsi positivi e quindi ottimizza la produttività della cella perchè evita fermi robot quando non sono necessari. Le informazioni elaborate dal sistema di visione vengono, poi, processate da un nuovo sistema di controllo basato su logica fuzzy che modula la velocità del robot a seconda della cosiddetta "distanza minima di separazione", cioè la minima distanza che garatisce l'applicazione della norma ISO/TS 15066 secondo il paradigma collaborativo *Speed and Separation Monitoring*. Questa distanza è il frutto di una valutazione del rischio eseguita in tempo reale che considera non solo la distanza effettiva uomo-robot ma anche una serie di condizioni euristiche che rappresentano le reali condizioni operative, per esempio il movimento predittivo del robot e dell'operatore, la distinzione fra le parti effettivamente appartenenti al corpo umano che possono impattare con la macchina rispetto ad eventuali altri strumenti di lavoro maneggiati dall'operatore e le reali velocità relative delle due parti coinvolte. Gli esperimenti su una cella di lavoro collaborativa industriale, infine, dimostrano l'efficacia dei metodi presentati in questo lavoro e tracciano nuovi orizzonti applicativi.

Abstract

Autonomous robotic workcells have played a crucial role in industrial manufacturing for many years. These robots are used to carry heavy loads and perform a variety of tasks that need extreme precision, accuracy and repeatability. However, they are unable to quickly adapt to changes in the work process or environmental characteristics. Furthermore, according to the actual regulations, industrial robots must operate within safety fences aimed at preventing human operator access to dangerous areas. In recent years, manufacturing processes are increasingly demanding a paradigm changing, as discussed in the human-centered design principles of Industry 5.0. The short life cycles of the product and a strong customization of the production process, in fact, require more flexible production techniques: the result is that light off-the-shelf robots have been born to allow human operators to work in their shared workspace, thus combining the advantages of a robot with the flexibility and dexterity of human beings, as well as their decision-making capabilities which certainly cannot be delegated to a machine. However, nowadays collaborative applications within companies are having lots of difficulties in replacing traditional processes, to the detriment of potential qualitative process and cycle time optimizations, due to the safety regulations which are very restrictive and entail low efficiency and excessively long cycle times.

The objective of this thesis work is to present a novel human-robot collaboration algorithm for industrial applications that maximizes process performance while preserving human safety. For this purpose, assembly workplaces are analyzed and the advantages of collaborative work cells are described. Reviewing current safety regulations, the evolution and scientific challenges proposed in the state-of-the-art have been presented as well as the most recent technologies of new sensors for the development of safe human-robot collaboration algorithms. Particular attention has been paid to risk assessment before presenting the new collaborative method which combines a robust algorithm for human detection in the shared workspace with an optimal control strategy for the speed of the industrial robot which is modulated by processing the results of the risk analysis to which the operator may be subjected, evaluated in real time. The technology complies

with the current safety standard ISO 10218-1/2.

The proposed vision strategy for monitoring the robot workspace is based on a multimodal perception system composed of a depth camera integrated with a thermographic camera. The system adopts artificial intelligence algorithms to identify the human operator within the workcell. In particular, an innovative convolutional neural network is able to robustly recognize the humans present in the scene, after being trained on images combined with spatial data and thermal data. This approach minimizes false positives and therefore optimizes cell productivity because it avoids robot stops when they are not necessary. The information elaborated by the vision system is then processed by a new control system based on a fuzzy logic which modulates the speed of the robot according to the so-called "minimum separation distance", i.e. the minimum distance that guarantees the application of the standard ISO/TS 15066 according to the collaborative *Speed and Separation Monitoring* paradigm. This distance is the result of a risk assessment performed in real time which considers not only the actual human-robot distance but also a series of heuristic conditions that represent the real operating conditions, i.g. the predictive direction of motion of the robot and the operator, the distinction between the parts actually belonging to the human body that can impact the machine compared to any other work tools handled by the operator and the real relative speeds of the two parts involved. Finally, experiments on an industrial collaborative workcell demonstrate the effectiveness of the methods presented in this work and outline new application horizons.

Ai miei genitori
e alle mie sorelle

Contents

Notation

The following acronyms have a special meaning:

AGVs	Automated Guided Vehicle
AI	Artificial Intelligence
AMR	Automated Mobile Robots
ANSI	American National Standards Institute
ADAS	Advanced Driver Assistance Systems
CCD	Charge-Coupled Device
CNN	Convolutional Neural Networks
DMU	Digital Mock-Up
DoF	Degrees of Freedom
DLT	Direct Linear Transformation
DT	Depth-Thermal
EAP	EtherCAT Automation Protocol
FIS	Fuzzy Inference System
FMEA	Failure Mode and Effects Analysis
FN	False Negatives
FP	False Positives
HAZOP	Hazard and operability study
HCAI	Human-Centered Artificial Intelligence
HDT	Human Detection and Tracking
HG	Hand-Guiding
HMI	Human Machine Interface
HOG	Histogram of Oriented Gradients
HR	Human-Robot
HRC	Human-Robot Collaboration
HRI	Human-Robot Interaction
IE	Intention Estimation
IEC	International Electrotechnical Commission
IoU	Intersection Over Union
IRM	Institute of Robotics and Mechatronics
ISO	International Organization for Standardization

LABOR	Lean robotized AssemBly and cOntrol of composite aeRostructures
LKF	Linear Kalman Filter
LWR	Light Weight Robots
maP	Mean Average Precision
MSAC	M-estimator SAmple Consensus
NaN	Not-a-Number
PCD	Point-Cloud Data
PFL	Power and Force Limiting
PHA	Preliminary Hazard Analysis
PL	Performance Level
PLr	Required Performance Level
PLS	Programmable Limit Switch
PMD	Photonic Mixed Device
PMLA	Passive Mechanical Lower Arm
PPE	Personal Protection Equipment
RANSAC	RANdom SAmple Consensus
RCA	Root Cause Analysis
RGB-MA	RGB Mapping Approach
RIA	Robotic Industries Association
ROI	Region Of Interest
ROS	Robot Operating System
SIL	Safety Integrity Level
SCAT	Systematic Cause Analysis Technique
SMS	Safety-rated Monitored Stop
SRP/CS	Safeguarding and safety-Related Part of a Control System
SSLs	SphereSwept Lines
SSM	Speed and Separation Monitoring
SVM	Support Vector Machine
SoA	State-of-Art
ToF	Time of Flight
TRL	Technology Readiness Level
TS	Technical Specification
TUSL	Testo Unico per la Sicurezza sul Lavoro
UWB	Ultra-WideBand
VIA	Variable Impedance Actuator
YOLO	You Only Look Once

1. Introduction

1.1. Background and motivations

The importance of discussing human-robot collaboration (HRC) in the field of industrial robotics has grown over the past few years, especially in Industry 4.0 and Industry 5.0. Figure 1.1 shows the types of industrial revolutions, from 1.0 to 5.0, with dates and details on the research topics of each one.

The main focus of Industry 4.0 was the integration of mechanical, electronic and digital technologies for machine interconnection and interaction through a continuous flow of data. Robotics in Industry 4.0 has already constituted one of the enabling technologies of the I4.0 smart factory, in terms of flexibility and reconfigurability of processes, integration with the digital factory system and also interaction with the operator. However, although HRC was a trend related to Industry 4.0, new robots have been exploited more for other characteristics (flexibility, ease of programming). This is why the topic has been then taken up again in Industry 5.0, in which the main goal of this new trend is to improve the safety, ergonomics, productivity and quality in the field of industrial robotics. "Industry 5.0 complements the existing Industry 4.0 paradigm by highlighting research and innovation as drivers for a transition to a sustainable, human-centric and resilient European industry. It moves focus from share-holder to stakeholder value, with benefits for all concerned. Industry 5.0 attempts to capture the value of new technologies, providing prosperity beyond jobs and growth, while respecting planetary boundaries, and placing the wellbeing of the industry worker at the centre of the production process." [1].

The current industrial research topic aims to fill the gap between manual manufacture and fully automated production, which means human shares the workspace with the robot which assists him by executing hazardous, repetitive, and non-ergonomic tasks. According to this point of view, industrial processes obtain benefits both from people and robots. While the human operator increases the process adaptability and intelligence, robots are efficient, fast, durable and highly accurate machines and they can per-

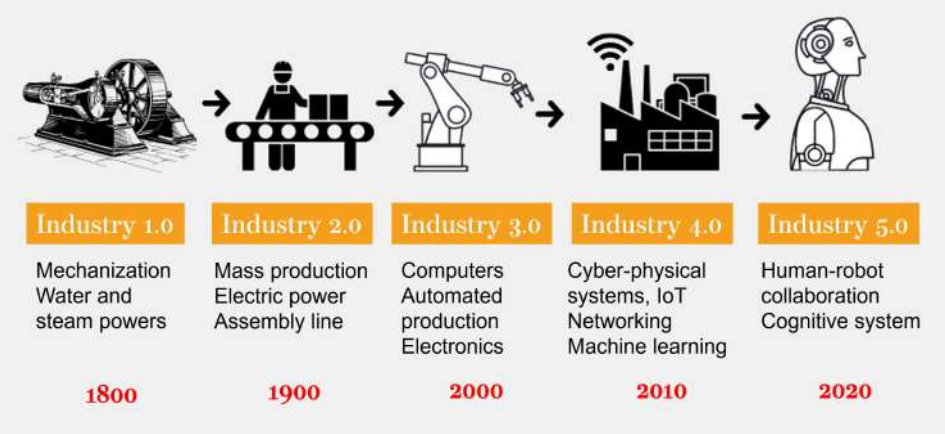


Figure 1.1.: Types of industrial revolutions: dates and details of the research topics from 1.0 to 5.0.

form tasks more quickly, more effectively and more affordably than people [2]. On the other hand, some duties need operators able to think and must be adapted to the situation at hand and can not be delegated to a machine: robots are not capable of thinking, they just execute commands and accomplish pre-learned movements. In other words, whereas humans are more flexible —for instance, the upper limb of a human body has thirty DoFs— robots are created with six or seven DoFs and they are constrained by their predetermined programming.

HRC breaks through these barriers and benefits from the advantages of robot to succeed in challenging applications where operators are still needed [3]. The efficiency and productivity of the process have always represented significant challenges in production lines in which all tasks were initially delegated to the human operator: since the first industrial revolution the use of mechanics, and then of robotic technology, has been a strong support for improving these requirements, which is why many solutions have been proposed and are still the subject of regulatory discussion today [4]. As discussed, the ability to give the operator more flexibility to complete tasks with fewer payload objects helps him to focus more on tasks that require higher levels of human capabilities. However, this collaboration may be extremely dangerous due to possible unpredictable, wrong motion of the robot which can cause irreversible injuries to the operators [5].

1.2. Research objectives and market relevance

Generally, manufacturing assembly processes are dominated by high-precision, high-throughput, and heavy-payload robots integrated into the production lines. According to the actual safety regulations for industrial robots, the working area is enclosed by fences to prevent the human access to the danger areas. Especially in the automotive and aerospace sector, the assembly process is often composed of very complex sub-processes that require human intervention, such as wiring, harnesses, seals, limp components, dismantling and consequent deburring of the metal parts [6]. So to avoid interrupting the automatic process of the system and therefore stopping the robots, thus lengthening production times, it would be desirable to implement HRC strategies to allow humans to carry out these operations while robots continue their work cycle.

In recent years, robotics hardware has advanced significantly to enable safe human-robot interaction (HRI). Innovative light-weight technology enables minor human-robot collisions while preventing against serious harm brought on by robots, the so-called *cobots*, robotic assistants or cooperative robots. They are made to work alongside people, they are typically small and safe robots, they do not require a design that is significantly different from the industrial robots used in typical workplaces but they already meet ISO EN 10218 safety standards [7, 8] though the integration of safety components. As a result, many companies are interested in using collaborative robotic systems, which allow humans and robots to work side by side, improving the working conditions and the process productivity.

Unfortunately, no great breakthrough occurred and only a few cooperative applications were developed for industrial environment. This is a result of the strict safety requirements in industrial fence-less applications. The safety constraints significantly reduce the efficiency of *cobots*, making them inapplicable for the most of industrial assembly processes. More in detail, international guidelines state that if the robot moves objects with weight up to several tons or moves pieces with acceleration up to 98 m/s^2, it must be adequately secured, so it is enclosed in fences and marked with warning icons. Robots have to immediately stop their activity and motions whenever their operating area is violated in order to avoid collisions, damage, or fatal injuries. To employ HRC in a production line, a few variables, such as robot payload and velocity, must be readjusted: the standard load capacity of a collaborative robot is in the range of 10 kg, and its maximum speed is often limited to 250 mm/s. These restrictions, however, are not enough to totally prevent the collision, so the *cobot* should be further protected with detecting

sensors or other collision avoidance techniques and, most of all, when the task requires both precisions or loads particularly high and specific human dexterity and decision-making capabilities, *cobots* become unusable due to these limitations.

The solution is the use of an industrial robot equipped with a HRC system compliant with the current international standards which saves the process productivity, optimizing the lead-time while preserving human safety. This is the industrial research objective of this thesis:

Maximizing industrial robotic workcell performances while preserving human safety.

This thesis is motivated by the lack of efficiency in industrial collaborative applications. State-of-Art (SoA) research predominantly focused on the robot system, developing human-robot collision avoidance path re-planning methods. This assumption is often not applicable for assembly processes where it is necessary to follow a predefined path that is the same as the traditional non-collaborative process. Therefore, this thesis is motivated to transfer the attention on the safety of the human operator who enters the collaborative cell. The innovative HRC system proposes a new collaborative workspace monitoring method and a new robot speed control algorithm based on human-in-the-loop considerations. The basic idea is to keep unchanged the traditional robotic process but integrate it with an additional HRC module to transform the standard process into a collaborative process, obtaining maximum performance.

The reference markets are multiple, ranging from the manufacturing industry -e.g. aerospace, automotive, household appliances- with collaborative assembly applications but also collaborative transport, maintenance and inspection, quality control, up to logistics. In automotive sector, in particular, the technology can find large application for pedestrians recognition in advanced driver assistance systems, the so-called *ADAS*. Finally, the services market could also invest in this technology, such as systems integration companies and surveillance and monitoring companies.

1.3. Contribution to robotics research

The work contributes to the robotic industrial research as follows:

- **Application and market trends studies**

The preliminary work is the study of existing supplier trends and consumer needs which highlights the variety of context in which industrial HRC can be employed to obtain process advantages (Chapter 2). In an on-site analysis, industrial work place and an industrial assembly application are analyzed to identify the risk assessment of a HRC task (Appendix B) and a suitable collaborative robotic workspace design has been discussed (Chapter 3) for the industrial research objective.

- **Development of a novel multimodal perception system**

  The pros and cons of current vision systems for robot workspace monitoring and human detection are analyzed in Chapter 4. An innovative solution which combines depth data with thermographic data is developed to decrease false positives during the human worker detection, with the objective of increasing the workcell productivity due to the reduction of robot downtime.

- **Risk assessment evaluation under human-in-the-loop analysis**

  In Chapter 5, a *Speed and Separation Distance* collaborative scenario is analyzed for the workcell design. A problem considering human-in-the-loop analysis is introduced: the calculation of the separation distance actually belongs to the operator's body and behaviors and the system's ability to identify even multiple subjects at the same time. The thermal point-cloud of the human clusters is analyzed in real-time to predict human and robot directions of motions in order to execute a risk assessment analysis that is faithful to the real scenario.

- **Fuzzy control logic for robot motion**

  At the robot control level, described in Chapter 6, the peculiarities of the technology consist in the evaluation of the safety distance based on the real speeds of the human and of the robot, the real-time risk analysis based on the actual operating conditions of the cell and, finally, a fluid modulation of the robot speed to allow greater acceptability of the machine by the operator.

- **Development of a workcell for HRC**

  Originating from the research work review, the innovative technology is integrated into the industrial robotic workcell LABOR to build an industrial research platform for a future scientific work in the field of HRC. LABOR adopts medium-size robots to perform both the assembly and the inspection of regional aircraft panels in a lean system.

After the drilling operation, some hybrid metal and composite structures have to be manually removed for further manual operations, like deburring, and then re-installed on the skin panel before the sealing and riveting operations. In this view, the workcell has been used as a demonstrator for the HRC experiments, as discussed in Chapter 7.

The author of this work is also the co-author of several publications, most of which are included in this thesis. Therefore, graphics, texts, contents, etc. have been used, which are cited whenever applicable.

1. Costanzo, M.; De Maria, G.; Lettera, G.; Natale, C. and Pirozzi, S., *Flexible Motion Planning for Object Manipulation in Cluttered Scenes*, 2019 In Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics - Volume 2: ICINCO; ISBN 978-989-758-321-6; ISSN 2184-2809, SciTePress, pages 110-121, doi: 10.5220/0006848701100121

2. Costanzo, M., De Maria, G., Lettera, G., Natale, C., Pirozzi, S., *Motion Planning and Reactive Control Algorithms for Object Manipulation in Uncertain Conditions*, 2018 In Robotics (Vol. 7, Issue 4, p. 76). MDPI AG, doi:10.3390/robotics7040076

3. Campomaggiore, A., Costanzo, M., Lettera, G., Natale, C, *A Fuzzy Inference Approach to Control Robot Speed in Human-robot Shared Workspaces*, 2019 In Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics - Volume 2: ICINCO, pages 78-87, 2019. INSTICC, SciTePress, doi: 10.5220/0007838700780087

4. M. Costanzo, G. De Maria, G. Lettera, C. Natale and D. Perrone, *A Multimodal Perception System for Detection of Human Operators in Robotic Work Cells*, 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, 2019, pp. 692-699, doi: 10.1109/SMC.2019.8914519.

5. M. Costanzo, G. De Maria, G. Lettera and C. Natale, *Grasp Control for Enhancing Dexterity of Parallel Grippers*, 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 2020, pp. 524-530, doi: 10.1109/ICRA40945.2020.9196873.

6. M. Costanzo, G. De Maria, G. Lettera and C. Natale, *Can Robots Refill a Supermarket Shelf?: Motion Planning and Grasp Control*, 2021 in IEEE Robotics and Automation Magazine, vol. 28, no. 2, pp.

61-73, June 2021, doi: 10.1109/MRA.2021.3064754.

7. Caterino, M., Chiacchio, P., Cristalli, C., Fera, M., Lettera, G., Natale, C., and Nisi, M., *Robotized assembly and inspection of composite fuselage panels: the LABOR project approach*, 2021 In IOP Conference Series: Materials Science and Engineering, Vol. 1024, Issue 1, p. 012019, IOP Publishing. https://doi.org/10.1088/1757-899x/1024/1/012019

8. M. Costanzo, G. De Maria, G. Lettera and C. Natale, *A Multimodal Approach to Human Safety in Collaborative Robotic Workcells*, in IEEE Transactions on Automation Science and Engineering, vol. 19, no. 2, pp. 1202-1216, April 2022, doi: 10.1109/TASE.2020.3043286.

9. Lettera, G.; Natale, C. *An Integrated Architecture for Robotic Assembly and Inspection of a Composite Fuselage Panel with an Industry 5.0 Perspective*, Machines 2024, 12, 103. https://doi.org/10.3390/machines12020103

2. Evolution of HRC

The importance of discussing HRC has grown over the past few years. Leading robot manufacturers have begun to create robots for workplaces that are accessible to human operators in order to combine the productivity of robots and the flexibility of humans: collaborative production has immense promise. Applications for collaboration range from simple coexistence to fully interactive industrial operations.

2.1. Terminology for interactive applications

Especially in the last two decades, much focus has been placed on collaborative robotics and HRC has been introduced in very different context [9]. Industrial robots are mostly used in manufacturing to carry out monotonous, risky, and tiring repetitive activities. They are typically installed in physically fenced area to prevent direct human-robot interaction. Contrary to these typical scenarios, the most recent industrial robotics research is increasingly focusing on how humans and robots may work together [10].

Schmidtler et al. [11] introduces the term *interaction* as a general term to represent every kind of interactions between humans and robots. However, Chandrasekaran et al. [12] and Goodrich et al. [13] distinguished *interaction* from *collaboration*: HRI has been defined as a situation in which many agents (humans and robots) react or communicate with each other to accomplish a common work task, while for closer physical and cognitive interaction it is necessary to extend the working system to HRC [14] which includes the extension of the degree of autonomy of the robot and the allowance of spatial proximity between the human and the robot during the operation [9]. This particularly requires advances in interactive and adaptive safety devices to guarantee human integrity.

The following Figure 2.1 clarifies the possible interaction types to illustrate the increasing demands on safety devices, based on the definitions contained in [4].

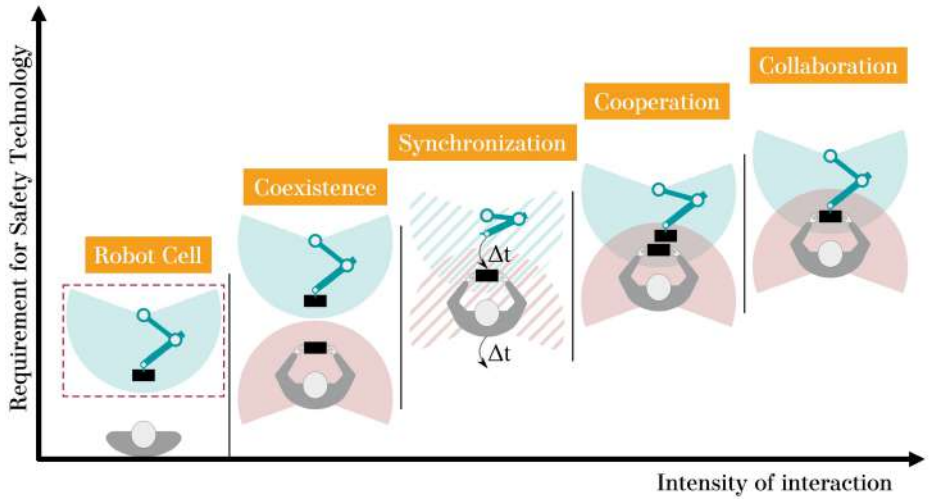


Figure 2.1.: Types of human-robot interactive scenarios with increasing safety technology requirements: from the robot cell to the collaborating shared workspace.

Robot cell
The installation of protective fences in the physical working area separation physically forbids human and robot contact during the robot motion. The robot moves with full power in absence of humans, and it is stopped in presence of them. Human access to the robot working area is handled by peripheral safety functions or systems, such as emergency buttons installed at the robot workspace entry. The robot is stopped as soon as the human operator requires to cross the fences to enter the working area.

Coexistence
In a virtual working area separation, also called coexistence or coaction, no real fences are installed because they are replaced by light barriers, ground sensors or vision-based safety devices which monitor the robot working area and ensure that no people are present. Behrens et al. [15] envision no sharing of the workspace between humans and robots, and no common task and contact, nor the coordination of actions and intentions.

Synchronization
In a syncronized workspace, both the human operator and the robot use the working area simultaneously and the two actors work on the same task. However, the work in the overlapping area, the so-called *collaboration space*, takes place with a time delay (temporal separation). Physical contact is not intended but possible.

Cooperation
In cooperation, the robot and the human operator work simultaneously on a common goal in a shared interactive workplace [16]. This means that they collaborate closely on the same task and a frequent contact between each other is permitted. Every kind of dangerous operations have to be avoided, thus this interaction context requires advanced safety devices such as advanced vision systems, force sensors and complex sensing for collision detection [17].

Collaboration
Collaboration is defined as a joint execution of a complex work task with direct interaction between humans and robots [18]. In collaboration, humans and robots work simultaneously on the same workpiece. There are two types of collaborative contexts:

- Physical collaboration: there is explicit and intentional physical contact between humans and robots in terms of force exchange [9, 19, 12].
- Non-contact collaboration: no physical interaction takes place but the actions are coordinated through information exchanged via direct communication (speech, gestures, etc.) [20, 21, 22] or indirect communication (recognition of intentions, gaze, facial expressions, etc.) [23, 24]. Usually, the human performs tasks that require dexterity or decision-making competence, while the robot takes over tasks such as repetitive, precise, dangerous, or force requiring applications [25].

The last one scenario describes the most common interest in collaborative robotics and is mainly discussed in this industrial research work.

2.2. Types of collaborative operations

The basic safety requirements on the robot and the robot system are described in the standards for the safety of the robot ISO 10218-1, and for the safety of the robot system ISO 10218-2. See Appendix A for more details. They also specify four basic protective principles for HRC, as shown in Figure 2.2. Each kind of collaborative operation deploys special technologies and methods for maintaining a safe operating environment. Types of collaborative operations are divided into the following levels:

- Level 1: Safety-rated Monitored Stop
- Level 2: Hand-Guiding

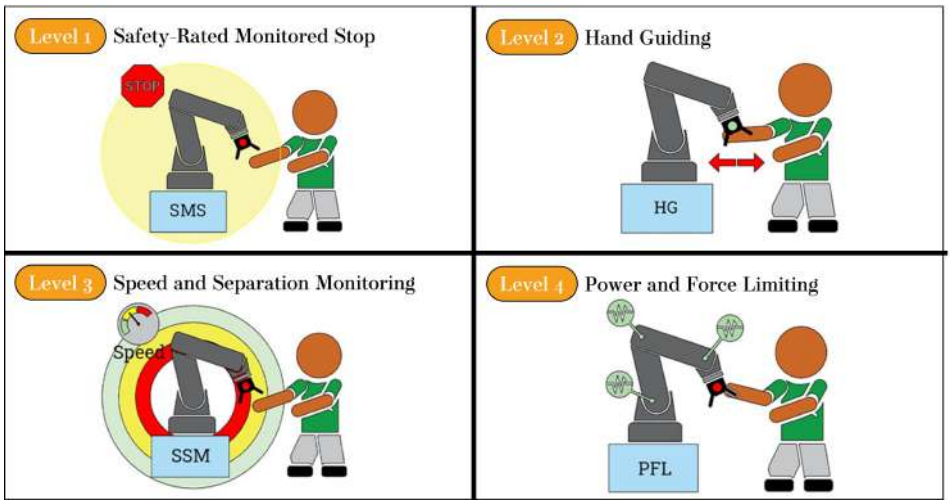


Figure 2.2.: The four levels of collaborative operations described in ISO 10218.

- Level 3: Speed and Separation Monitoring
- Level 4: Power and Force Limiting

From Level 1 type of standards to Level 4, the interaction level is increasing, and the operator and the robot are close to each other, thus the HRC system need to be more sophisticated. So for the different integration level, the corresponding safeguard method is different from each other.

2.2.1. Level 1: Safety-rated Monitored Stop

The Safety-rated Monitored Stop (SMS) is the simplest form of HRC: the robot executes its task into a supervised area in the collaborative workspace, the space shared by the human operator and the robot. During the execution of the robot task, the worker can enter the collaborative workspace and carry out his task, such as placing/removing parts into/from the end effector. When he completes this task, the worker leaves the collaborative workspace and the robot can resume the non-collaborative work automatically. In other words, human and robot share the same workspace, but the robot does not move while the human is present.

These type of applications uses a traditional industrial robot attached with different safety sensors to sense when a human enters into the robot workspace. The SMS collaborative scenario is largely used for applications which will

have the least human interaction with the robot, i.g. when the operator has to manually insert objects to the robot end effector or for visual inspection that may be required during a lot of robotic operations or for some operations which may require human involvement, such as finishing or automating complex processes that are expansive.

2.2.2. Level 2: Hand-Guiding

HRC scenario according to hand guiding (HG) puts the worker into direct control of robot motion. The mitigation of the risks associated with robot motion is to effect robot motion only as a result of dedicated input from the operator. This input is given for example via a joystick or by means of direct input of external force onto a compliant robot, which then moves accordingly. Motion can also be constrained by programmed limits, such as "virtual walls", fixed trajectories, and speed limits or similar, as the application risk assessment might require. For improving the safety, there is an enable button (death switch) in the grabbing area.

This collaborative scenario is frequently utilized to easily and quickly program new cobot paths and positions using a hand. It's perfect for mobile applications where the robots travel between stations and need retraining to complete different jobs, and when the robot has to be reprogrammed regularly for a new task. Applications of hand guiding method are as following: robotic lift assist, highly variable applications (acts like a manually "tool") or limited and small-batch production.

2.2.3. Level 3: Speed and Separation Monitoring

The speed and separation monitoring (SSM) collaborative scenario allows the human worker and the robot to move in the same workspace in such a way during which the robot speed and distance between the robot and the human is actively adjusted to be able to stop any hazardous motions of the robot before a contact with the human occur.

It is clear that this generates the need for safety-related detection of the worker in the collaborative work space by properly sensors. Safety laser range scanners are generally adopted to set up areas around the robot workspace. When a person gets close to the robot, the sensor process the situation and this information is used by the robot controller to adjust its speed to avoid moving contact with the worker. This means that the robot motion slow down and then halt if that human gets closer to

the robot workspace. As soon as the human operator moves away from the robot workspace, the collaborative robot automatically reverts to the normal speed when the protective separation distance is less then the real separation distance between the two actors. Thus, online estimation and computation of the separation distance is needed.

The case study which will be presented in Chapter 3 is based on SSM mode.

2.2.4. Level 4: Power and Force Limiting

Power and Force Limiting (PFL) collaborative robots are what many people think when they hear the name collaborative robots, the *cobots*. These robot are designed to enable direct collaboration with human employees without needing extra safety barriers, external scanners, or vision systems. Further, improving human safety, PFL robots are designed without exposed motors, sharp corners or pinch points.

The main design characteristic is that they are provided of sensitive built-in collision monitors. The robot motion parameters are monitored with high precision and even the tiny deviation between the robot and the actual position can be detected compared with the program control. The high precision encoder and high resolution allow the robot to accurately monitor its speed and position. By analyzing the current generated by the actuators and by measuring the response to the ground or by using tactile and torque sensors in the robot joints, the robot can recognize the impact of obstacles, analyze and react to them in a very short time. The robot can brake and stop immediately after a collision, or it can move in the opposite direction, minimizing the impact energy.

The main principle of PFL collaborative scenario is that human and robot are working so closely that incidental impacts can occur, but the mechanical and control architecture of the robot system are designed to make such contacts "harmless". Therefore, it is important to underline that to apply PFL requires that the robot system should be specifically designed.

Note that the proper biomechanical limits for harmless contacts, both for sustained and for short contact, are provided in the ISO/TS 15066 [4], also distinguishing quasi-static (pressure) or transient (dynamic) contacts. These limit values are based on pain sensation threshold research, a literature study on the estimated onset of minor superficial injury and a simple model-based interpretation.

PFL collaborative robots are presently limited to small or highly variable

applications and conditions requiring frequent operator presence.

2.3. A literature review

The safety collaborative scenarios can be divided into two categories: post-collision and pre-collision [26]. A post-collision system reacts after the physical impact occurs between the robot and the human operator. Three main drawbacks can be highlighted:

1. a collision could be dangerous if the robot limits are poorly defined;
2. any collision will halt task execution, leading to the decrease of the production time;
3. especially in industrial applications, the robot is equipped with sharp tools, e.g. drilling tools, whose impact can cause serious injuries.

Human safety can be assured by minimizing the energy transmitted during the contact [27], [28] by using robots endowed of compliant sensors for assessing force exchange when the impact occurs, e.g. force or tactile sensors [29], or by using an industrial manipulator having a closed control architecture and no additional sensors [30]. On the other hand, a pre-collision scheme makes use of exteroceptive sensors to detect humans and prevent collisions. Motion capture systems combined with range sensors [31] or artificial vision systems [18] are crucial in the case of distance monitoring, which is the most suitable approach for pure coexistence in a collaborative workspace. The dangerous zone around the robot is monitored and any operator that accesses it, makes the robot slowing down, until the full stop when the human is too close.

In order to provide a structured framework about the evolution of HRC, a classification of the main safety systems proposed in literature has been provided in Figure 2.3. The following sections describe the main hardware and software systems that are employed to solve three macro-objectives:

1. the quantification of the level of injury caused by collision between a human and a robot;
2. the minimization of the injury in human-robot collisions through specific devices and actions;
3. collision avoidance for safe pre-collision strategies.

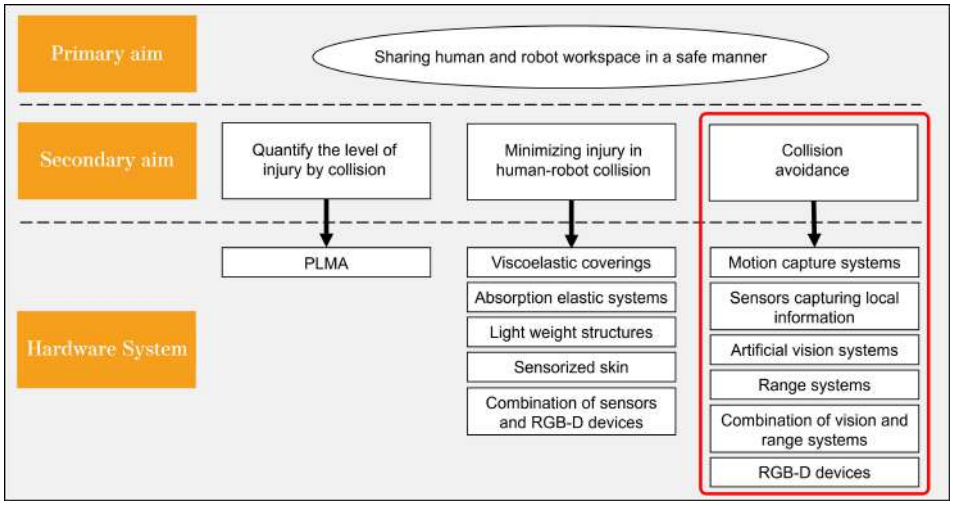


Figure 2.3.: A framework for classifying SoA Human-Robot research works with proposed hardware safety systems.

Note that Section 2.3.3 contains a general overview of some pre-collision strategies but it will be taken up in Chapter 4, in which the sensors proposed in literature for HRC collision avoidance algorithms (cointained into the red rectangle of Figure 2.3) will be largely explored with the aim of supplying a comparison with the multimodal vision system proposed by this thesis work.

2.3.1. Quantify the level of injury by collision

The information about the quantification level of injury caused by robot collision with the human body is a starting point to take the necessary steps to minimize risks and can be used for testing new robot safety systems.

To imitate relevant human characteristics, a mechanical device called Passive Mechanical Lower Arm (PMLA) has been built and proposed to be used to measure human pain during collisions with an industrial robot [32, 33]. Indications of pain was evaluated using perceived pain of volunteers and correlated measurements obtained by PMLA. The authors found that the *impact energy density*, which is a function of the values of impact force, the contact surface area, and distance between the robot end effector and the center of the PMLA, correlated well with the perceived pain.

2.3.2. Minimizing injury in human-robot collision

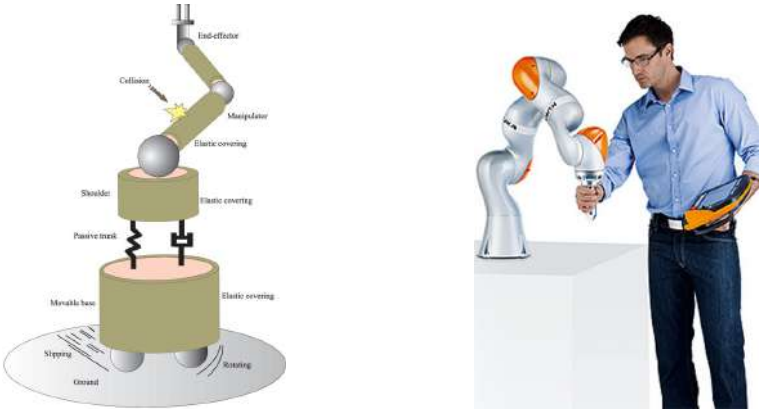
As in some cases a robot-human collision during the execution of collaborative tasks can be unavoidable, an important line of research focuses on the minimization of injuries in humans caused by such collisions. The proposed methods that reduce the effects of collisions can be classified in two categories: mechanical compliance systems, which aim at reducing the collision energy, according to [34], and safety strategies involving collision/contact detection, as shown in Figure 2.3.

Viscoelastic coverings

In [35], the robot was equipped with torque sensing and links which are covered with a viscoelastic material (e.g. rubber or urethane foam): the cover aims at reducing the impact force whilst maintaining contact sensitivity. The authors found that although the use of soft compliant rubbers or urethane foam allows to mitigate collision forces, the achieved reduction is not sufficient to ensure that pain tolerance thresholds are not exceeded. Viscoelastic coatings are also used as a suitable component for contact force reduction in [36], along with a deformable trunk consisting of springs and dampers, which is located between a fixed base and the robotic arm, see Figure 2.4a. This type of passive redundant DoF in an *absorption elastic system* helps prevent excessively strong impacts but it is recommended to use a light robot with low inertia in conjunction with a soft surface because the majority of conventional industrial robots are challenging to use in human-robot collaboration due to their weight and significant inertia. The evolution of viscoelastic covering has been treated in [37], where a novel tactile skin covering has been made: a low-noise capacitive force sensing array embedded in the soft foam that covers the whole body of a manipulator was proposed. The use of these sensorized skin sensors is treated in next subsection as a way to obtain information about intentional or unintentional human-robot contact to be used to trigger a suitable collision reaction strategy.

Absorption elastic systems

Another type of mechanical compliance system is mechanical absorption elastic systems in industrial robots, which are able to absorb part of the energy in a human-robot impact but not enough to effectively reduce injuries to the human. Traditional industrial robots normally introduce stiff actuators with a high impedance in order to guarantee a high accuracy and



(a) (b)

Figure 2.4.: Mechanical compliance systems: from the viscoelastic coverings of the conceptual model of a Human Friendly Robot HFR, [36](a) to the light weight robot KUKA LBR iiwa (b).

precision. But for collaborative operations a more compliant system is necessary to reduce potential impact force, e.g a Variable Impedance Actuator (VIA). [38] describes non-VIA (traditional stiff actuator) as a device able to track a certain trajectory and once a desired position is reached it will strive to hold this position. The research work highlights that a VIA, otherwise, deviates from its set equilibrium position, i.e. where the actuator generates zero force and torque, depending on the external forces and the mechanical properties of the actuator. The main classes of different VIA are active impedance by control, inherent compliance and damping actuators and inertial.

Light weight structures

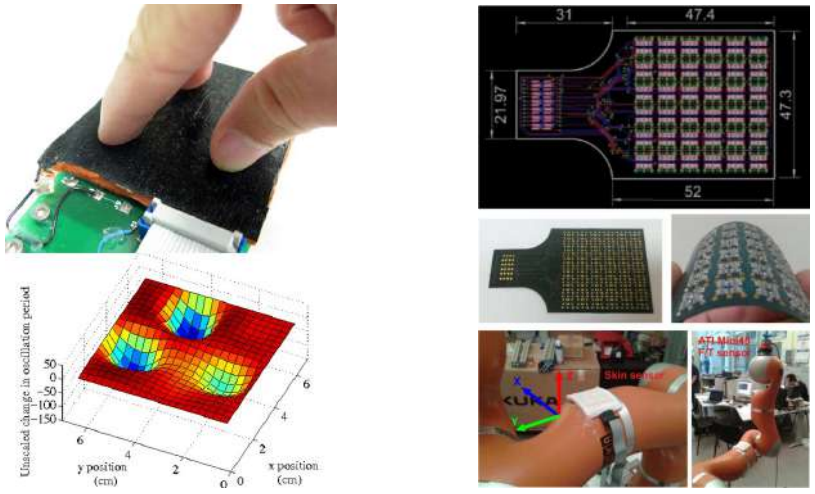
Finally, also light weight structures represents another type of mechanical compliance system which are largely used. In the 1990s the first generation of Light Weight Robots (LWR) was presented by the Institute of Robotics and Mechatronics (IRM) at DLR. The design concept and first steps towards a light weight robot generation were presented in [39] and [40], where the goal was to design a multi-sensory, modularly configurable light weight robot in a unified and integrated way. Later, control approaches were reported till the second generation of lightweight robots [41, 42]. The third generation LWRIII was presented in 2002. After two years, in 2004, DLR transferred this technology to the KUKA Robot Group [43], where several robot generations were developed, from KUKA LBR3 to LBR iiwa (Figure 2.4b).

With the development of systems for mechanical compliance, the use of lightweight materials such as light carbon fibres [44], and the use of sensor skin based on capacitive sensing developed by MRK-Systeme for Kuka robots or the capacitive skin developed by Boch for the APAS robot, nowadays they are suitable for collaborative human-robot tasks. These robots have been the basis of many investigations into the discussed quantification of the level of injury by collision, minimization of injury in human-robot collision and development of safety strategies for collision/contact detection.

Sensorized skin

Some mechanical devices have been developed as safety measures involving collision detection for HRC to minimize injury in collision. As discussed in previous subsection, teh first sensorized skin was developed as the evolution of viscoelastic coverings in [37]. A grid of nxn capacitive sensors was created and through a quad NAND chip, row-column addressing was implemented, as shown in Figure 2.5a. 100 sensors can be scanned continuously with a 1 kHz update rate with a total of 5 interface wires. The sensor was constructed from physically resilient materials, but is still soft enough to allow energy absorption in case of a collision. The idea of mounting the sensorized skin on the robot link surface to obtain information from the impact point and to reduce robot velocity when indispensable, is currently under analysis. In [29] a conformable sensor skin made of optoelectronic components mounted on a flexible PCB was presented. This sensorized skin has the capability of measuring not only the position of the contact point but also the three components of the applied force, and thus it allows to estimate the force vector of the impact. The results of experimental tests on a light robot show a suitable behaviour of the robot in intentional and unintentional human-robot contact situations, with a controlled human-robot interaction and a safe reaction in case of accidental contacts being guaranteed, as shown in Figure 2.5b.

Recently, in [45] has been presented a research work describing the development of another flexible sensor skin, made of a conductive fabric doped with Carbon Nanotubes. The authors validated the prototype through a series of experiments which characterized the skin in terms of applied force and positioning of applied force: the sensor error is within 2.5 ± 1.5 mm, even during applied stretch (max 150%). The skin also proved to absorb energy from impacts, while the accuracy is below one centimeter, which is enough for most collaborative tasks. In a collaborative control demonstration the sensorized skin maintained a collaboration force of 10 N and for



(a) (b)

Figure 2.5.: Main SoA proposed sensorized skin: tactile display from the capacitive array prototype, [37] (a) and the flexible PCB, [29] (b).

an emergency stop application the skin was consistently reacting to 0.5 N applied force.

Combination of sensors and RGB-D devices

A unified and very interesting approach for safety in HRC was presented in [46]: the research develop an algorithm not only for collision detection but also for robot reaction out of the collision zone, while allowing for intentional physical interaction. The basis of this approach is the application of residual signals as suggested in [47]. Additionally, the system has a collision avoidance approach that makes use of an RGB-D sensor. Moreover, the system uses a safe collaboration phase that combines the residual method for collision detection with the localization of the contact point obtained from the RGB-D image to enable both intentional physical contact between human and robot as well as contactless interaction via voice and gesture recognition. In later work, Magrini et al. [48] present motion and force control at any point of contact on the robot and offer generalizations of classical impedance and direct force control schemes that don't require a force sensor to operate—instead relying on a quick estimate of contact forces.

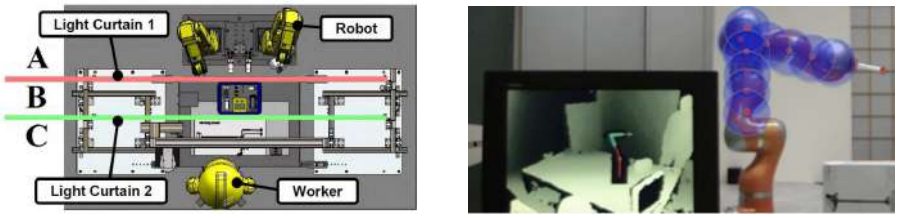
2.3.3. Collision avoidance

While reducing injuries in the event of a human-robot accident is crucial, in most of industrial contexts it is preferable to avoid collisions altogether, not only for safe reasons but also because the psychological state of the operator is an important aspect for safe HRC. Indeed, an effective coworking in smart factories requires that the operator feels comfortable and safe near the robotic system. In a paper written by Arai, et al. [49], an assessment of mental stress of a human operator working in a collaborative robotic manufacturing system has been done. Three influential factors have been considered: separation distance between the robot and the human operator, the robot speed and the warnings of motion. The result was that the velocity of the manipulator as well as the relative distance during the robot motion is related to the amount of induced stress. Furthermore, if the test subjects are not informed in advance of the manipulator speed, the stress levels get significantly higher. Thus, near the operator training which is very crucial for industrial collaborative scenarios, implementing collision avoidance systems to improve safety is the third main goal in HRC. Several approaches have been proposed in literature for this aim, as discussed in the next subsections.

Safety pre-collision strategies

Many techniques have been developed for estimating the distance between the robot and an object or generating an alternative path or/and trajectories, or varying robot acceleration and velocity before a collision occur. Some of them are still under analysis or they have been used to demonstrate how to prevent unfavorable robot-human collisions when humans and robots are collaborating.

An example of such early work is described in [50], where the study uses danger estimates obtained in real-time during the execution of a HRC task to generate an alternative trajectory if the danger level was higher than a predefined threshold. Based on [51], the danger index was developed as the product of the velocity, distance, and inertia variables. It was evaluated for the closest point between each link and the human. Inspired by [52] and [53], the alternative trajectory generation relies on a Virtual Force that aims to push the robot away from the warning area. This virtual force is a function of both the effective impedance at the closest point and the relative distance between the robot and the object. In further development, a human monitoring system -composed of a safe path planner and a safe motion controller- was integrated in [54]. In addition to psychological char-



(a) (b)

Figure 2.6.: Safety pre-collision strategies: scheme of light curtains establishing three safety working areas in a cellular manufacturing operation, [57] (a) and 3D data used to compute the distances between the robot and the dynamic obstacles, [58] (b).

acteristics like skin conductance, heart rate, pupil dilation, and brain and muscle reactions, the monitored data also includes physical characteristics like facial expression, head posture, hand gestures, and human eye gazing that are gathered by vision systems.

Other methods use the concept of *energy dissipation* to propose an injury index to be integrated in a pre-collision control strategy [55], or to propose a control algorithm that includes a safety indicator which depends on the relative distance between the human operator and the robot end effector [56]. In both cases experimental tests were carried out on lightweight robots.

Another pre-collision approach is based on the use of the same devices traditionally used in industrial environment to delimit the robot workspace, such as photoelectric sensors and light curtains. Indeed, in [57] a proposed and experimentally tested safety design for a HRC workcell for manufacturing was presented based on the use of two light curtains which distinguished three safety working areas, as shown in Figure 2.6a. More in detail, in zone A, which is the closest area to the robots, high speed movements of the robots are allowed; in zone B, which is the intermediate zone, low speed movements are permitted; in zone C, which is the closest zone to the human, the robot speed is limited to be below $150mm/s$ as to be compliance with the actual regulations for HRC applications. The system used two cameras to estimate the human operator position by localizing some color marks installed on the operator head and shoulders.

The evolution of the use of IP cameras has been proposed in [50, 54, 57] where stereo camera systems have been integrated into the collaborative workcell, or in [59] where IR-LED sensors have been applied in safety pre-collision strategies. However, 3D cameras quickly gained popularity for

collaborative applications to estimate the distance between the robot and the human operator. For example, in [58] 3D data were used to compute the distances between the robot and the dynamic obstacles and generate repulsive vectors which implemented a collision avoidance algorithm. More in detail, the repulsive vectors obtained for each control point (the center of the spheres that model the robot arm) were employed to modify the robot motion at the control point of interest, obtaining smooth and feasible joint trajectories which prevent collisions. Note that to perform the motion control while executing the initial task, two approaches were tested: a first one, which is a variation of the artificial potential field method, focuses on the avoidance of collisions involving the end effector; a second one aims to avoid collisions involving the robot body. Figure 2.6b shows the red centers of the blue spheres modeling the robot body, while the image on the screen below left visualizes the depth image of the moniored collaborative scenario in which the manipulator projection has been removed.

RGB-D cameras have been also adopted in [60] to estimate the distance between the human operator and the FRIDA robot. Figure 2.7 distinguishes two scenarios: the human ca be on side of the robot or in front of the robot. During the fist case scenario, the robot operates with its programmed speed while no human is inside the workspace. When a human operator enters the shared workspace the robot turns to a collaborative behaviour and the RGB-D cameras start to compute the human-robot distance in real-time. According to the computed value, when the human approaches the robot and the relative distance becomes less than a *warning distance threshold*, the robot speed is decreased to a pre-programmed one. Furthermore, in case the human is too close to the robot (the computed relative distance is less then a *dangerous distance threshold*) its behaviour turns to a safety exception, suspending the execution of the task. Finally, when the human leaves the work area or the distance can be considered to be safe, the task is resumed at the programmed speed. Otherwise, when the human is working in front of the robot, the scenario is considered more dangerous than lateral work and therefore the collaborative behaviour always involves a reduction of speed. This research is a first approach to the goal of maintaining the highest level of productivity of the collaborative workcell, distinguishing *autonomous behaviour* from *collaborative behaviour*, while guaranteeing human safety.

Different sensors and devices closely linked to the safety pre-collision strategies (the ones contained into the red rectangle of Figure 2.3) adopted to capture information from a collaborative scenarios and, as well as the way in which each type of data has been processed to develop a HRC algorithm,

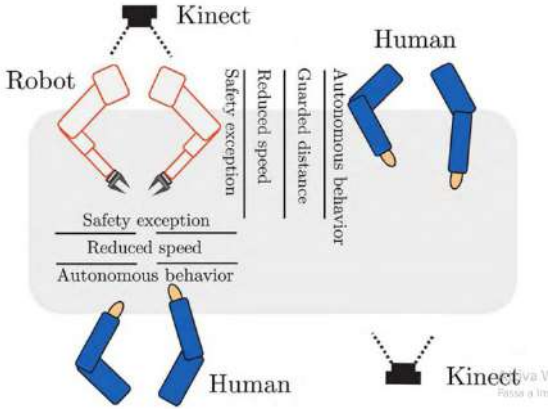


Figure 2.7.: Experimental set-up for HRC tasks using RGB-D devices, [60]

will be largely discussed in Chapter 4 and they will be compared with the multimodal vision system proposed by this thesis work.

3. A Human-centered workspace design

Human-centered manufacturing is an important pillar of Industry 5.0: humans should be at the center of manufacturing and understanding how technology can serve people, prioritizing their well-being. It does not have to be man who adapts to machines and work tools, but it is the environment in which he operates that must be designed and made for the worker. Instead of viewing robots and automation as replacements for human workers, the aim is to create a harmonious collaboration where technology complements human capabilities and enhances productivity. It is in this context that the SSM-type collaborative scenario is framed, which is the starting point from which the research products have been developed.

3.1. Industry 5.0 principles

Industry 5.0, a new and emerging notion of production settings that benefits workers, industry, and society as a whole, is based on the Industry 4.0 paradigm. This new framework places a strong emphasis on workers empowerment and designing a human-centric environment in which to operate. This paradigm will enable the transfer of human knowledge to robots, the adaptation of robot behaviors based on the human needs and the enhancement of working conditions for employees.

In fact, human-centered robotics places humans at the center of the design and the implementation of the robotic systems. The goal is to develop a technology that supports and augments human abilities, rather than replacing them. This approach emphasizes collaboration, safety and usability to create systems that work alongside humans effectively and efficiently. The following steps are essential for putting the Industry 5.0 paradigm into practice:

- adopting human-centric approaches for digital technologies, including artificial intelligence (AI);

- up-skilling and re-skilling workers, in particular considering digital skills;
- modern, resource-efficient and sustainable industries, allowing the transition to a circular economy.

Techniques that implements this industrial paradigm will make it possible to introduce a novel idea for production environments where humans play a central role. Although there have been some interesting developments in collaborative robotics scenarios, there are still no completely integrated applications aimed at the Industry 5.0 paradigm, i.e. solutions for designing, developing and validating human-centered problems. Thus, the goal of the study subject that is being suggested here is to begin filling the gap between the potential advantages that robotics might offer the industrial domain and the real applicability of that proposed solutions.

3.1.1. Human-AI collaboration

Systems and algorithms that provide a productive human-robot interaction and are always learning from human input are referred to as Human-Centered Artificial Intelligence (HCAI). This innovative branch of robotics pushes the boundaries of previously constrained AI solutions to bridge the gap between machine and human by developing machine intelligence with the aim of understanding human motions, language, emotion, and behavior.

HCAI based manufacturing is becoming more and more necessary for future industries to be ready for the ever-changing market issues. It would increase factories competitiveness, flexibility, and agility [61]. In fact, one of the key advantages is the potential to enhance productivity, efficiency and create safer work environments. By automating repetitive and mundane tasks, workers can focus on more complex and creative aspects of their work. Robots and automated systems can perform physically demanding or dangerous tasks, reducing the risk of injuries and increasing overall productivity.

As previously mentioned, the majority of past and present research is system-centric, with the goal of maximizing system performance and guaranteeing a safe working environment for all actors [62]. Human-centric AI-based manufacturing seeks to rekindle humans essential position in the manufacturing cycle while preserving their cognitive and psychological well-being in addition to providing them with greater opportunities for personal development. An anthropocentric paradigm for human–machine symbiosis was

proposed in [62]. Briefly, the work offers a reference model of human-centric manufacturing where people collaborate with empathetic technologies in a symbiotic way. Simply, "human and machine agents form intelligent teams to collectively sense, reason and respond to incoming manufacturing tasks to ensure productivity and workforce well-being". The authors state that the following are given priority in this reference model:

1. **Human centrality**
   the capacity to pay attention to human desire and judgment;

2. **Social wellness**
   the capacity to recognize and react to human physical and mental performance in order to maximize human wellness;

3. **Adaptability**
   the ability to learn from the environment and change the automatic behavior based on that learning.

This philosophy is also called *Human in the loop* and the success of this approach depends on how much we can learn from human behaviour: when the human operator is working with the automatic production plant; when he is writing the meta language to describe the production process and when he is designing the tasks for the standardisation of the robotic actions. The human being must be used in order to tune the entire manufacturing process and improve the general performance and safety. It is precisely from this perspective that this thesis work is placed, which aims at proposing a new robotic solution for the optimization of production processes that lies between the human operator and machines based on artificial intelligence.

3.1.2. A harmonious future of collaboration: application contexts

Applications for human-centered automation and robotics can be found in many different industries. Robots and humans work together on assembly lines in manufacturing to increase output quality and production time. Robots help in the medical field with procedures, patient care and rehabilitation. Collaborative robotics improves procedures like picking, packing, and inventory control in logistics and warehousing. Let's explore some of these sectors and how the HRC can benefit them.

- **Manufacturing and Assembly**
  Collaborative robotics support human workers on assembly lines in

the industrial sector by helping with duties including drilling, fastening, sealing, welding, material handling, and quality monitoring while the worker can dismantling some other metal parts for deburring purposes which need particular cognitive skills. Because these robots are capable of performing physically demanding and repeated activities, worker fatigue and injury risk are decreased. The output is that manufacturers can increase production and accuracy, reduce costs, and produce better work by automating some processes.

- **Healthcare and Medical Assistance**
  Surgical robots help surgeons in the healthcare sector execute minimally invasive treatments with greater accuracy and dexterity, which leads to smaller incisions, faster recovery periods, and better patient results. Robots help patients restore their motor skills and movement as part of rehabilitation therapy. Furthermore, robots in healthcare environments can carry out duties like cleanliness, drug administration, and patient monitoring, freeing up medical staff to concentrate on patient care.

- **Logistics and Warehousing**
  Robotic automation plays a major role in optimizing operations in the logistics and warehousing sector. Automated mobile robots (AMRs) and automated guided vehicles (AGVs) efficiently select and move items within warehouses. In jobs like order fulfillment, inventory management and packing, robots assist human workers. Automating these procedures lowers expenses, minimizes errors, and enhances overall operational speed and accuracy.

- **Agriculture and Farming**
  Automation and human-centered robotics have various advantages for the agriculture sector. Autonomous drones with sensors and cameras monitor crops and collect useful information on pest activity, hydration requirements, and plant health. Planting, harvesting, and weeding are among the jobs that robots can help with. This increases productivity and lowers the need for human labor. Robots may learn from and adapt to the specific requirements of each plant with the integration of AI and machine learning, which maximizes crop yields and resource efficiency.

- **Construction and Infrastructure**
  Autonomous drones can survey and check construction sites to ensure quality compliance and safety in the construction industry. Robotic exoskeletons relieve strain and avoid injuries during heavy lifting for

construction workers. Robots for construction may carry out repetitive jobs like pouring concrete and installing bricks, making building operations quicker and more accurate. Timeliness, worker safety, and efficiency are all enhanced by HRC developments.

- **Retail and Customer Service**
  Within the retail sector, both physical and virtual establishments are undergoing change. In warehouses and distribution centers, robots can help with order fulfillment, shelf stocking, and inventory management. Customers' in-store buying experience can be improved by in-store robots that can offer advice and information. AI is used by chatbots and virtual assistants to deliver online customer service and tailored recommendations. These technologies increase consumer involvement, expedite processes, and enhance customer service.

All of these uses highlight the adaptability and significance of human-centered robotics. These applications will grow as technology develops, opening up new avenues for human-machine cooperation. In this thesis work an innovative human-centred HRC system will be proposed which will then be applied in a real context of assembly of aerial fuselage panels that fall into the first macro-sector mentioned above.

3.1.3. Research contributions to I5.0

The research is primarily focused on the following topics, albeit not exclusively:

- vision-based and sensor fusion approaches applied to human-robot collaboration improvement;
- non-physical human-robot collaboration;
- AI applied to safety monitoring and control in working environments;
- human-centric intelligent design and control for industrial applications;
- work-place ergonomics evaluation, optimization and implementation for human-robot collaboration applications;
- integrated human-centric workplace design;
- real-time human monitoring to improve the working conditions, the mental health, and the workcell productivity;

- human-based optimization of processes and applications in collaboration with robots;
- interconnected human-robot data exchange for the optimization of the tasks allocation, robot motion planning and speed modulation.

3.2. Towards a novel SSM scenario

3.2.1. *Pre-collision* and *post-collision* scenarios

As described in Chapter 2 and in Appendix A, the ISO analysis classifies four collaborative scenarios. They can also be divided into two categories: *post-collision* and *pre-collision*.

A *post-collision* system reacts after the physical impact occurs between the robot and the operator. Human safety can be assured by minimizing the energy transmitted during the contact and by using robots endowed of a compliant structure or sensors for assessing force exchange when the impact occurs, as described in Section 2.3.

A *pre-collision* scheme makes use of exteroceptive sensors to detect humans and prevent collisions. Motion capture systems, range sensors or artificial vision systems are crucial in the case of distance monitoring, which is the most suitable approach for pure coexistence in a collaborative workspace. Any operator who enters the monitored dangerous zone around the robot causes it to slow down until it comes to a complete stop when the human operator approaches too closely. Several researchers proposed different methods for representing humans and robots geometry and implement pre-collision strategies.

There are three main negative aspects for a *post-collision* scenario:

1. if the robot limits are not well defined, a collision may be dangerous;
2. any collision will stop task execution, resulting in a decrease in production time;
3. especially in industrial applications, the robot is equipped with sharp tools, such as drilling tools, whose impact could cause serious injury.

For these reasons the thesis work aims at finding a pre-collision solution like. More specifically, since in industrial environments it is often recommended to stick to the robot's predetermined path without modifying it, the choice

of the collaborative scenario most consistent with a general industrial environment fell on the SSM. In an industrial scenario, in fact, the relative position and speed of the robot and the human operator can be used to build a safety metric for scaling the speed of the robot for the execution of its task and its trajectory.

3.2.2. Robot stopping functions

A crucial factor in *pre-collision* solutions is the stop function. Stopping any powered, mechanical system relies on one of the possible procedures described in IEC60204-1 [63]: (1) removing power to the drives, (2) applying brakes and (3) actively controlling motors to counter motion. The stopping function can be triggered via automatic mechanisms internal to the equipment (e.g., software watchdog systems, electronic monitors, or mechanical limit switches), external safeguards or manual switches (push buttons, pull cords or pedal-operated switches).

According to the standard, three categories of stop functions are specified, as summarized in Table 3.1:

1. **Stop category 0 (STOP0)**: the equipment is stopped immediately by removing power to the actuators and applying the brakes. This leads to a "uncontrolled stop" since the actuators' motion does not follow a predetermined course; instead, the goal is to stop the equipment as soon as possible;

2. **Stop category 1 (STOP1)**: the system stops under control, meaning that power is kept to the actuators until it stops completely, at which point the brakes are engaged and the power is removed;

3. **Stop category 2 (STOP2)**: the system comes to a controlled stop, but neither the power nor the brakes are deactivated.

In this work, a stop category 2 (STOP2) has been considered.

Table 3.1.: Robot stopping categories as defined in IEC60204-1, [63].

Stop category	**Controlled**	**Power removed**	**Brakes applied**
STOP0	No	Yes	Yes
STOP1	Yes	Yes	Yes
STOP2	Yes	No	No

Figure 3.1.: Human operator in a dangerous zone (red). The warning zones are in orange and yellow.

3.2.3. SSM SoA development

As previously described in Section 2.2, SSM enables simultaneous movement of the operator and the robot system in a collaborative workspace. Here are some implementation details. Ensuring that there is always at least the protective separation distance, S_{ISO}, between the human operator and the robot helps to reduce risk. The robot system never approaches the operator closer than S_{ISO} while it is moving. As illustrated in Figure 3.1, the robot system stops before it may affect the operator when the Euclidean separation distance, d, is equal to or less than S_{ISO}. The robot system may automatically resume motion while preserving at least the protected distance to separation S_{ISO} when the operator moves away from it.

The first study to look into the problem of placing safeguards to protect people from stationary, active machinery is ISO13855 [64], in which specific parameters based on values for approach speeds of parts of the human body are provided. The process outlined in ISO 13855 is used to determine

the minimum distances to a hazard zone from the detection zone or from the devices that activate safeguards. The speed parameters for techniques like walking pace and upper limb movement have been tried and tested in real-world scenarios. This is how the standard has been developed: it provides guidelines for common approaches, with the exception of running, jumping or falling. The equipment for electro-sensitive or pressure-sensitive protection, two-hand control devices and interlocking guards without guard locking are among the safeguards that are taken into account in this standard. The document recommends calculating S_{ISO} as

$$S_{ISO} = vT + C, \tag{3.1}$$

where v is the approach speed of human body parts and it varies according to the separation distance d:

$$v(d) = \begin{cases} 2.0m/s, & \text{if } d \leq 0.50m \\ 1.6m/s, & \text{otherwise} \end{cases}$$

Note that $2.0m/s$ is assumed to be as the maximum operator speed. T is the total system stopping performance time, in seconds, and it is a combination of the time required by the machine to respond to the operator's presence (i.e., T_R) and the response time of the machine which brings the robot to a safe, controlled stop (i.e., T_S). C is the intrusion distance safety margin, which represents an additional distance, based on the expected intrusion toward the critical zone prior to the actuation of the protective equipment.

From Equation 3.1, ISO/TS 15066 updates the S_{ISO} meaning by including robot dynamic properties. At time t_0, when the robot system reduces its speed, the minimum protective separation distance, S_{ISO}, decreases correspondingly, i.e.

$$\begin{aligned} S_{ISO}(t_0) \geq & \int_{\tau=t_0}^{\tau=t_0+T_R+T_S} v_H(\tau)d\tau + \int_{\tau=t_0}^{\tau=t_0+T_R} v_R(\tau)d\tau + \\ & + \int_{\tau=t_0}^{\tau=t_0+T_R+T_S} v_S(\tau)d\tau + (C + Z_S + Z_R) \end{aligned} \tag{3.2}$$

In Equation 3.2, v_H is the *directed speed* of the closest operator which travels toward the robot, v_R is the speed of the robot in the direction of the operator, v_S is the directed speed of the robot in course of stopping. The remaining terms represents uncertainties: the intrusion distance C is based on the operator reach, Z_R is the robot position uncertainty and Z_S is the operator position uncertainty (i.e., the sensor uncertainty). For all practical safety purposes, the time, t_0, is considered the current time.

An external observer system built within a robotic workcell ensures operator safety. As people, or "operators," get closer, the observer system scans the area around the robot and gives commands to slow down or stop it. The safety system starts a safety-rated, controlled stop if the separation distance d is less than the value of S_{ISO} at time t_0. When the separation distance is more than S_{ISO}, the robot may then start moving again.

The main issue of ISO 13855 is that the separation distance was initially intended for static machinery, not for dynamic and reconfigurable robotic systems. Therefore, extending what is contained in the standard to the case of industrial robotics is not trivial. Nevertheless, ISO/TS 15066 tries to make a contribution to the HRC problem and describes S_{ISO} using the following linear function

$$S_{ISO} = (v_H T_R + v_H T_S) + (v_R T_R) + (B) + (C + Z_S + Z_R) \tag{3.3}$$

where B is the Euclidean distance travelled by the robot while braking. Note the one-to-one correlation between Equation 3.2 and the linear relationship Equation 3.3. The first term in parentheses describes the contribution attributable to the operator's change in location in the time necessary to bring the robot to a full stop from its current speed. The second term describes the contribution attributable to the robot system reaction time, before it initiates the braking sequence. The third term describes the distance travelled by the robot during its braking. Finally, the fourth term describes the possible distance of intrusion into the robot work volume as a function of the operator reach and the uncertainty of the sensory system and robot kinematics. The values of v_H, T_S, B and C can be found in the safety standards: the values of v_H and C are given in ISO 13855, while guidelines for evaluating T_S and B are given in Annex B of ISO 10218-1 and they result from measurements that directly depends on the robot system under test.

3.2.4. Research contributions to SSM

Reducing the protective zone around the robot while maintaining safety standards is the main objective of this work. Only in situations where there is an immediate risk of collision the robot should change its behavior in terms of trajectory scaling. In order to handle the extreme variety and unpredictable nature of human behavior, a thorough analysis of the operator approach into the collaborative workspace has been conducted. The developed solution calculates the points at minimum distance between the robot and the nearest human operator and has a lot of advantages over the other current SoA techniques, many of which rely on evasive actions to increase safety. Nonetheless, in an industrial environment, it is usually advised to adhere to the robot's predetermined path without changing it, particularly in complex work cells where collisions are likely to occur. The main characteristics of the proposed approach are described below, while the implementation details will be largely described in next chapters:

- it considers the whole surface of human operators, without skeleton-based techniques and without approximating the body to a single point of mass;
- it considers the whole robot kinematic chain, the entire volume and possible tools, without factoring only a singular representative coordinate of the robot (e.g., the end effector);
- it explicitly takes into account the regulations;
- it allows for the provision that the human speed, v_H, may be estimated directly and it is not assumed constant;
- it considers the relative directions of velocities, which are not factored into the Equation 3.3 proposed by the actual ISOs;
- it does not modify the robot programmed path and it does not require the task to be aborted.

4. Sensor technologies for *Human Detection*

A typical industrial robot's workspace is at least partially unknown because of environmental changes in terms of objects and tools shifting or of human operator motions within the workplace. There are two basic approaches in literature that have been taken for analyzing these unpredictable situations. The first one approach suggests using a geometric representation of the human body, thus using a motion capture system to localize human position and to estimate his direction of motion to be used for HRC tasks. The second approach, otherwise, makes use of contactless devices that offer 3D and/or visual data. The innovative sensor fusion strategy proposed by this thesis work extends the second approach for a HCAI collaboration.

4.1. Motion capture systems and simulated environments

Geometric representations of the environment and objects within it have been developed in earlier studies, before the pure research on HRC. Nowadays, they represent the basis of much work employing motion capture systems for collaborative tasks: [65] and [66] are two main examples. The first one work describes a fast collision detection technique for a manipulator task based on the *Octree Representation Scheme* technique which is a recursive workspace decomposition. The second study, otherwise, suggests a method for detecting collisions between an industrial robot and solid objects in its workspace that uses a hierarchical spatial representation based on spheres.

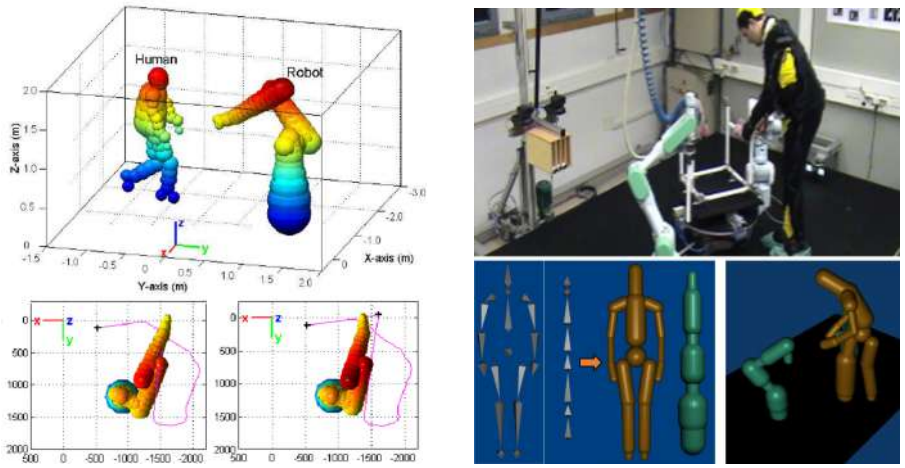
Focusing on a more complete human-robot collaborative tasks, the first idea of using a geometric representation of the environment is to compute the minimal human-robot distance, allowing for the search for collision-free paths. In [67], a novel solution to the 3D dynamic human-robot collision problem is presented. Sphere-based geometric models are used to represent

both the human and robot. The collision avoidance algorithm searches for collision-free paths by moving the end effector along a set of pre-defined search directions. An optimization algorithm selects the search direction considering the goal location and maximizing the distances between the human and robot models. The optimization incorporates predictions of the motions of the robot and human to reduce the negative effects of a non-instantaneous robot time response. The simulations employ data from a motion capture system to animate the human models, as shown in Figure 4.1a). This method may prevent all human-robot collisions when the person is approaching the robot within the robot workspace, as shown by simulations involving a PUMA robot arm and a human.

Following the philosophy of geometric model representation of the robotic workcell, the SphereSwept Lines (SSLs) technique has been introduced in [68] and then expanded in [69]. The approach was based on bounding volumes with the idea of overcoming the drawbacks of the previous model:

- the high computational cost of systems using a mesh of polygons;
- the imperfect fitting of spheres to human and robot surfaces;
- the lack of accuracy of systems based on linear skeletons of humans and robots.

More in detail, the structure of the body of the human operator and the robotic manipulator are represented with the same model: a linear skeleton. The skeleton of the human operator is composed by 18 nodes while the skeleton of the robot is composed by 8 nodes. Each bone of the skeleton of the human operator and the robot manipulator is covered with a SSL bounding volume which models the geometry of the corresponding link, as shown in Figure 4.1b) bottom left. To localize the human in the robot workspace, the information from a inertial motion capture system is combined with the data from a UWB (Ultra-WideBand) localization system which uses the technique of triangulation of information from four fixed camera sensors located in the workspace and a series of small tags attached to the human body wear to estimate the human position, thus the environment simulation is provided in real-time (see Figure 4.1b) bottom right) to compute the minimum distance between human and robot. This value is measured to identify the risk of impact: when it is sufficiently high, the robot nominal trajectory is stopped and the control system generates an alternative collision-free trajectory. The strategy was successfully implemented and tested in an assembly and disassembly collaborative tasks, as shown in Figure 4.1b) top, guaranteeing human safety in conditions of close



(a) (b)

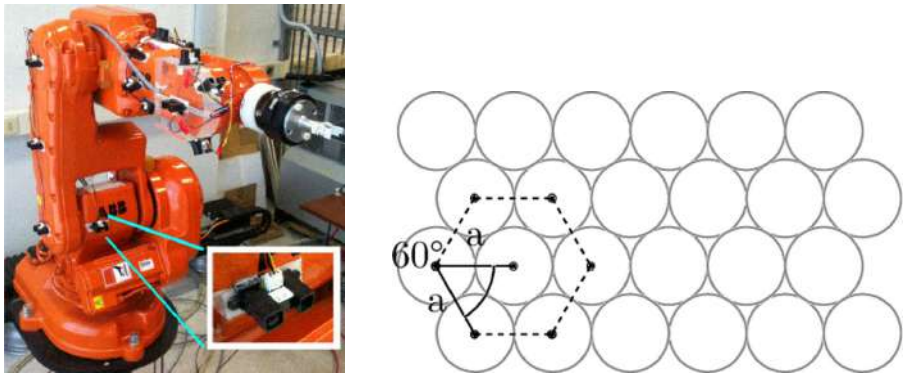
Figure 4.1.: Spherical geometric model of a workcell containing a Puma-762 robot arm and a human, (top) for collision-free path searching (bottom), [67] (a); a collaborative task demonstration of SSL bounding volumes technique for the human and the robot representation, [69]

proximity between human and robot.

4.2. Sensors capturing local information

Some researches on avoiding object-robot collisions (and therefore also human-robot collisions) were focused on obtaining information about the local space surrounding the industrial robot. For this purpose, robot arms were provided with sensors like infra-red proximity [70], capacitive sensors [71, 72, 73], ultrasonic sensors [74] or laser scanner systems [75]. However, the disvantages of these sensors is that they does not cover the whole scene, thus they can only provide a limited contribution to improving safety in collision avoidance tasks, thus therefore HRC tasks.

An example of sensors capturing local information applications is described in [59], where 20 distributed Sharp IR-Led sensors were mounted on-board of an industrial robot arm, as shown in Figure 4.2a, by using the optimization method suggested in [76] and based on the hexagonal compact packing disposition of distance measurement spots which provides the maximum spot density. Their signals were used for the calculation of danger field



(a) (b)

Figure 4.2.: ABB IRB 140 robot equipped with 20 Sharp IR-LED sensor spots, [59] (a); hexagonal compact packing disposition of distance measurement spots, providing maximum spot density, [69]

vectors associated to each sensorized link. This data was incorporated into a reactive control scheme to force the robot to search and execute a collision-free path when indispensable. The outcomes demonstrate that this strategy maintains task consistency while boosting human safety by guaranteeing about 90% of the chance of human detection.

4.3. Artificial vision systems

Human-robot collisions have also been avoided with the use of artificial vision systems. In the literature there have been multiple studies that fall into this category. [77], for example, was one of the first robotic application based on vision system: it developed a two-module safety system for HRC, as shown in Figure 4.3. The first one module uses a CCD camera and computer vision techniques (the shading algorithm presented in [78]) to compute the difference of the gray level between current image and a reference image to detect changes in the monitored scene. This module not only determines the human's position but also executes three different robot behaviors based on the computed distance between the human and the robot itself. More in detail, if no humans are present in the working area, the robot operates at its highest preprogrammed speed. Otherwise, if a human is found within the robot working area, the robot speed is limited to a safe value with respect to the preprogrammed speed. Finally, there is a predefined minimum safe distance between the robot and the human:

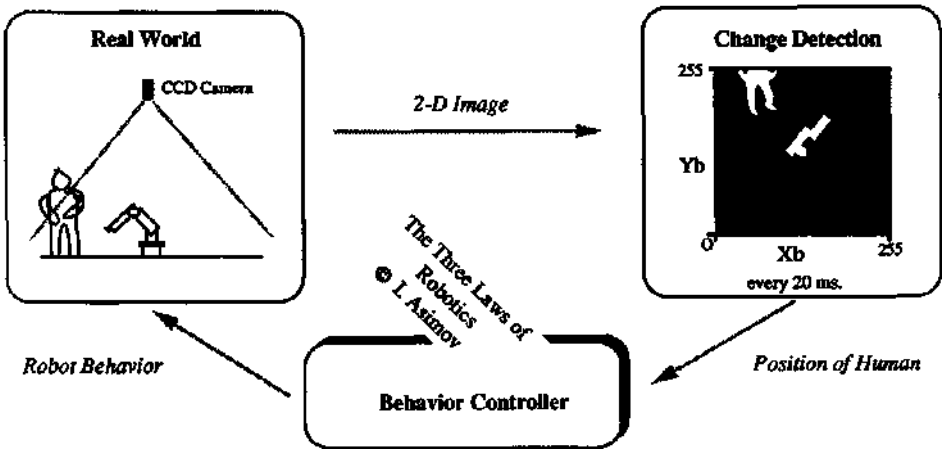


Figure 4.3.: The control scheme implemented in [77] for HRC based on image processing module which uses data from a CCD camera installed in the monitored workspace.

the robot is stopped if the gap between it and the human gets less that this threshold. The other module is used to prevent an unexpected robot motion that could result from a hardware malfunction in the robot. It is based on joint position information and accelerometers on the robot.

An evolution of [77] was [79] in which not only one camera has been installed to monitor a scenario but several fixed cameras, to prevent image occlusions. The current image frame is given as the input to an algorithm that calculates image differences to detect the position of the human. In order to distinguish the robot entity from the human one, data from the robot position sensors is projected into each camera's image. This, along with estimates of the human location, enables the calculation of the robot-to-human distance.

A new methodology to estimate human position in a monitored workspace was based on the use of color markings on the human head and shoulders and two IP cameras in [80]. A color filter method, based on the HSV components, was applied to overcome the drawback of different light conditions. The Direct Linear Transformation (DLT) Method was used to estimate 3D data on each mark's center position. In the subsequent work [57], this vision system was used on an HRC prototype production cell to demonstrate its efficacy (Figure 4.4).

The off-the-shelf safe camera system SafetyEye, developed by Pilz company, comprises three components, as shown in Figure 4.5: a sensing device made up of three extremely dynamic cameras that don't require calibration; a

Figure 4.4.: Operator monitoring system for active HRC prototype production cell design, [80].

high-performance computer that serves as the analysis unit and creates a 3D image from the images taken by the cameras; and a programmable safety control system that manages the entire SafetyEYE operation while simultaneously serving as the robot controller's interface. The security camera system allows continuous 3D monitoring and control of a dangerous workspace: it allows an operator to delimit one or more danger zones inside the monitored area and define the speed of the robot for each of them. After the configuration, the system identifies any object that enters one of the designated zones and adjusts the robot's movement to stop or decrease speed in response to the configuration itself.

4.4. Range systems

Range systems create a 3D range map of objects in the surrounding environment by using light patterns, IR-LEDs, lasers, and even stereo camera systems.

[81] provides the first example of range imaging cameras integrated into a safety system for a robotic industrial environment. The paper uses two 3D Photonic Mixed Device (PMD) cameras for fast collision detection of multiple unknown objects. The depth sensors are placed around the workcell to observe a common surveilled 3D space. The visual hull technique and the fusion of 3D data from the two range imaging cameras are used

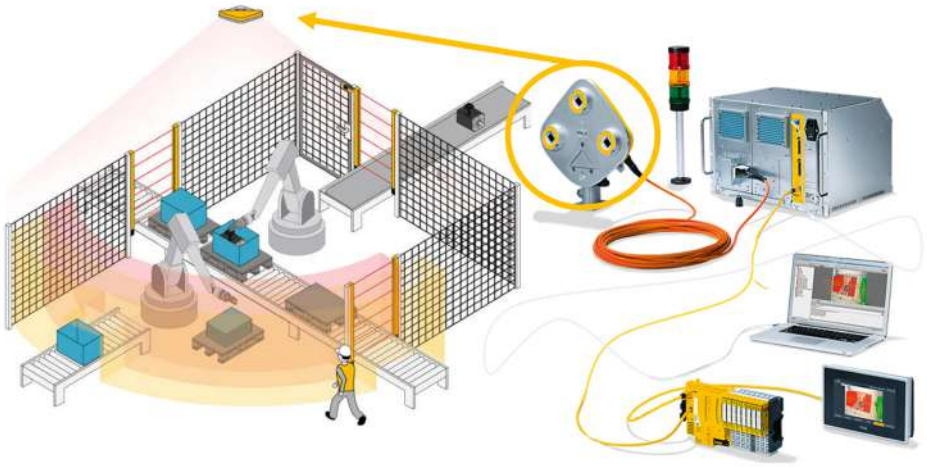


Figure 4.5.: SafetyEye.

to calculate a conservative approximation of all detected obstacles within the area of interest. To check for potential collisions, the configurations of a robot model and its projected route are combined with data on the observed impediments. More in detail, the minimum distance to any obstacle is used to limit the maximum velocity. In Figure 4.6 (left) is shown a prototype setup of a workcell containing an industrial robot Stäubli RX130, whose CAD model is known (consisting of one cube per robot link). Two PMD cameras have been placed in the upper corners of the workcell to monitor the robot workspace. At any time, the current robot configuration can be requested from the robot controller to move the robot model into the simulated environment and a proper velocity for the current trajectory segment can be sent to the controller. On the right it is possible to observe the minimum distance calculation between the robot, the orange 3D model, and the detected unknown objects, the blue entities. The small images display the results of the object detection in the depth images. The minimum distance (red line) is used for velocity control (diagram), thus to limit the maximum velocity of the robot (accuracy of measurements is within an interval of centimetres).

[82] is the extension of the previous work, introducing a general approach for surveillance of robots using depth images of multiple distributed smart cameras which can be either standard colour cameras or depth cameras. Unknown objects intruding the robot workcell are detected in the camera images and the minimum distance between the robot and all these detected unknown objects is calculated. The authors present a comparison between

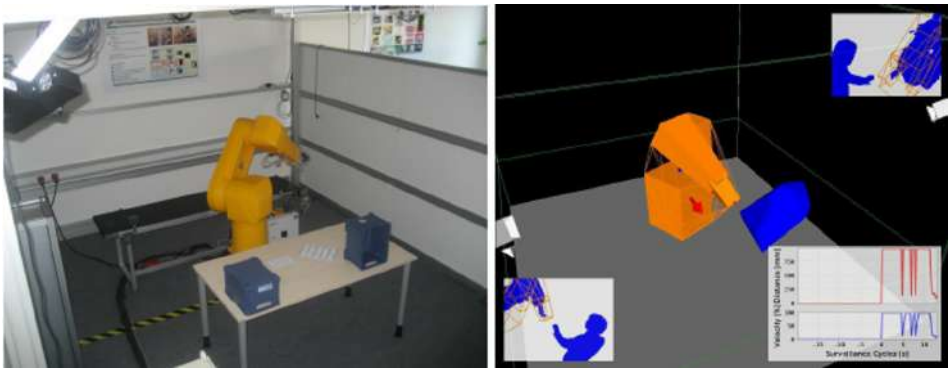


Figure 4.6.: Experimental validation setup in [81]: the robot workspace is monitored through two PMD cameras (right); minimum distance calculation (red line) between robot (orange) and detected unknown objects (blue).

distance information from color and ToF cameras, as well as between a single ToF camera and ToF information fusion from multiple cameras, using an experimental setup consisting of multiple range and color cameras monitoring the workspace of a robot arm. One of the conclusions of this work is that the predicted distance between a robot model and the unknown item is longer than the distance obtained using information from one colour camera. Another conclusion is that the fusion of information from several ToF cameras provides better resolution and less noise than the information obtained from a single camera.

A framework for HRC based on range systems has been given in [83]. A key element of this research work is the reconstruction of a robotic environment's workspace not through marker kinematic estimation but by using a range imaging camera fixed at the top of the workcell. Background subtraction techniques based on Gaussian Mixture Models were developed. Using background data, algorithms based on eigenvalue analysis, depth measurements of pixel distributions, connected component distribution, and motion features produced from optical flow estimation are used to determine whether a pixel cluster is produced by a human presence. A fuzzy logic module that assesses risks and regulates the robot uses this data to derive more reliable features than modeling human kinematics.

Figure 4.7.: Hardware configuration of Microsoft Kinect with detailed location of each sensor (left); additionally, two image samples captured by the RGB camera and the depth camera are provided (right).

4.5. RGB-D devices

2010 saw the commercial release of the Microsoft Kinect, a new type of 3D sensor known as RGB-D that is equipped with both an RGB camera and a 3D depth sensor, as shown in Figure 4.7. This hardware overcame the ToF cameras first appeared on the market. Compared to ToF cameras, the RGB-D sensor offers several number of benefits, including better resolution, a lower cost and the availability of both color and depth information, for widespread use. The complementary nature of the depth and visual information provided by the Kinect sensor openned up new opportunities to solve fundamental problems in computer vision, as described in [84]. The paper presented a comprehensive review of Kinect-based computer vision algorithms and applications: preprocessing, object tracking and recognition, human activity analysis, hand gesture analysis and indoor 3D mapping. The researh work served as a tutorial and source of references for Kinect-based computer vision researchers.

In collaborative robotics, the Kinect sensor can be used as a replacement to range systems or conventional vision cameras. For instance, the depth sensor of the Kinect is modeled as a traditional pin-hole camera in [58] in order to record the depth space of the surroundings. To develop a real-time collision avoidance algorithm, the robot-obstacle distance was estimated and used to modify on-line the trajectory of the robot arm, as shown in Figure 4.7. Subsequent research provided and evaluated an enhanced method for calculating the separation distance between items sensed by the RGB-D sensor and a point of interest in Cartesian space in a highly dynamic human-robot collision avoidance task, [85].

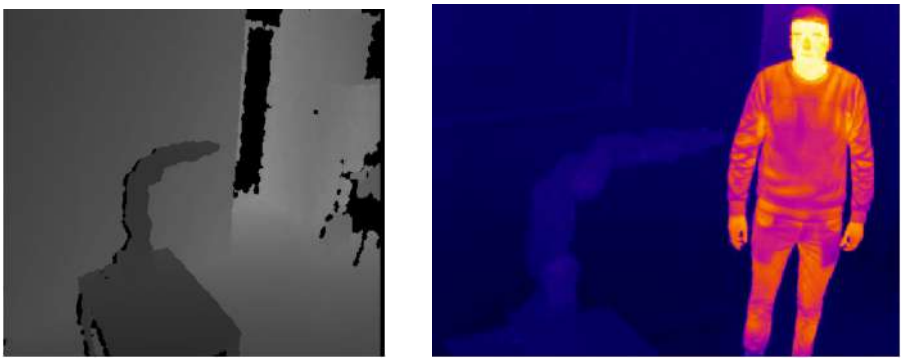
Figure 4.8.: A robot arm reacts instantaneously to motions of humans and other dynamic obstacles that are detected in depth space, such that collisions are avoided, [58]

4.6. Innovative multimodal vision system

As previously mentioned, motion capture systems, range sensors and artificial vision systems can all be used to solve the problem of distance monitoring. However, robust localization of human operators is a challenging task. Fusing many sensors with distinct characteristics can be required.

Thermal cameras were first created for military use, but they now offer useful information for a variety of tasks, including infrastructure and electrical system monitoring, gas and liquid detection, inspection and control in industrial applications [86], breast cancer diagnostic [87]. Thermal cameras are perfect for detecting objects at specific temperatures. Since human body temperature is approximately 37° C, human detection and tracking (HDT) is a good fit in this situation. Unfortunately, depth information is not supported by thermal cameras, which makes it impossible to accurately calculate the separation distance between the operator and the robot and implement a SSM collaborative scenario applying the current regulations.

The primary challenge in combining thermal and spatial images is determining the correspondence between related pixels. A comparable sensor fusion method has been developed in [88] for indoor human detection integrating RGB data with depth information: this method emploied the depth feature that describes the self-similarity of an image along with the Histogram of Oriented Gradients (HOG) suggested in [89]. Different strategies are based on Convolutional Neural Networks (CNN), widely used for object recogni-



(a) (b)

Figure 4.9.: Multimodal vision system: depth image from the Microsoft Kinect v1 depth camera (a) and thermal image from Optris PI 450 thermographic camera (b)

tion [90] and human detection [91]. A CNN-based RGB-D human detector exploiting the depth information to develop a region of interest selection method (ROI) is proposed in [92]. However, the fusion of thermal and spatial information has gained attention in the last few years, especially in fields where the spatial data are used as the main source of information, but nowadays there are no standardized methods to robustly combine them.

4.6.1. Hardware components

The following section is an overview of the hardware devices selected to be combined in a multimodal vision system to solve the human detection task: the Microsoft Kinect v1 and the Optris PI 450.

Depth sensors

The first sensor chosen for the research activity has been a depth sensor, with the goal of developing not only a human detection algorithm but also to compute the minimum human-robot separation distance in collaborative SSM scenarios. The primary feature that sets apart a depth camera from a traditional 2D video camera is its capacity to estimate the distance, also referred to as "range" or "depth," of a specific point in a scene from the sensor itself. Many different kinds of technology exist, such as the typical Laser Range Finders (which can use triangulation or Time of flight methods but can only produce a single depth point), or the scanning Lidar (which scans an area with a few beams and gradually builds up a depth map).

The depth data is calculated using one of the three methods proposed in literature.

The *Time of flight* (ToF) technology measures the amount of time that the light emitted from an illumination unit needs to travel to an object and back to the sensor array. To do this, the scene is pulse-illuminated with periodic radiation, such as near-infrared light, infrared light or laser. Assuming sensor and lighting placed in the same position, due to the distance between the sensor plane and the object, a time shift is caused in the optical signal which is equivalent to a phase shift in the periodic signal. This shift is detected in each sensor pixel by a so-called mixing process.

The *structured light* approach is an active stereo vision technology that projects patterns that are known to exist on scene objects in order to compute the objects' surface and depth information. Parallel stripe patterns are popular because they can precisely retrieve the 3D coordinates of any details on an object's surface due to their ability to deform upon impacting surfaces, even if there are many other variations of structured light projection that can be used.

The *triangulation* can be seen as a process of determining a point in 3D space given its projections onto two images coming from two different cameras which frame the same scene from two different views. This method, which is based on the visual perception system of humans, is the foundation of stereovision and 3D reconstruction. By using equipolar geometric consideration, the corresponding 3D point may be determined given the two images and the spatial relationship between the cameras, as shown in Figure 4.10.

The adopted depth camera is the Microsoft Kinect v1 shown in Figure 4.7, which combines an RGB visible spectrum camera and an infrared (IR) spectrum 3D camera, i.e. the depth image shown in Figure 4.9a. The Kinect v1 camera combines structured light and stereovision by utilizing an IR projection and two IR image sensors to provide more accurate depth information: IR emitter project a light pattern and an IR sensors detect the deformations in the projected pattern for resolving the depth. More in detail, the camera analyzes the shift of the speckle pattern by projecting from one location and observing from another. The depth map can be combined with a color camera to produce point clouds. The specifications of the Microsoft Kinect v1 are shown in Table 4.1.

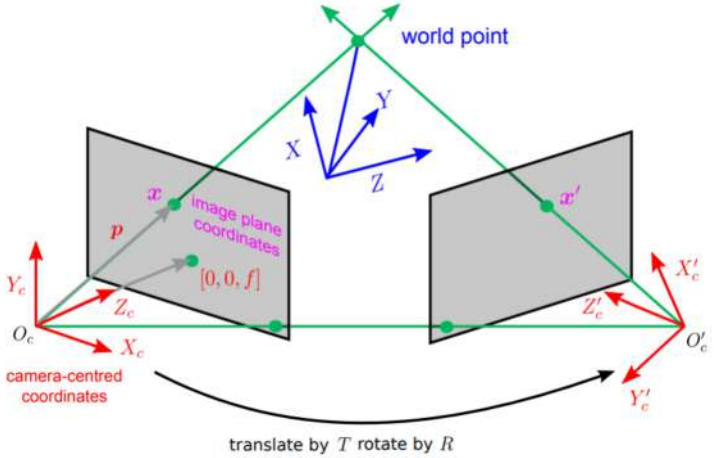


Figure 4.10.: Stereo Vision Model

Table 4.1.: Microsoft Kinect v1 specifications.

Feature	**Microsoft Kinect v1**
Focal length:	6.1 mm
Color camera	640×480 at 30 fps
Depth camera	320×240
Max depth distance	∼4.5 m
Min depth distance	0.4 m
FOV	57×45

Thermographic Camera

The second sensor chosen for the research activity has been a thermographic camera, with the aim of identifying all and only the human clusters present in the scene, then developing a robust human detection algorithm that has a low percentage of false positives, in order to optimize productivity time. A thermographic camera, also known as a thermal imaging camera, uses infrared radiation to identify heated objects. In contrast to standard cameras, which function within the visible light spectrum of 400–700 nanometers, thermal cameras can detect light at up to 14 µm in wavelength.

Any item with a surface temperature higher than zero produces electromagnetic radiation which can be detected by a special camera in a way similar to the way an ordinary camera detects visible light. The emitted radiation is characterised by two features: the wavelength, λ, and the intensity (or

thermal emission), Q. The following Equations 4.1 and 4.3 connect these two parameters to the surface temperature of the body. The first formula is:

$$Q_{tot} = \sigma T^4 \tag{4.1}$$

where Q is the total radiation emitted by a body and depends on both the surface temperature, T, and σ, which is the Stefan-Boltzmann constant and its value is:

$$\sigma = 5.67 \times 10^{-8} W m^{-2} K^{-4} \tag{4.2}$$

Otherwise, λ_{max} is the peak wavelength of emission and also its value depends on the surface temperature, T:

$$\lambda_{max} = 2.9/T \tag{4.3}$$

Images captured by infrared cameras are typically displayed in monochrome, though occasionally they are presented in pseudo-color, which uses color changes instead of intensity changes to indicate signal changes. This method, known as *slicing*, aids in the human ability to perceive subtle variations in intensity.

The thermographic camera used in this vision system is the Optris PI 450 (Figure 4.9b). The detectable temperature is within a range of -20°C up to 900°C while other technical specifications are reported in Table 4.2.

Table 4.2.: Optris PI 450 specifications.

Feature	**Optris PI 450**
Focal length	15 mm
FOV	38×29
Spectrum	7.5 to 13 m
Image size	382×288
Frame rate	80 Hz

4.6.2. Software components

The following section is an overview of the software platforms adopted to develop the sensor fusion strategy for the human detection algorithm: ROS, KERAS, YOLO and MATLAB.

ROS

The open-source Robot Operating System (ROS) [93] is a software library and toolkit that facilitates the development of robot applications that are compatible with a broad range of robotic platforms. Standard operating system functions including hardware abstraction, low-level device control, process message forwarding and package management are provided by the operating system side. The peer-to-peer network of processes, known as the ROS runtime "graph", is made up of a large number of small computer programs that are constantly communicating with one another, exchanging messages: they allow visualization, logging, plotting data streams, etc. The ROS software modules can be written in any language for which a client library has been written (C++, Python, etc.).

The framework is based on concepts as nodes, topics, messages and services. The nodes are single-purposed executable programs (e.g. sensor drivers, actuator drivers, mapper and planner) that are individually compiled, executed and managed. They can publish or subscribe to a Topic and also provide or use a Service. Nodes communicate with each other by publishing messages to topics so the topics represent a stream of messages with a defined type.

KERAS

A popular Python high-level neural network API for quick prototyping is called Keras [94]. This deep learning library is capable of supporting both recurrent networks and convolutional networks, and also their combinations. Additionally, it enables the CNN to operate on either the CPU or the GPU, depending on the required performance. Depending on the network model selected, Keras offers two options for configuring the CNN's layers. The most basic kind of model is the sequential one, which is equivalent to a linear stack of layers made up of several convolutional, activation and pooling layers. The number of channels and the resolution of the input images must be known by the model.

The input image size for the case study is approximately 388x288, that is the

resolution of the thermal camera. There are two channels involved: one for the depth image and another for the thermal image. The learning procedure has been configured to maximize the output classification accuracy in order to train the chosen model. To put it more precisely, the learning process must be assembled by choosing the optimization technique, the loss function that needs to be optimized, and the metric to apply to the classification task. The training procedure is initiated by the initial fit function, which obtains the training picture data set and the associated labels. When the batch size, epoch count, or validation data are reached, the training phase is over.

YOLO

Current detection techniques suggest using image classifiers to locate and recognize objects. More modern methods, such as R-CNN [95], combine region proposals with CNNs: it extracts about 2000 bottom-up region proposals for each image, uses a convolutional neural network to compute a proposal's features, and then uses class-specific linear Support Vector Machines (SVMs) to classify each region. Subsequent post-processing is needed to remove duplicate detections and improve the bounding boxes. Because each component of these complicated pipelines needs to be trained separately, they are too slow and difficult to optimize.

You Only Look Once (YOLO) [96] is a novel, incredibly quick, and accurate real-time object detection system. This system uses a single neural network to segment the image into regions and forecast bounding boxes and probabilities for each region. It approaches object detection as a single regression problem. Unlike R-CNN systems, which require thousands of evaluations for a single image, YOLO also generates predictions with a single network assessment. Because of this, it is incredibly fast, more than 1000 times quicker than R-CNN. The YOLO framework was selected for the case study because it emphasizes an extraordinary speed/accuracy trade-off.

MATLAB

Developed by MathWorks [97], MATLAB is a scientific computing environment with numerous applications, ranging from the pocket calculator to the simulation and analysis of complex systems. Operating systems like Windows, Mac OS, Linux and Unix can handle MATLAB, which is intended to address scientific and design challenges. This software is used by scientists and engineers worldwide to study and create the goods and sys-

tems that change the world: active safety systems in cars, interplanetary spacecraft, health monitoring gadgets, smart power grids and LTE cellular networks are actually made of MATLAB. This platform finds applications in a wide range of fields, including robotics, control design, computational finance, machine learning, signal processing, image processing and machine vision. Since its programming language is built on matrices, MATLAB is an acronym for MATrix LABoratory, which stands for the most natural approach to describe computational mathematics in the world. A vast toolkit library and integrated graphics make visualization easier while providing a thorough comprehension of the data.

4.6.3. Sensor fusion strategy

Sensor fusion combines sensory data to obtain output data of grater information content [98]. Original sensor data typically have limits in terms of uncertainty, imprecision and limited coverage in terms of both space and time. These issues could be partially addressed by a sensor fusion technique, which creates an enhanced single view of a scene with expanded information content [99] by combining multi-modality sensor images of the same observed scene. Image fusion is divided at the hierarchical level, e.g. by processing the features extracted from the input images, or at the low level, e.g. by processing original sensor images. The second way has been applied in the case study, wherein a single fused image was created by combining the original depth and thermal data. The methods used to calibrate the sensors, the method created to map the pixels from the depth picture onto the corresponding pixels of the thermal image, and, lastly, the suggested approach for performing image fusion are all explained below.

Vision system calibration

Camera calibration is an essential task for extracting metric information from 2D images. A camera must be calibrated in two steps: the *intrinsic calibration* describes the internal features of the sensor and it is needed to reject the optical distortions introduced by the camera lenses, while the *extrinsic calibration* describes its position in space. A calibration target (such as a perforated grid) can be used for the first type of calibration and these procedures are standardized [100, 101]. On the other hand, in literature a lot of solutions to be adopted for the extrinsic calibration of the camera have been proposed but each one suffers from many limitations and constraints that describe the calibration approach just for the aim of solving it, without taking into account its applicability at all or providing an

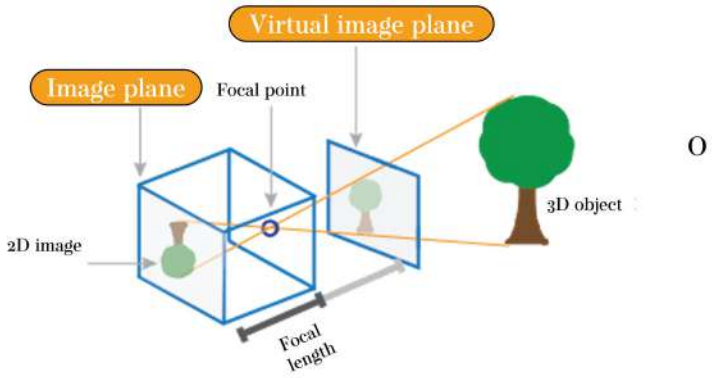


Figure 4.11.: Pinhole camera model: image plane and virtual image plane.

appropriate contextualization, all based on rather conservative assumptions.

For this reason, while the intrinsic calibration of the sensors follows a classic procedure, new approaches have been proposed for the extrinsic calibration of them, more suitable for industrial contexts. More in detail, two developed approaches for extrinsic camera calibration are explained in this section. The goal is to get the two cameras poses as precisely as possible for the merging process to be successful and the subsequent human detection and separation distance computation to be completed.

1. Pinhole camera model

A pinhole camera is a basic camera with a single, tiny aperture and no lens. As light beams enter the aperture, they display an image that is reversed on the other side of the camera. The *virtual image plane* is in front of the camera and containing the upright image of the scene, as shown in Figure 4.11.

The geometric model of a pinhole camera is shown in Figure 4.12 and consists of two coordinate systems: a reference system Σ_c centered in O_c called the *optical center* and with the z-axis coinciding with the optical axis; and a second plane, called *image plane* with a reference system Σ_p centered in O_p and with u and v axes oriented respectively as x and y. The image plane is at focal length from the origin O_c of the camera reference system.

The pinhole camera parameters are represented in a 3×4 matrix called the *camera matrix*. The 3D world scene is mapped into the *image plane* via this matrix. The camera's position in the world frame Σ_w is represented by the extrinsic parameters, R and t. Otherwise, the focal length f and optical

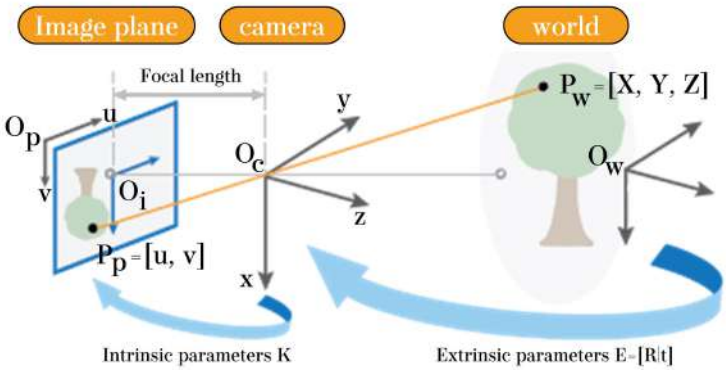


Figure 4.12.: Pinhole camera model: geometric model relationships.

center of the camera are represented by the intrinsic parameters, K. Below the geometric model details.

Using similar triangles as:

$$\frac{f}{Z} = -\frac{u}{X} = -\frac{v}{Y} \tag{4.4}$$

which can be rewritten as:

$$u = -\frac{f}{Z}X \tag{4.5}$$

$$v = -\frac{f}{Z}Y \tag{4.6}$$

Given a point in world space $P_w = [X, Y, Z]^T$, the homogeneous coordinates of P_w can be written as $\tilde{P}_w = [X, Y, Z, 1]^T$ while $\tilde{P}_p = [u, v, 1]^T$ denotes the homogeneous coordinates of $P_p = [u, v]^T$, where P_p is the correspondent point of P_w into the image plane.

Using homogeneous coordinates and matrix notation, the Equation 4.5 can

be written as:

$$Z\begin{bmatrix}u\\v\\1\end{bmatrix} = \begin{bmatrix}-f_x\\-f_y\\Z\end{bmatrix} = \begin{bmatrix}-f & 0 & 0 & 0\\0 & -f & 0 & 0\\0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}X\\Y\\Z\end{bmatrix} = P\begin{bmatrix}X\\Y\\Z\end{bmatrix} \tag{4.7}$$

P is the so-called *camera matrix*.

The above is the geometric model of a camera with the assumptions described at the beginning of the section. In the real case, however, the shape and size of the pixels and the position of the image plane relative to the optical center must be considered. If the origin of the 2D image coordinate system does not coincide with the z-axis that intersects the image plane, you must perform a translation defined by (t_u, t_v). Furthermore, a rescaling of the u and v axes is defined through the parameters m_u and m_v corresponding to the inverse of the pixel size. Thus:

$$u = m_u\frac{-fX}{Z} + m_u t_u \tag{4.8}$$

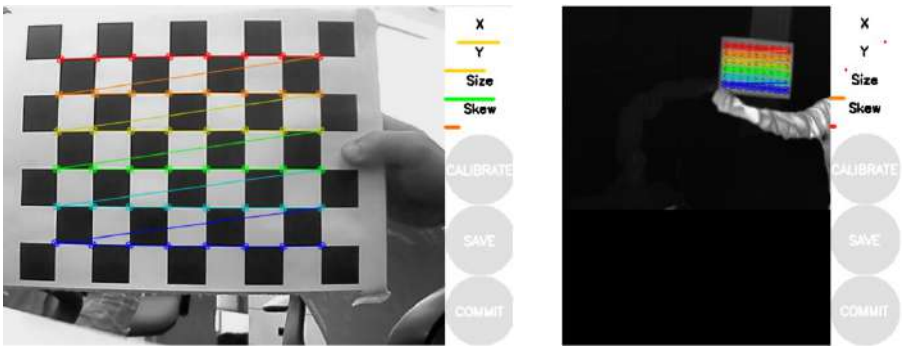
$$v = m_v\frac{-fY}{Z} + m_v t_v \tag{4.9}$$

They can be expressed in matrix form as:

$$Z\begin{bmatrix}u\\v\\1\end{bmatrix} = \begin{bmatrix}-m_u f & 0 & m_u t_u & 0\\0 & -m_v f & m_v t_v & 0\\0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}X\\Y\\Z\end{bmatrix} = \begin{bmatrix}a_x & 0 & u_0 & 0\\0 & a_y & v_0 & 0\\0 & 0 & 1 & 0\end{bmatrix}\tilde{P} = K[I|0] \tag{4.10}$$

where K only depends on the intrinsic camera parameters like its focal length, principal axis and thus defines the intrinsic parameters of the camera.

$$K = \begin{bmatrix}a_x & 0 & u_0\\0 & a_y & v_0\\0 & 0 & 1\end{bmatrix} \tag{4.11}$$



(a) (b)

Figure 4.13.: Intrinsic calibration of the multimodal vision system: the chessboard pattern adopted for the depth camera calibration (a) and the heated perforated plastic grid with a circular pattern used for the thermographic camera calibration (b).

In general, the reference system of the camera, Σ_c, is in a different position from the world reference system, Σ_w, thus a rigid transformation that connects the two reference systems made up of a rotation R and a translation t must be introduced. The extrinsic parameters E represent the coordinate transformation between the reference frame and the sensor frame.

$$t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}, R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, E = [R|t] \tag{4.12}$$

2. Intrinsic calibration of the cameras

An open-access ROS program [102] has been used to perform the intrinsic calibration of the two cameras. The software generates a *.yaml* file with intrinsic parameters that are compatible with the ROS camera drivers. To use this tool it is needed a table with a certain pattern that is clearly distinguishable from the camera: for the intrinsic calibration of the depth camera, the chessboard pattern shown in Figure 4.13a has been used. On the other hand, a perforated plastic grid with a circular pattern has been used for the thermal camera calibration: the grid has been properly heated to be clearly distinguishable from the thermal image, as shown in Figure 4.13b.

The pattern set-up (column count, raw count, actual size, and pattern type,

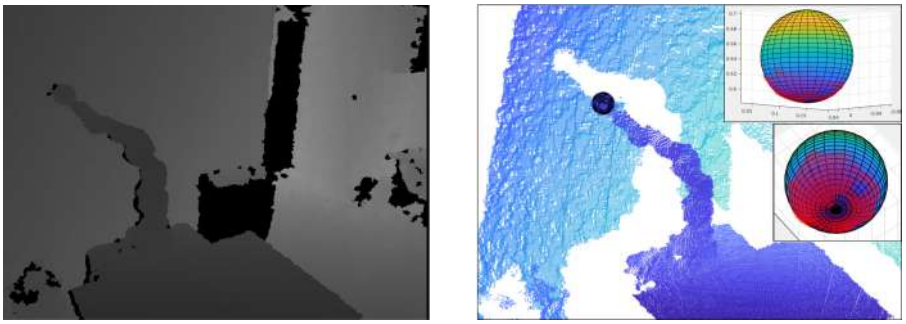
such as square or circular) is required by the software before starting the calibration procedure. After identifying the pattern, the operator moves in both linear and angular directions the pattern to enhance the calibration parameter estimation.

3. Extrinsic calibration of the depth camera

The first one procedure developed for this research activity aims at localizing the depth camera pose with respect to the robot base frame, T_c^b. The literature addresses this problem with a variety of calibrating strategies, especially in object recognition applications. Identifying objects that are around 0.5 meters away from the camera frame is their normal objective. On the contrary, in a typical industrial scenario, the robot and its operators operate at a distance of about 2.5 meters from the camera. The use of sphere tracking, as will be discussed below, in the depth camera extrinsic calibration process is a novel approach to achieving the necessary accuracy.

The 3D tracking method involves a 0.12 m diameter polystyrene sphere. To align the sphere center with the end-effector reference frame origin, the red sphere has been mounted at the robot end effector as shown in Figure 4.14a. After that, the robot has been positioned at different, specific configurations that allow to accurately identify the target sphere within the field of view of the camera. At each configuration, reading the robot joint values, the robot forward kinematics quickly calculates the robot end effector position (which corresponds to the red sphere center position) with respect to the robot base frame, P_i^b. Concurrently, the strategy obtains the depth image (Figure 4.14a) and transforms it into point-cloud data (PCD) [103] (Figure 4.14b) in which the software finds a sphere that fulfills a radius dimension constraint and gives its geometric model by using the M-estimator SAmple Consensus (MSAC) algorithm [95], which is an extension of the well-known RANdom SAmple Consensus (RANSAC) algorithm [96] (red points inliers of Figure 4.14b). The algorithm makes an approximation of the intended spheric model and computes the red sphere center position with respect to the depth camera reference frame, P_i^d. The process is iterated for different robot configurations to minimize calibration error and cover the full collaborative area. Finally, the resulting corresponding couples of points have been analyzed and an optimization technique of a cost function that integrates the collected data has been used to evaluate the transformation matrix T_c^b between the robot frame Σ_b and the depth camera frame Σ_c.

The pseudo-code of the procedure is reported in Algorithm 1. *pcfitsphere* is the MATLAB function that allows to compute a geometric model that describes the red sphere: it returns the center position and the radius of



(a) (b)

Figure 4.14.: Depth camera extrinsic calibration procedure: the original depth images (a) is converted in point clouds to allow MSAC to estimate the sphere model pose by finding correspondences, i.e. red points inliers (b).

the sphere. In order to use the *pcfitsphere* function, it is necessary to provide the ROI, thus the function *pcseg* has been applied on the original point cloud. As previously said, the resulting related couples of points have been examined, and the transformation matrix T_c^b has been assessed using the optimization function *fmin_ optimizer* using a cost function, named *cost_function* in Algorithm, that integrates the data gathered.

4. Extrinsic calibration of the thermal camera

After the depth camera calibration, the extrinsic calibration of the thermal camera with respect to the depth camera has been solved. In [100] and [101] a thermal camera and a depth camera are calibrated by using a perforated grid placed in front of the sensors. However, the proposed solutions propose the calibration approach with the only purpose of solving it, without a suitable contextualization or not considering the applicability at all in a specific scenario, basing on fairly conservative assumptions.

The first experimental setup of this work is shown in Figure 4.15 and consists of two cameras rigidly attached together through two cross zip ties. They have been arranged in a way that their optical axes were aligned. The adopted cameras have different FOVs and this implies that some depth pixels are outside the thermal image which is smaller (thus, they will not used in the merging step of Section).

The goal of the extrinsic calibration of the system is to obtain an accurate identification of the relative camera poses, which guarantee the minimum

Algorithm 1 T_c^b computation

Input:

$r_n \leftarrow$ Nominal sphere radius ; $r_e \leftarrow$ Estimated sphere radius

$\Delta r \leftarrow$ Acceptable radius displacement ; $ROI \leftarrow$ Region Of Interest

$C \leftarrow$ Number of acquired samples for each robot configuration

$PC_o \leftarrow$ Original point cloud ; $PC_s \leftarrow$ Segmented point cloud

$acc_i \leftarrow$ Accumulator matrix of estimated sphere models of j-th sample

$P_i^d \leftarrow i$-th target point expressed in depth frame

$q_i \leftarrow i$-th robot configuration in the joints space

$Q \leftarrow$ number of different robot joint configurations

$T_{init} \leftarrow$ T_c^b initial estimation

Output:

$T_c^b \leftarrow$ depth camera reference frame expressed in robot base frame

1: **procedure** T_c^b ESTIMATION
2: **for** $i = 1$ to $size(Q)$ **do**
3: **for** $j = 1$ to $size(C)$ **do**
4: $PC_o = pcrec$
5: $PC_s = pcseg(ROI, PC_o)$
6: $[p, r_e] = pcfitspehere(PC_s)$;
7: **if** $|r_e - r_d| <= \Delta r$ **then**
8: $acc_i[j, 1:3] = p; acc_i[j, 4] = r_e$
9: **end if**
10: **end for**
11: $P_i^d = mean(acc_i[:, 1:3])$
12: $K = forward\_kinematic(q_i)$
13: $P_i^b = K.transpose$
14: **end for**
15: $CF = cost\_function(T_{init}, P^b)$
16: $T_c^b = fmin\_optimizer(CF, T_{init}, P^d, P^b, K)$
17: **end procedure**

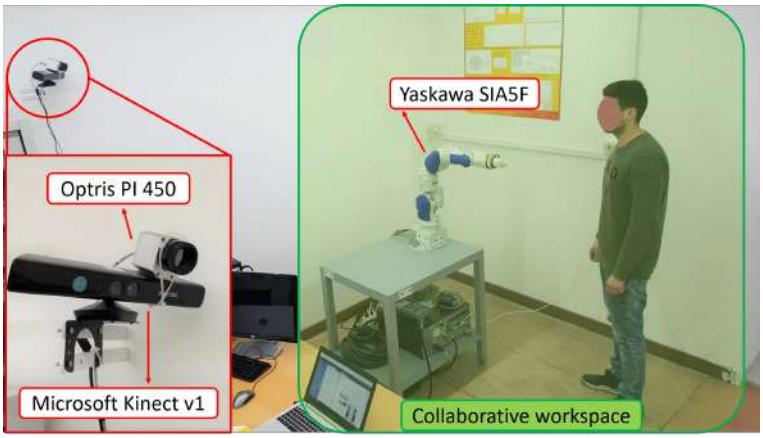


Figure 4.15.: The experimental perception system composed of a depth camera (Microsoft Kinect v1) and a thermal camera (Optris PI 450).

accuracy error when the two camera views are merged. To obtain an estimation of the transformation matrix T_t^d, between the depth camera frame Σ_d and the thermal camera frame Σ_t, three spheres have been attached to a flat cardboard support and heated to be distinguishable by both the depth and the thermal cameras. The pattern has been moved inside the collaborative workspace in 10 different configurations at distances from the cameras in the range where the human operator is expected to act during the collaborative task. At every acquisition, the calibration target has been suitably heated to be detectable from both cameras. The coordinates $p_k^d = [x_k^d, y_k^d, z_k^d]^T$ of the k-th center of the target sphere have been directly calculated from the depth image, while the corresponding thermal point coordinates have been calculated from the thermal image, assuming the distance from the lens equal to the depth value, $z_k^d = z_k^t$ and

$$x_k^t = \frac{(a_k - c_{x_t}) z_k^t}{f_{x_t}} \tag{4.13}$$

$$y_k^t = \frac{(b_k - c_{y_t}) z_k^t}{f_{y_t}} \tag{4.14}$$

where a_k and b_k are the pixel coordinates of the sphere center in the thermal image, c_{x_t} and c_{y_t} are the pixel coordinates of the thermal image center and f_{x_t} and f_{y_t} are the focal lengths expressed in pixel-related units. Finally, the

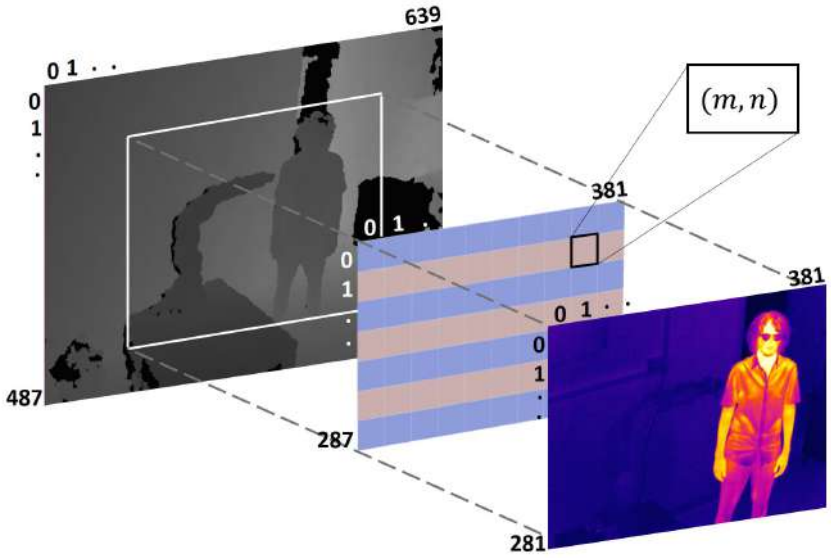


Figure 4.16.: The mapping algorithm which merges depth and thermal images finding pixel-by-pixel matches between the two images: the result is a 382×288 matrix.

transformation matrix T_t^d has been estimated by minimizing a cost function that combines the corresponding data with an optimization function similar to the one presented in Algorithm 1.

5. Depth-Thermal mapping matrix

Preserving all accurate and valuable data from the sources to be merged while avoiding distortions is the primary prerequisite for a sensor fusion strategy. Using a mapping matrix procedure, the depth image and the corresponding thermal image have been combined for the purposes of this work to create an improved single view of a scene with more information content.

The extrinsic calibrations explained above are a first step towards a correct mapping, that means finding matches between the depth image and the thermal image. Since the adopted cameras have different FOVs and resolutions, the resulting map size must correspond the mapping step builds a 382×288 matrix, the size of the thermal image shown in Figure 4.16.

The mapping step has been solved through a pixel-by-pixel procedure: the pixel of the depth image, of indices (m, n), contains the depth value, $z_{m,n}^d$,

which can be acquired to compute the corresponding Cartesian point coordinates $p^d_{m,n} = [x^d_{m,n}, y^d_{m,n}, z^d_{m,n}]^T$, by using the formulas similar to Equations 4.13 and 4.14:

$$x^d_{m,n} = \frac{(m - c_{x_d}) z^d_{m,n}}{f_{x_d}} \tag{4.15}$$

$$y^d_{m,n} = \frac{(n - c_{y_d}) z^d_{m,n}}{f_{y_d}}. \tag{4.16}$$

where c_{x_d} and c_{y_d} are the pixel coordinates of the center (on the focal axis) in x and y directions of the image plane, f is the focal length of the depth camera.

The Cartesian point is then expressed with reference to the thermal camera frame through the relation:

$$\tilde{p}^t_{m,n} = \begin{bmatrix} p^t_{m,n} \\ 1 \end{bmatrix} = T^t_d \begin{bmatrix} p^d_{m,n} \\ 1 \end{bmatrix}. \tag{4.17}$$

from which $p^t_{m,n}$ is immediately obtained from the first three components of $\tilde{p}^t_{m,n}$.

Using the intrinsic parameters of the thermal camera, the corresponding pixel indices of the point $p^t_{m,n}$ into the thermal image (a, b) are finally computed by inverting (4.13)-(4.14). If they are contained in the FOV of the thermal image, the corresponding depth pixel indices (m, n) are written into the mapping matrix at the indices (a, b); otherwise, they are discarded because they are outside the mapping image size. Note that, if the observed object is far enough, the mapping matrix does not depend on the distance. Thus the mapping matrix can be computed offline by using a fixed value of $z^d_{m,n}$ compatible with the working area, hence saving computational load.

6. The proposed *RGB Mapping Approach*

Different image fusion algorithms can be used to merge multi-modality sensor images. They operate at different merging levels: pixel-by-pixel, combining signals, using relevant features or at symbol levels. This section offers a pixel-level image fusion technique that, while similar to the most popular pixel-level image fusion algorithms, takes a novel approach.

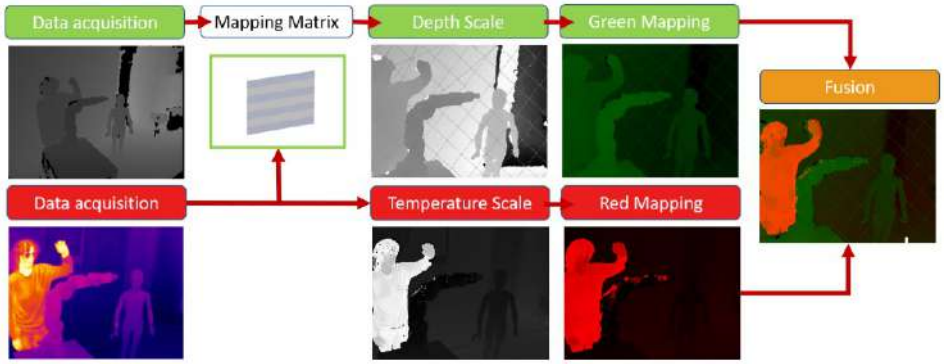


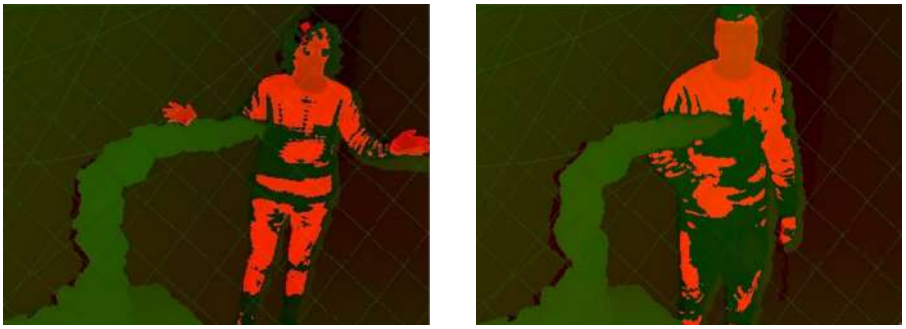
Figure 4.17.: Depth image and thermal image mapping pipeline.

The proposed approach, callable *RGB Mapping Approach* (RGB-MA), consists in defining the intensities of empty RGB channels. This is also because, to the best of the author knowledge, CNNs work better with RGB images. RGB-MA strength is that it assigns the same priority to the input sources, both depth and thermal data. The result is no longer a grayscale, as a depth image alone could be, or a weighted average image which assigns different priorities to the sources, but it is an RGB image where the depth data have been mapped on the green channel and the temperature values have been mapped on the red channel, as shown in Figure 4.17.

Specifically, the original depth sensor value, s^d, and the corresponding temperature sensor value, s^t, have to be normalized into the interval [0, 1]. To do this, a *minimum* and a *maximum* variability ranges of the source values have been defined for both thermal, min_t, max_t, and depth, min_d, max_d, cameras. They do not actually correspond to the ranges of the sensors technical specifications, but they have been chosen according to the values detectable into the considered workspace. More in detail, the detectable depth values are included between 0.30 m and 4.0 m, while the detectable temperature values are within the range $[0, 50]^\circ$ C, which are suitable for any type of human detection task.

The color information inserted into the specific channel of the *(i,j)-th* pixel of the output image must be mapped to 8 bits. The R (red) value is computed, by acquiring $s^t_{i,j}$ from the thermal image, as

$$R_{i,j} = \text{round}\left(255\frac{s^t_{i,j} - min_t}{max_t - min_t}\right); \tag{4.18}$$



(a) (b)

Figure 4.18.: Two examples of the resulting image of the depth-thermal sensor fusion.

the G (green) value is computed by acquiring $s^d_{m,n}$ from the depth image, where m and n are contained into the *(i,j)-th* value of the mapping matrix,

$$G_{i,j} = \text{round}\left(255\frac{s^d_{m,n} - min_d}{max_d - min_d}\right); \tag{4.19}$$

the B (blue) value of the resulting image is always zero:

$$B_{i,j} = 0; \tag{4.20}$$

The result is shown in the last picture of Figure 4.17 and in Figure 4.18. Note that the proposed image fusion technique leaves another channel that could be used for a further input source.

4.6.4. DT-CNN for Human Detection

Nowadays there are many architectures and framework with a pre-trained CNN that allows to recognize some categories of objects and these networks are generally trained on the RGB images, so they can be used just on the traditional RGB cameras.

To assess the effectiveness of the fused images described in the previous section, YOLOv3 [104] has been used for real-time human detection. The selected framework is an open source neural network framework written in C and CUDA that supports CPU and GPU libraries and it is fully compatible

with ROS environment. YOLOv3 is an off-the-shelf SoA 2D object detector pre-trained on ImageNet [105] and fine-tuned on the MS-Coco [106] data-set. It is an extremely fast and accurate object detection system, which is born to detect semantic objects of a certain class, e.g., humans, buildings and cars. In general, YOLOv3 is trained by using RGB images and there are many pre-trained weights (available online) to test the detection on the topic of the RGB camera.

On the other hand, nowadays there are no neural networks which have been trained on combined images such as those proposed by this research activity, so the YOLOv3 CNN model has been re-trained to adapt the detection system to the Depth-Thermal (DT) images. In practice, after the definition of a *Human class*, the YOLOv3 pre-trained classes have been removed from the prediction and a new training data-set has been developed. Finally, YOLOv3 CNN weights have been correctly retrained. Details are described below.

Training dataset building

The Yolo training dataset is composed by a set of properly labeled images. In practice, the training dataset is a folder containing N images and N corresponding text files. More in detail, for each image (i.e., named *image_ name.jpg*) there is a corresponding text file (named *image_ name.txt*) which contains two data: 1. the label (category) of the object present inside and 2. the coordinates of the (manually drawn) bounding boxes containing the object itself.

In order to create the Yolo training dataset, a video stream of the sensor fusion frames have been saved, recording the experimental collaborative workspace and moving inside objects and different human operators, simulating a short film of a real collaborative scenario. After that, a Python program has been used to automatically saves the frames of the video itself with a 2Hz frequency. Once the images have been saved, the available software *YOLO-Annotation-Tool* has been used to manually label the images, as shown in Figure 4.19. The software allows the user to load an image and to manually draw the bounding box on it around the object to be labelled. The output of the tool is the text file of the corresponding image. Note that for the experimental validation of the software, one thousand images have been manually labelled image (this step has been the longest time consuming operation: it needed about one hour).

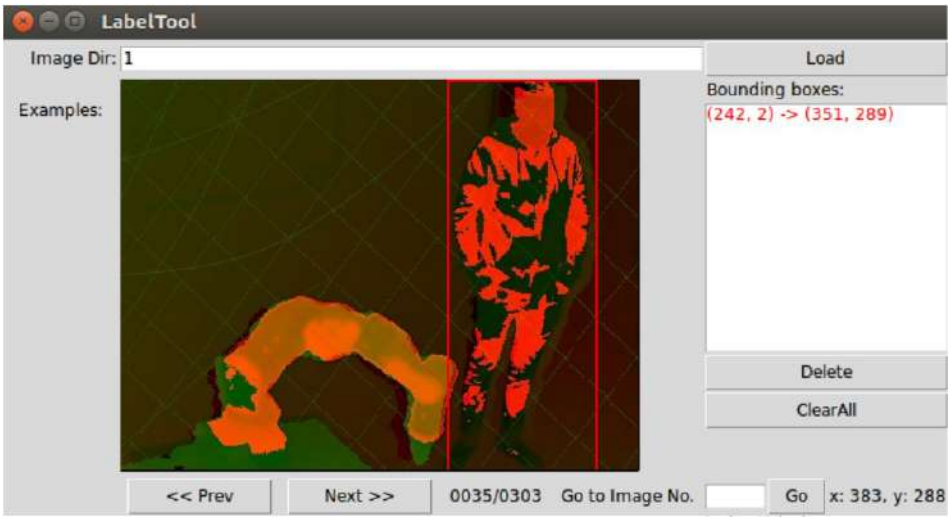


Figure 4.19.: YOLO-Annotation-Tool during the manual labelling operation.

Pre-training configurations

Once the dataset has been built, the following steps have been performed:

- the labelled images of the dataset have been divided in two groups to build the *train.txt* and *test.txt* files, containing the train (80% of dataset) and test (20% of dataset) image paths respectively;
- configuration of the *cfg/obj.data* file, which specifies the number of classes to train, as well as the path of the *train.txt* and *test.txt* files;
- configuration of *cfg/obj.names* file, which lists the names of the categories to be trained (every new category should be on a new line, its line number should match the category number in the *image_ name.txt* label files): for the case study, the only defined category is *HUMAN*;
- configuration of *cfg/yolov3.cfg* file which contains the convolutional neural network model. Default Yolov3 model has been used as a model but batch size and subdivision parameters have been modified. The batch size indicates the number of images used for each training step while the subdivision indicates the number according to which the batch will be divided to decrease GPU VRAM requirements.

At this point it has been possible to send all these files to Yolo to start the weights retraining.

Training step

The Titan V, a special GPU, has been used for the training phase since it can do tasks up to ten times more quickly than a CPU. Every hundred iterations, the trained weights are saved in a backup file and the training continues without ever pausing. The framework publishes a number of parameters during the training phase, allowing the user to choose whether or not to halt it. Here are the published parameters:

- *Region Avg IOU* is the average of the IOU of every image in the curren subdivision. IOU is a metric described in detail in next subsection. At 100% we have a perfect overlap of the manually drawn bounding box and the one selected by the CNN model. A good value to stop the training is around 80%.
- *Avg Recall* is a metric which counts how many positives Yolo detected with respect to the total ones present into the dataset.
- *Count* is the amount of positives (objects to be detected) present in the current subdivision of images.
- *Obj* is the average confidence score on those grids where there is an object;
- *No Obj* is the average confidence score for the locations where there's no object.

Test step

To quantify the performance of the training results, another small data set, called *test set*, has been used to test the model retrained weights. This performance can be measured by using one statistics methods such as accuracy, precision, recall. Depending on the specific application, one metric may be used instead of another. To prove the performance of the training set in the experimental use case published in [107], the Mean Average Precision (maP) has been adopted. The maP is specifically chosen for algorithms that tackle object *detection* problems, which predict the position of the item along with the classes, in contrast to *classification* problems for which simpler precision and recall metrics are advised.

The mAP computation merges different metrics such as precision, recall and IoU.

- Intersection Over Union (IoU) is a very useful metric which determines

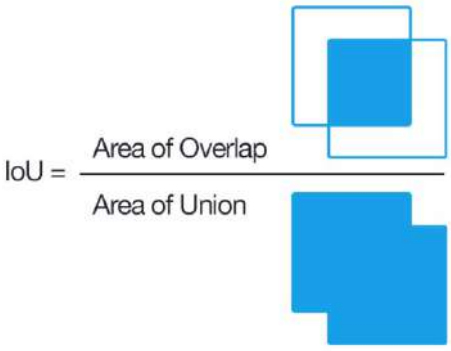


Figure 4.20.: IoU computation.

how accurately the model has detected a certain object. Figure 4.20 shows the IoU computation: it consists in the division of the area of overlap between the bounding boxes by the area of union. In practice, two bounding boxes *metches* when they share the same label and an the $IoU \geq 0.5$.

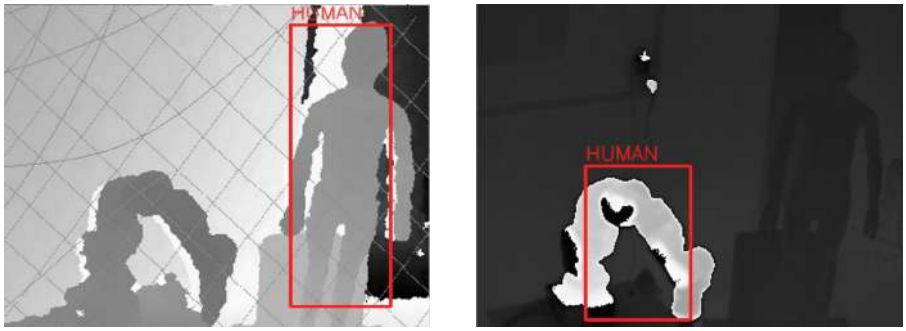
- Precision, also called positive predictive value, is the fraction of relevant instances among the retrieved instances.
- Recall, also known as sensitivity, is the fraction of relevant instances that have been retrieved over the total amount of relevant instances.

Finding the area under the precision-recall graph by substituting the greatest precision for any recall for the precision value is conceptually similar to mAP. The percentage of false detections was estimated as an extra statistic because the mAP does not account for false positives (FP) and false negatives (FN).

The idea of mAP is conceptually similar in finding the area under the precision-recall graph replacing the precision value with the maximum precision for any recall. Since the mAP does not take false positives and negatives into account, the percentage of the false detected was estimated as an additional metric.

Experimental Results

Two alternative single-source CNNs, a D-CNN based only on depth images and a T-CNN based only on thermal images, have been trained in order to evaluate, through comparison, the performance of the proposed DT-CNN



(a) (b)

Figure 4.21.: Two cases of human FP: a dummy labelled as *HUMAN* by the D-CNN (a) and a hot moving robot labelled as *HUMAN* by the T-CNN.

which uses the innovative sensor fusion approach suggested by this study work. Note that the identical input data set has been used for training and testing the two CNNs: the only difference was about the source data.

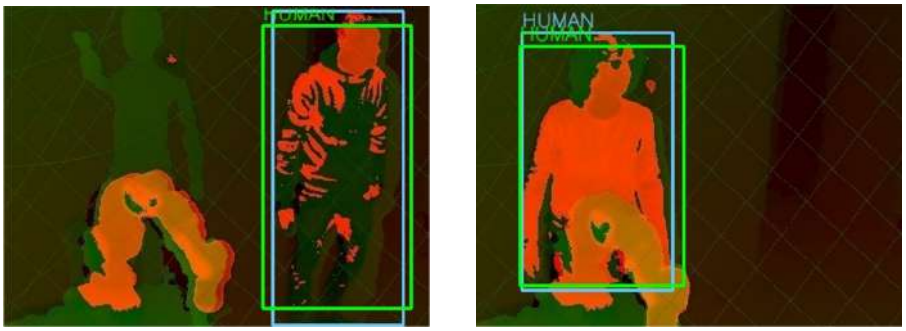
Table 4.3.: Optris PI 450 specifications.

Approach	**mAP**	**FP**	**FN**
D-CNN	65.09%	26.87%	17.53%
T-CNN	62.76%	64.35%	4.37%
DT-CNN	50.54%	7.47%	11.85%

It has been demonstrated by comparing the Table 4.3 results that the mAP in the DT-CNN is comparable to that of the other approaches. FP's percentage is far smaller than the others, though. The percentages of FN are low, in general.

More in detail, Figure 4.21 illustrates two cases of FP of single-source techniques: the second picture represents the heated robot which has temperature similat to the human body surface and it has been confused with a human operator by the T-CNN, while the first picture represents a plastic mannequin with human-like shapes which can be incorrectly labelled as human by the single-source D-CNN.

Otherwise, as shown in Figure 4.22, by using a DT-CNN there is a good overlap between the predicted and the ground truth boxes and also both the hot robot and the presence of the plastic dummy do not cause false



(a) (b)

Figure 4.22.: Ground-truth and predicted boxes overlapping in sensor fusion based application.

positives.

As a result, one significant outcome of this suggested method is that, with a training dataset of just 1000 samples, the DT-CNN can detect any human operator with good accuracy and a very low percentage of FP and FN.

5. Real-time workspace monitoring

CNN for human detection has a wide range of applications, such as enhancing safety systems and surveillance systems. This chapter describes how to use a multimodal perception system and the developed DT-CNN to robustly detect human operators in collaborative robot work cells based on the *Speed and Separation Monitoring* concept. In fact, the human prediction has been obtained by applying the DT-CNN to the real-time DT video stream. For every human detected in the scenario, an estimated bounding box is created and its coordinates are ready to be processed in the collaborative pipeline's subsequent stages. Note that the training and prediction processes require a significant amount of processing power and have been carried out on a suitable GPU (NVIDIA Titan V). The robot's separation distance from the nearest human operator is calculated by the system in real-time and it will be used by the next robot control step to be compliant with the current regulations to ensure the safety of human operators by adjusting the robot motion speed.

5.1. Background segmentation

The basic idea of the developed pipeline was to refer to the *SSM* collaborative scenario in which the speed of the robot is modulated according to the separation distance between the robot itself and the human operator who is closest to it, as detailed in Chapter 3. Therefore, the next step after human detection is the aforementioned separation distance computation. The technique adopted in this research is a point-cloud-based technique which was found to be the most suitable with respect to the other SoA strategies proposed in literature and described in Chapter 4.

Due to the fact that any point-cloud-based approach always entails a computationally demanding process, the first assumption is to handle only the data pertaining to the dynamic objects that are part of the scene under observation. Figure 5.1 shows the initial *Background Segmentation* phase which removes the static environment and leaves active only the pixels of the dynamic entities, then transforms them into point clouds.

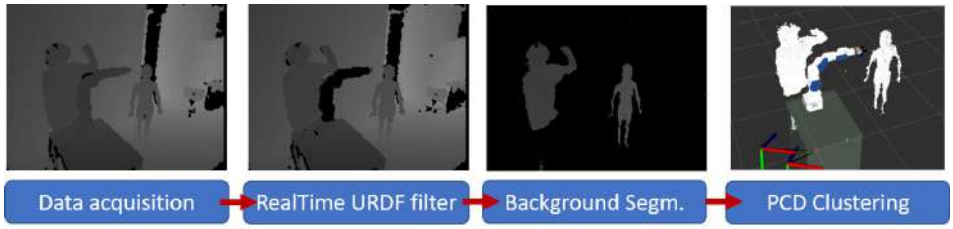


Figure 5.1.: *Background segmentation* step of the collaborative pipeline.

To be more specific, take note of the fact that the robot kinematic chain is fully visible from the multimodal vision system view, becoming a dynamic entity too. Consequently, at the beginning of the step, the package *Real-time URDF Filter* [108] has been executed to identify the depth pixels that belong to the robot model and to assign them a Not-a-Number (NaN) value. An efficient algorithm that subtracts a stored background at the pixel level is used to develop the background filtering. First of all, 50 frames of the static background have been recorded and the average value of each pixel is kept in a memory area to create the static background model frame that will be subtracted from the dynamic frames, in real time. After converting the depth image to a PCD, the cloud density can be chosen to apply a uniform sampling filter which will increase the algorithm's reactivity. Finally, a *PCD Clustering* step is used to detect dynamic entities: it analyzes the point-cloud image and generates a number of clusters based on the number of single dynamic areas that are detected in the foreground. To distinguish each cluster into the collaborative workspace, the *Euclidean cluster extraction* method is used. It should be noted that the areas in the foreground should be large enough to depict a human body in order to compensate for the measurement noise of the sensors, which occasionally could yield misleading clusters. Consequently, an experimental minimum PCD cardinality threshold has been established to discriminate the validity of the cluster.

5.2. Human validation check

The *Euclidean cluster extraction* method distinguishes the dynamic entities into single clusters moving into the collaborative workspace. In the same time, the DT-CNN is running. When a *HUMAN* is detected, the DT-CNN returns the predicted bounding boxes that contain him. These coordinates are sent to the segmentation pipeline to verify the human cluster through the *Human Validation* step.

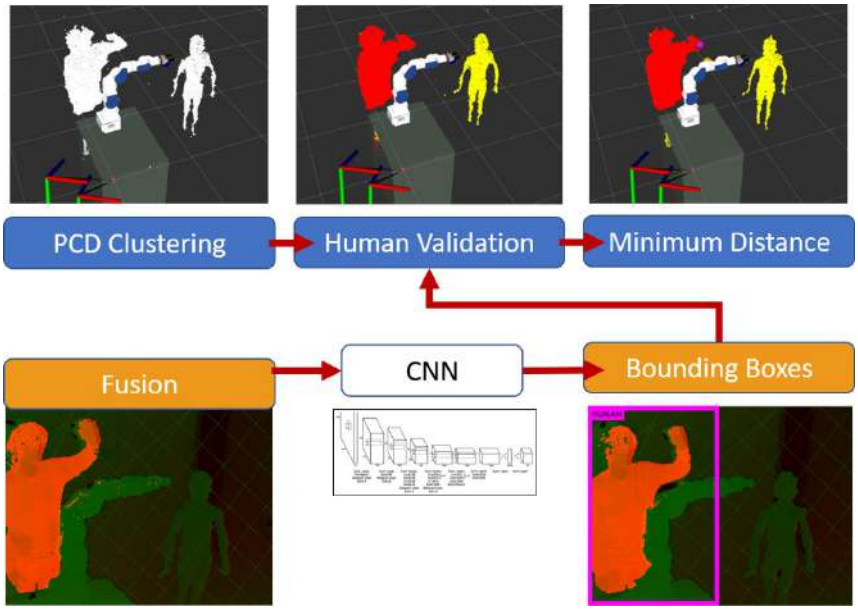


Figure 5.2.: *Human validation* step of the collaborative pipeline.

More in detail, Equations 4.15, 4.16 are inverted, thus each point of the cluster is transformed into the corresponding depth pixel coordinates. Since the bounding box coordinates coming from the DT-CNN detection is expressed in the thermal image plane, the selected pixels are converted into the depth image pixel coordinates through the mapping matrix described in Section 4.6. The pixels of the cluster are checked to be contained in the pixel box of the depth image. The cluster is labelled as *HUMAN* and passes the check if at least 50% of the cluster points are contained within the DT-CNN bounding box. Figure 5.2 displays an example of two clusters: the yellow plastic mannequin, which is correctly not identified as human, and the red human operator, who is correctly detected by the DT-CNN. *Human Validation* check is an essential step to compute the subsequent separation distance between human operators and the robot.

5.3. Human-Robot real time data processing

Prior to calculating the pints at minimum distance, a special integration of the human-robot separation distance computation has been to handle a thermal point cloud: the clusters are transformed into *RGB-D point clouds*, which contain both spatial and temperature data. Next, by calculating the *Human-Robot separation distance*, a novel strategy finds the points at

the shortest distance. Finally, a *Human-Robot speed estimation* process is created to determine the real-time speeds of the robot and the nearest human operator.

5.3.1. Thermal point-clouds

The main idea in this step is that not all the points of a cluster could actually belong to the body surface. When the worker holds a tool for his manual operation, the tool is also included into the identified human cluster. Thermal information can distinguish the points representing the tool from points belonging to the human body surface inside the single cluster.

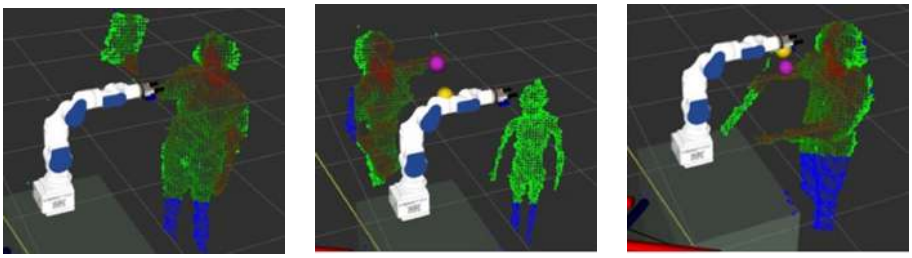
This consideration led the research to use a thermal point cloud: every point of the cluster is represented not only by its position but also by a RGB color depending on its own temperature. The color scale is between green (temperature value below or equal to 25° C) and red (temperature value above or equal to 40° C), as shown in Figure 5.3. When the point at minimum distance from the robot is identified, the purple sphere of Figure 5.3b/c contains also its temperature value. This information is sent to the fuzzy inference system (Chapter 6), which distinguishes if the point actually belongs to the body surface or to an handheld tool. In the latter scenario, a little reduction of the minimum protective distance S_{ISO} can be applied in the control strategy in order to improve production time of the workcell because an impact of the robot with an handheld tool is considered potentially less hazardous to the worker than a direct impact with the body surface. A collision is still not permitted, but the robot's speed is reduced less if the point at the minimum distance is not part of the surface of the human body.

5.3.2. HR separation distance computation

At this point, the algorithm has all the inputs available for the identification of the nearest pair of points, one belong to the robot (P_R) and the other one belonging to the operator (P_H), that minimize the distance, i.e.,

$$P_H \in \mathcal{H}, P_R \in \mathcal{R} \mid d(P_H, P_R) \leq d(P'_H, P'_R) \quad \forall P'_H \in \mathcal{H}, P'_R \in \mathcal{R} \tag{5.1}$$

where $d(\cdot, \cdot)$ is the Euclidean distance between two points, $\mathcal{H}$ and $\mathcal{R}$ rep-



(a) (b) (c)

Figure 5.3.: RGB-D point cloud: the thermal point cloud is composed by points associated not only to the depth value but also to the own temperature information whcih determine the point color (a). These data distinguish if the purple point belonging to the human cluster and at minimum distance with respect to the yellow robot point properly belongs to the human body surface or not (b)(c).

resent the set of all points that belong to the human operator and to the robot, respectively.

A robot model has been devised to take into account its physical volume and to correctly compute the separation distance d. As stated in the SoA assumptions described in Chapter 4, just one representative coordinate of the robot (such as the end effector) is taken into consideration, or geometries that are not accurate representations of the robot's kinematic chain -like spheres and ellipses- are used in literature to represent the robot model, [109] and [110]. The distance between the operators and the robot kinematic chain is computed inefficiently by selecting one of these strategies.

The same concept served as the inspiration for the modeling method proposed in this research work: in order to protect the robot uniformly, the kinematic chain of the robot has been padded through 12 dummy frames because the robot links may have variable lengths. Additionally, a 0.10 m diameter security sphere has been created around each frame, taking into account the final end-effector frame that may contain a tool or an attached object (during a pick and place task, for example). In this case, the robot model covers the entire robot's kinematic chain and volume.

A point-cloud-based approach may quickly identify the pair of human-robot points that are closest to one another. In fact, compared to the previously mentioned suggested ways, the point cloud offers far more precise, accurate and comprehensive information content. Furthermore, in contrast to

popular skeleton-based methods, the suggested strategy enables tracking of people even when they are carrying objects. Moreover, human operators do not need to be in front of the camera view in order to be recognized and it is very fast to identify the two human-robot points at the lowest separation distance (Equation 5.1). The algorithm calculates the distance between all points of a clusters point cloud and the origin of every robot frame. Eventually, the robot point P_R will be the one on the surface of the virtual sphere, around the identified frame, which lies on the line connecting the origin of this frame and the closest point in the cluster. From these results, the closest human cluster is indirectly selected if more than one human have been detected. The pseudo-code of the procedure is reported in Algorithm 2.

Figure 5.4 shows the results. The proposed approach is able to identify more detailed body parts that are exposed to the robot, e.g., a elbow, the head, an hand, the chin or the chest (the purple sphere of Figure 5.4), as well as that P_R can be detected along the whole robot kinematic chain (the yellow sphere of Figure 5.4). Note that the developed DT-CNN distinguishes human operators belong to the Human class (Figure 5.4b, red clusters) from other clustered objects (Figure 5.4b, yellow clusters), i.e. a plastic mannequin and a chair, which are not labelled as *HUMAN* and they are not considered for the safety separation distance computation, even if they can possibly be closer to the robot. The algorithm effectiveness in multi-humans scenarios has been proved in Figure 5.4b.

5.3.3. Intention estimation

Intention estimate (IE), or the prediction of human movement, is another crucial element of the HRC problem in addition to the HDT technique. IE uses a set of previously acquired locations to estimate the future position and velocity of the trajectory that the human operator will conduct. The robot control system will use these data to calculate the ideal joint speed value, which is established for SSM collaboration scenarios, in order to avoid potentially dangerous scenarios. This section explains how to compute the operator and robot actual velocities, while the upcoming Chapter will use these parameters to go into great detail on the robot control strategy and to describe how to take advantage from these analysis for industrial collaborative applications with the aim to maximize productivity.

Algorithm 2 d_s computation and closest cluster identification

Input:

$S \leftarrow$ Original depth image frame

$S^0 \leftarrow$ Static background

$PCD \leftarrow$ Point cloud data corresponding to depth image S

$R_m \leftarrow$ URDF robot model

$clusters \leftarrow$ PCD clusters extracted from the current scene

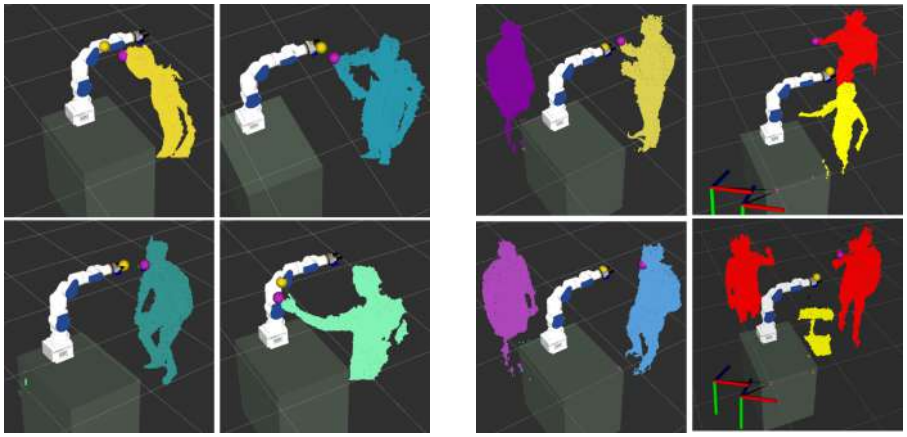
$acc \leftarrow$ Accumulator matrix of clusters separation distances

Output:

$d_s \leftarrow$ Robot-closest operator separation distance

$cluster \leftarrow$ Closest operator cluster

procedure SEGMENTATION PIPELINE
  $bool\ first\_frame = true$
  $S = depthAcquisition() - R_m$
  if $first\_frame == true$ **then**
    $S^0 = S$
    $first\_frame = false$
  end if
  $S = S - S^0$
  $PCD = pcdConversion(S)$
  $clusters = EuclideanClusterExtraction(PCD)$
  for $i = 1$ to $clusters.size$ **do**
    $acc_i = computeMinDist(R_m, clusters_i)$
  end for
  $(d_s, cluster) = min(acc)$
end procedure



(a) (b)

Figure 5.4.: Identification of the minimum distance points between the whole robot (yellow sphere) and the closest human operator (purple sphere) in scenarios in which there is only one human operator (a) and also in multi-humans and multi-entities scenarios (b).

Estimation of the human operator speed

The Linear Kalman Filter (LKF) [111], which attempts to tackle the challenge of estimating the state of a discrete-time process defined by the equations, is at the base of the IE developed method for the human operator speed estimation. The LKF problem is described by the following equations:

$$x_{k+1} = \begin{bmatrix} I_3 & \Delta t I_3 \\ O_3 & I_3 \end{bmatrix} x_k + w_k \tag{5.2}$$

$$y_k = \begin{bmatrix} I_3 & O_3 \end{bmatrix} x_k + n_k, \tag{5.3}$$

where Δt is the sampling time, I_3 and O_3 are the identity and zero matrices of size 3×3, respectively; w and n are the process and measurement noises with covariance matrices W and N, respectively. Finally, x is the state vector of the system, i.e., the position and the velocity of the operator $x = \begin{bmatrix} p_H^T & \dot{p}_H^T \end{bmatrix}^T$, and the measured output y is a vector containing the coordinates of the point P_H described in Section 5.2. The covariance matrix N is experimentally estimated, while the covariance matrix Q has been chosen as

$$Q = \begin{bmatrix} I_3 \Delta t^2 & O_3 \\ O_3 & Q_2 \end{bmatrix} \tag{5.4}$$

where Q_2 quantifies the uncertainty on the velocity dynamics (assumed constant) of the state equations. Based on the vector nature of the velocity, it is possible to make some considerations about the direction (trend) of the operator, that is to say, to predict in which direction he/she is traveling to.

The LKF equations implemented in this research work are the standard ones and they are reported in Table 5.1.

Predict	
Predicted (a priori) state estimate	$\hat{x}_{k\|k-1} = F_k \hat{x}_{k-1\|k-1} + B_k u_k$
Predicted (a priori) error covariance	$P_{k\|k-1} = F_k P_{k-1\|k-1} F_k^{\mathsf{T}} + Q_k$
Update	
Innovation or measurement pre-fit residual	$\tilde{y}_k = z_k - H_k \hat{x}_{k\|k-1}$
Innovation (or pre-fit residual) covariance	$S_k = R_k + H_k P_{k\|k-1} H_k^{\mathsf{T}}$
Optimal Kalman gain	$K_k = P_{k\|k-1} H_k^{\mathsf{T}} S_k^{-1}$
Updated (a posteriori) state estimate	$\hat{x}_{k\|k} = \hat{x}_{k\|k-1} + K_k \tilde{y}_k \hat{x}_{k\|k} = \hat{x}_{k\|k-1} + K_k \tilde{y}_k$
Updated (a posteriori) estimate covariance	$P_{k\|k} = (I - K_k H_k) P_{k\|k-1} (I - K_k H_k)^{\mathsf{T}} + K_k R_k K_k^{\mathsf{T}}$
Measurement post-fit residual	$\tilde{y}_{k\|k} = z_k - H_k \hat{x}_{k\|k}$

Table 5.1.: Linear Kalman Filter equations.

More details about the notation are described below:

- $\hat{x}_k$ is the a *posteriori* state estimate at time k given observations up to and including at time k;
- $\hat{P}_k$ is the a *posteriori* error covariance matrix;
- F_k is the state-transition model which is applied to the previous state x_{k-1};
- H_k is the observation model which maps the true state space into the observed space;

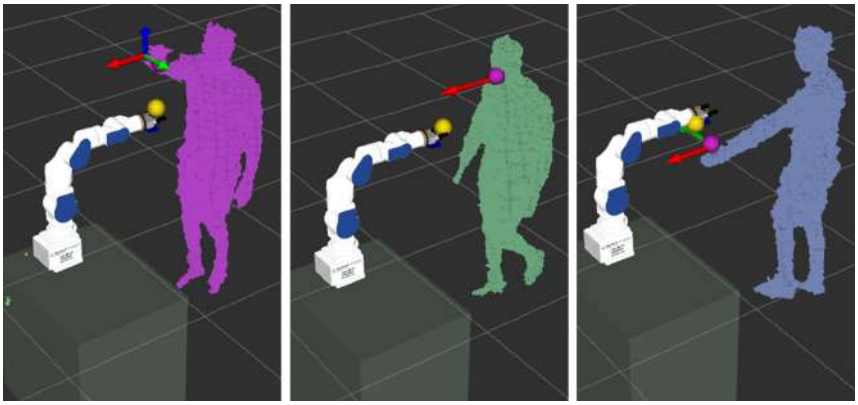


Figure 5.5.: Estimation of human operator velocity.

- Q_k is the covariance of the process noise;
- R_k is the covariance of the observation noise;
- B_k is the control-input model which is applied to the control vector u_k;

The tuned parameters are fully described in Chapter 7.

The result of the developed IE algorithm is shown in Figure 5.5, where the estimated speed of the human operator is represented through its three vector components (the adopted convention is $xyz \rightarrow RGB$) and, in particular, the length of each vector is related to the intensity of the value of the component itself.

Estimation of the robot speed

On the robot side, the intention estimation algorithm for the robot speed evaluation uses a common convention based on the knowledge of the robot joint angle values q [rad] and joint velocity values $\dot{q}$ [rad/s], read in real time. Actually, using the differential kinematics equation, the linear velocity $\dot{p}_R$ of the robot's point nearest to the operator may be calculated:

$$\dot{p}_R = J_p(q)\dot{q}, \tag{5.5}$$

where $q\,[\text{rad}]$ and $\dot{q}\,[\text{rad/s}]$ are the robot joint position and velocity vectors, respectively; J_p, on the other hand, is the position part of the Jacobian matrix calculated till the closest point.

Adaptation to the ISO15066 standard

The ISO/TS 15066 states that the "directed speeds" of the robot and the human should be used to compute the minimum separation distance S_{ISO} mentioned in Section 3.2. This means that, in Equation 3.3, v_h is the operator speed in the direction of the moving part of the robot and v_R is the robot speed in the direction of the selected operator. Note also that these speeds are vector magnitudes, hence they are always grater or equal to 0. Therefore, the velocity terms of Equation 3.3 can be computed as

$$v_H = \left| \hat{\dot{p}}_H^T \left(\frac{p_R - \hat{p}_H}{\|p_R - \hat{p}_H\|} \right) \right| \tag{5.6}$$

$$v_R = \left| \hat{\dot{p}}_R^T \left(\frac{\hat{p}_H - p_R}{\|\hat{p}_H - p_R\|} \right) \right|, \tag{5.7}$$

where $\hat{p}_H$ and $\hat{\dot{p}}_H$ are the operator position and velocity estimated by the LKF, respectively, and p_R is a vector containing the coordinates of the point P_R defined in Section 5.3.

6. Robot control for *Speed and Separation Monitoring*

Numerous methods have been offered in literature to reduce the robot speed in SSM scenarios based on specific metrics. Since the major goal is to find the best trade-off between operator safety and robotic cell productivity, defining these metrics is a crucial step. This Chapter proposes an innovative fuzzy logic approach to control an industrial robot speed by monitoring the human-robot separation distance and by taking into account a risk assessment analysis in real time. The solution generates a continuously modulated speed scaling factor which computes a dynamic minimum separation distance to be kept between the robot and the human operator, rather than a discretized value. Note that the fuzzy logic has been introduced to achieve a fluid modulation of the speed of the robot and therefore not only allow greater acceptability of the machine by the human operator into the collaborative application context, but also to optimize the productivity of the work cell, reducing the number of robot stops which are typical of a traditional logic, a deterministic logic based on predefined security areas.

6.1. Traditional Vs Fuzzy logic

The *fuzzy logic* is also called *faded logic* and it is a decision-making strategy in which each proposition represents a degree of truth into the interval $[0, 1]$ [112]. This approach was born as an alternative to the *traditional*, or *binary* or *crisp logic*, to reason with truth values that are not only 1 (true) and 0 (false). In other words, whereas in a *fuzzy logic*, the category definition intersects, allowing a particular subject to belong in variable percentages to either one or the other category, in a *traditional logic*, there are net thresholds beyond which one belongs to a category rather than another. For example, because they are imprecisely defined, phrases like "young" and "tall" belong to a fuzzy logic. In fact, it is impossible to determine whether or not a person fits into a set in order to categorize them as tall or youthful. Adding a level of relevance to the subset, the categorization preserves all of the data. As a result, this logic permits a more gradual

transition between categories based on the existence or absence of some characteristics.

This difference has led to attach a certain importance to the fuzzy logic, up to its applications in software (knowledge-based systems) with the aim of replacing as much as possible the human decision-makers, and in the hardware (systems control) to replace the human operator. Table 6.1 shows some pros and cons of the fuzzy logic.

Table 6.1.: Pros and cons of fuzzy logic.

Pros	**Cons**
Conceptually easy to understand	Manual tuning of membership functions and other parameters which might be tedious and time-consuming
Flexibility	Not well adaptable to big and complicated problems
Tolerance inaccurate data	Crisp models can be more efficient and even convenient
Modelling of non-linear functions complexity of arbitrary complexity system designing	No systematic approach to fuzzy

6.1.1. Fuzzy sets

A fuzzy set A is characterized by a *membership function*, denoted as $\mu_A(x)$, which maps the elements x of a universe X into a continuous real interval $[0, 1]$. More in detail, the value 0 (zero) indicates that the element x is not included in the fuzzy set at all, the value 1 (one) indicates that the element x is certainly included in the set (these two values correspond to classical set theory), while values between zero and one indicate the degree of membership of the element to the fuzzy set in question. The general notation is:

$$A = \{(x, \mu_A(x))\} \qquad x \in X, \tag{6.1}$$

where each pair $(x, \mu_A(x))$ is called *singleton*: in a crisp set singletons are only x, but in fuzzy sets a singleton is composed by the couple x and $\mu_A(x)$.

6.1.2. Types of fuzzy sets

Fuzzy sets can be classified into many types according to [113]. Some of them are listed below:

- *Finite set*: it contains a definite number of elements;
- *Infinite set*: it contains infinite number of elements;
- *Subset*: a fuzzy set X is a subset of the fuzzy set Y $(X \subseteq Y)$ if every element of X is an element of set Y;
- *Proper subset*: a set X is a proper subset of set Y $(X \subset Y)$ if every element of X is an element of set Y and $|X| < |Y|$.
- *Universal set*: it is a collection of all elements in a particular context or application. All the fuzzy sets in that context or application are essentially subsets of this universal set;
- *Empty set* or *Null set*: it contains no elements:
- *Singleton set* or *Unit set*: it contains only one element;
- *Equal set*: two fuzzy sets which contain the same elements;
- *Equivalent set*: two fuzzy sets which have the same cardinality.

6.1.3. Operations on a fuzzy set

Operations such as *intersection* and *union* are defined through the *min* $(\wedge)$ and *max* $(\vee)$ operators respectively, which are analogous to *product* and *sum* in algebra. Formally the *min* and *max* of an element are denoted by Equations 6.2 and 6.3. Figure 6.1 helps to understand how *intersection* and *union* work. Some fuzzy operations are reported in Table 6.2.

$$\mu_{\widetilde{A}} \wedge \mu_{\widetilde{B}} = min(\mu_{\widetilde{A}}, \mu_{\widetilde{B}}) \equiv \begin{cases} \mu_{\widetilde{A}} & \text{if and only if} \quad \mu_{\widetilde{A}} \leq \mu_{\widetilde{B}} \\ \mu_{\widetilde{B}} & \text{if and only if} \quad \mu_{\widetilde{A}} > \mu_{\widetilde{B}} \end{cases} \tag{6.2}$$

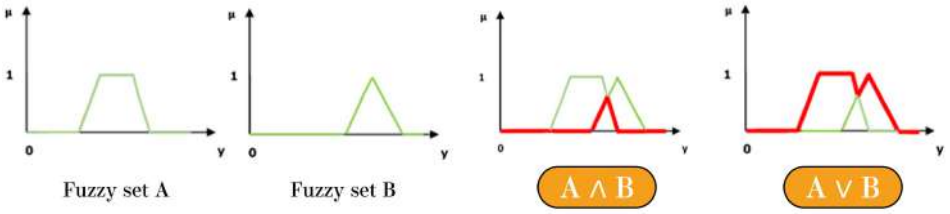


Figure 6.1.: Intersection and union of two fuzzy sets, A and B.

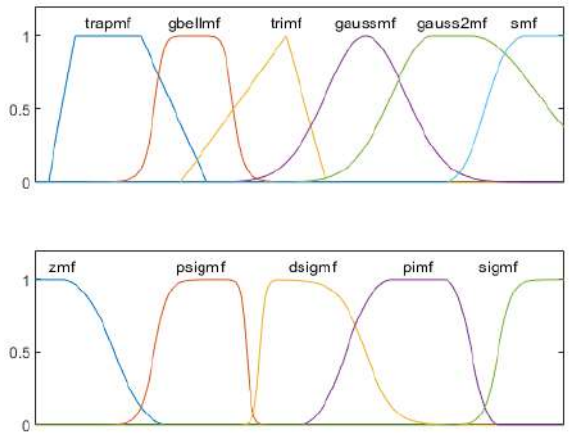


Figure 6.2.: Examples of membership functions.

$$\mu_{\widetilde{A}} \vee \mu_{\widetilde{B}} = max(\mu_{\widetilde{A}}, \mu_{\widetilde{B}}) \equiv \begin{cases} \mu_{\widetilde{A}} & \text{if and only if} \quad \mu_{\widetilde{A}} \geq \mu_{\widetilde{B}} \\ \mu_{\widetilde{B}} & \text{if and only if} \quad \mu_{\widetilde{A}} < \mu_{\widetilde{B}} \end{cases} \tag{6.3}$$

6.1.4. Membership function

Membership function represents the degree of truth in the fuzzy logic. They characterize the *fuzziness*, i.e. all the information in fuzzy set, whether the elements in fuzzy sets are discrete or continuous, and they are represented by graphical forms, as shown in Figure 6.2.

All membership function types are described in Table 6.3.

More in general, membership functions have different features, graphically shown in Figure 6.3. Below, the description of them relative to the fuzzy set $\widetilde{A}$.

- *Core*: the region of universe that is characterize by full membership in

Table 6.2.: The most important fuzzy operations.

Fuzzy set	**Description**	**Conditions**
Union of two fuzzy sets	Given A and B over the same universe of discourse X, their union is a fuzzy set $A \cup B$ in X with a membership function which is the maximum of the grades of membership of every x and A and B (OR operation).	$\mu_{A \cup B}(x) \equiv \mu_A(x) \vee \mu_B(x)$
Intersection of fuzzy sets	It is the minimum of the grades of every x in X to the sets A and B (AND operation).	$\mu_{A \cap B}(x) \equiv \mu_A(x) \wedge \mu_B(x)$
Complement of a fuzzy set	The complement of a fuzzy set A is denoted as $\bar{A}$.	$\mu_{\bar{A}}(x_a) \equiv 1 - \mu_A(x_a)$
Product of two fuzzy sets	$A \cdot B$ has a membership function that equals the algebric product of A and B.	$\mu_{A \cdot B}(x_a) \equiv \mu_A(x) \cdot \mu_B(x)$
Power of a fuzzy set	The β power of A (A^β) has the equivalence to linguistically modify the set with *VERY*.	$\mu_{A^\beta}(x) \equiv [\mu_A(x)]^\beta$
Concentration	Squaring the set is called *concentration* or *CON*.	$\mu_{CON(\beta)}(x) \equiv (\mu_A(x))^2$
Dilation	Taking the square root is called *dilation* or *DIL*.	$\mu_{DIL(A)}(x) \equiv \sqrt{\mu_A(x)}$

Table 6.3.: Fuzzy membership functions.

Membership Function Type	**Description**
gbellmf	Generalized bell-shaped membership function
gaussmf	Gaussian membership function
gauss2mf	Gaussian combination membership function
trimf	Triangular membership function
trapmf	Trapezoidal membership function
sigmf	Sigmoidal membership function
dsigmf	Difference between two sigmoidal membership functions
psigmf	Product of two sigmoidal membership functions
zmf	Z-shaped membership function
pimf	Pi-shaped membership function
smf	S-shaped membership function
constant	Constant membership function for Sugeno output membership functions
linear	Linear membership function for Sugeno output membership functions
String or character vector	Name of a custom membership function in the current working folder or on the MATLAB path. Custom output membership functions are not supported for Sugeno systems
Function handle	Handle to a custom membership function in the current working folder or on the MATLAB path. Custom output membership functions are not supported for Sugeno systems.

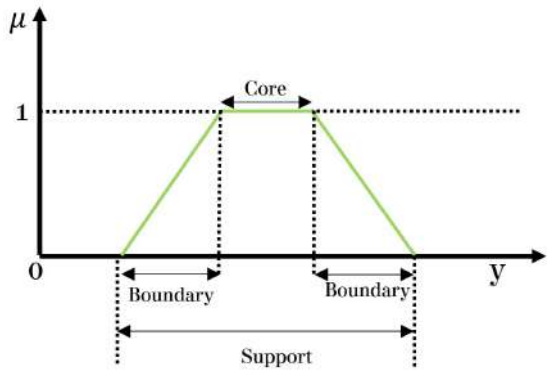


Figure 6.3.: Features of membership function.

the set, i-e. the elements y of the universe of information that satisfy:

$$\mu_{\widetilde{A}}(y) = 1;$$

- *Support*: the region of universe that is characterize by a nonzero membership in the set, i.e. the elements y of the universe of information that satisfy:

$$\mu_{\widetilde{A}}(y) > 0;$$

- *Boundary*: the region of universe that is characterized by a nonzero but incomplete membership in the set, i.e. the elements y of the universe of information that satisfy:

$$1 > \mu_{\widetilde{A}}(y) > 0;$$

6.1.5. Fuzzy Inference System

Any fuzzy logic system must have a Fuzzy Inference System (FIS). To determine the output that corresponds to crisp inputs, it applies fuzzy set theory, IF-THEN rules and fuzzy reasoning process. In IF-THEN rules, predicates are joined by logical connectives such as *and* or *or*. The main characteristics of FIS are:

- Read crisp value from the process;
- Maps the crisp value into fuzzy value using the fuzzy membership function;

- Apply IF-THEN rules from the fuzzy rule base and compute fuzzy output;
- Convert fuzzy output into crisp by applying some defuzzification methods.

Crisp input of any process (measuring temperature of air conditioner, measuring altitude, attitude, height, angle of direction for aeroplane etc.) is given to the fuzzifier, which applies fuzzy membership function and maps the actual readings into fuzzy value (i.e. the value between 0 to 1).

The inference engine applies fuzzy rules from the knowledge base and produces the fuzzy output, which is again between 0 and 1. This output can not be used directly in any process or system. It needs to be mapped into the original domain. Defuzzifier is the inverse process of fuzzification, which converts the fuzzy output into crisp output, which can be fed to the process. Crisp sets are internally converted to fuzzy sets.

There are may defuzzification methods [114] (see Figure 6.4) like:

- *Centroid*: it is the most used function and it is also known as *center of gravity* or *center of area*, it returns the center of area under the curve. This is the most commonly used technique; the only disadvantage of this method is that it is computationally difficult for complex membership functions;
- *Bisector*: it is the vertical line that will divide the region into two sub-regions of equal area. It is sometimes, but not always, coincident with the centroid line;
- *Middle*, *Smallest* and *Largest of Maximum*: MOM, SOM, and LOM stand for Middle, Smallest, and Largest of Maximum, respectively. These three methods key off the maximum value assumed by the aggregate membership function. If the aggregate membership function has a unique maximum, then MOM, SOM, and LOM all take on the same value.

6.1.6. Approaches for FIS

The following are the popular approaches to fuzzy inference systems. The antecedent part of all rules remains the same, they differ only in the consequent part.

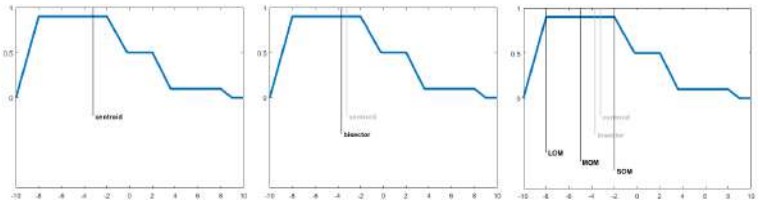


Figure 6.4.: Defuzzyfication methods.

- Mamdani FIS;
- Takagi-Sugeno FIS.

Mamdani FIS

Mamdani [115] FIS is the most commonly used in applications, due to its simple structure of *min-max* operations. Given the inputs (crisp values), their membership values are obtained. This process is called *input fuzzification*. If the antecedent of the rule has more than one part, a fuzzy operator (t-norm or t-conorm) is applied to obtain a single membership value.

Given the cosequent of each rule (a fuzzy set) and the antecedent value obtained in the previously step, a fuzzy implication operator is applied to obtain a new fuzzy set. Two of the most commonly used implication methods are the *minimum*, which truncates the consequent's membership function, and the *product*, which scales it.

Next, the outputs obtained for each rule are combined into a single fuzzy set, using a fuzzy aggregation operator. Some of the most commonly used aggregation operators are the *maximum*, the *sum* and the *probabilistic sum*.

Finally, the obtained fuzzy set is transform in a single numerical value (*defuzzification*).

Takagi-Sugeno FIS

In Takagi-Sugeno FIS [116], a typical rule has the form:

$$\text{IF } x \text{ is } A \text{ and } y \text{ is } B \text{ THEN } z = f(x, y)$$

Here, A and B are fuzzy sets in antecedents and $Z = f(x, y)$ is a crisp function in the consequent.

About the fuzzy inference process, it works in the following way:

- Step 1: fuzzifying the inputs. Here, the inputs of the system are made fuzzy;
- Step 2: applying the fuzzy operator. In this step, the fuzzy operators must be applied to get the output.

Comparison between the two methods

The main differences between Mamdani and Sugeno are shown below:

- Output Membership function: the main difference between them is on the basis of output membership function. The Sugeno output membership functions are either linear or constant;
- Aggregation and Defuzzification procedure: the difference between them also lies in the consequence of fuzzy rules and due to the same their aggregation and defuzzification procedure also differs;
- Mathematical rules: more mathematical rules exist for the Sugeno rule than the Mamdani rule;
- Adjustable parameters: the Sugeno controller has more adjustable parameters than the Mamdani controller.

6.2. Fuzzy control logic for innovative SSM robot speed modulation

The velocities of Equations 5.6 and 5.7 are used to determine the minimum protective separation distance S_{ISO} (Equation 3.3), which does not take into account the relative travel direction of the robot and the operator. This indicates that the value of S_{ISO} unnecessarily grows (proportionally to the estimated speed) even if the robot and the operator are moving away from each other.

The protective separation distance has been redefined as follows in order to improve the production time considering also this situation:

$$S = \alpha[(v_H T_R + v_H T_S) + (v_R T_R)] + (B) + (C + Z_S + Z_R) \tag{6.4}$$

where α is a scaling factor in the interval $[0, 1]$ that should be close to 1 when the current situation is actually dangerous for the worker, while it should be 0 when both the robot and the human worker do not move. The basic idea is to compute α by means of considerations about hazards for the workers. To chose the value of α, a fuzzy inference approach has been developed.

A simple fuzzy inference system has been develop for this purpose. The fuzzy system is based on Mamdami FIS in order to compute the scalar risk assessment parameter, α, to be used in Equation 6.4. The logic has been developed as a three-input, one-output, five-rule problem. Note that the FIS has been developed by using the *Fuzzy Logic Designer* available in MATLAB, which generates a `.fis` file.

6.2.1. FIS inputs for risk assessment

The three inputs identified for the proposed FIS are:

1. the time derivative of the human-robot distance, $\dot{d} = \dfrac{\mathrm{d}\,\|\hat{p}_H - p_R\|}{\mathrm{d}t}$, which is useful to distinguish cases when the operator and the robot are getting closer ($\dot{d} < 0$) and cases when they are moving away from each other ($\dot{d} > 0$);

2. the scalar product between the robot and the human velocity vectors, $\dot{p}_R^T \hat{\dot{p}}_H$, which specifies the relative direction of travel of the operator and the robot which can be aligned ($\dot{p}_R^T \hat{\dot{p}}_H > 0$) or opposite ($\dot{p}_R^T \hat{\dot{p}}_H < 0$);

3. the temperature value of the human point at minimum distance from the robot, which relaxes the speed monitoring if the point closest to the robot belongs to not-human surface, providing reduction of the production time.

Note that the scalar product (second input) represents a complementary information to the time derivative (first input) to distinguish situations shown in Figure 6.5. When the two entities are getting closer to each other and travel along opposite directions, a high hazard risk is occurring. On the contrary, when they are moving away from each other but they travel along the same direction, the risk level is intermediate. Only when they are moving away from each other, travelling along opposite directions, the risk level is low. This is the reason why it is necessary to combine the scalar product information with the time derivative of the distance between the

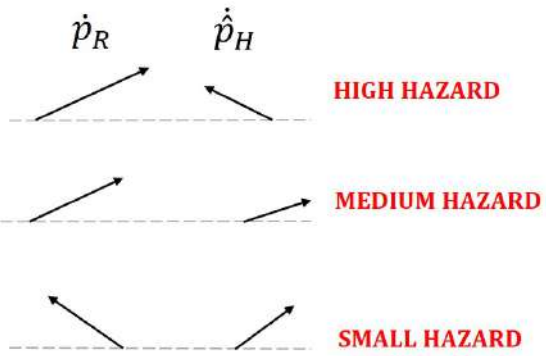


Figure 6.5.: Critical scenarios can be distinguished by considering both $\dot{d}$ and $\dot{p}_R^T \dot{\hat{p}}_H$.

human operator and the robot.

6.2.2. FIS rules for decision-making

The Mamdani FIS has been configured by defining five rules, reported in the Table below.

Table 6.4.: Fuzzy rules: [S] Small, [M] Medium, [H] High, [N] Negative, [P] Positive, [Yes] Human temperature, [No] Not-human temperature, [∼] any.

antecedent			consequent
$\dot{d}$	$\dot{\boldsymbol{p}}_R^T \dot{\hat{\boldsymbol{p}}}_H$	human	α
N	∼	∼	H
P	N	∼	S
P	P	∼	M
∼	∼	Yes	H
∼	∼	No	S

Two membership functions have been selected to represent positive (P) and negative (N) values, a Z-shape and a S-shape, respectively, as well as, to represent human (Yes) and not-human (No) temperature ranges. The defuzzification is performed according to the centroid method.

Note that the output value, α, is generated by analyzing different possible risk situations, corresponding to the three fuzzy sets high (H), medium (M) and small (S), with the twofold aim of avoiding any collisions between

human body surface and robot, and being in line with the current ISO/TS 15066.

6.2.3. Robot trajectory scaling

A standard SSM method usually sacrifices the production time because the minimum protective distance S_{ISO} is constant. On the contrary, the proposed strategy develops an algorithm to adapt the minimum protective distance S to the actual operating conditions and, in turn, changing the robot speed.

A typical industrial pre-programmed task, $\mathcal{T}$, is composed by N positions, $\tilde{q}_i$, associated to velocities $\dot{\tilde{q}}_i$, accelerations $\ddot{\tilde{q}}_i$ and time instants $\tilde{t}_i$ with $i = 1, \ldots, N$. Typically, the pre-programmed joint positions have to be interpolated according to the sampling time T_c required by the robot control interface. Nevertheless, the strategy described below simply translates into the computation of a scaling factor for industrial robots that allows the user to change the speed override in real time. In this work a quintic interpolation is used, i.e., the planned interpolated trajectory is

$$\tilde{q}_h = p_5(t_h; \mathcal{T}) \tag{6.5}$$
$$\dot{\tilde{q}}_h = p_4(t_h; \mathcal{T}) \tag{6.6}$$
$$t_{h+1} = t_h + T_c, \tag{6.7}$$

where t_h is the h-th discrete time instant, p_4 is the derivative of the polynomial p_5, $\tilde{q}_h$ and $\dot{\tilde{q}}_h$ are the *planned* joint position and velocity at time t_h, respectively.

The robot speed is modulated by scaling the trajectory time with a *safety scale factor* k, which can assume values in the interval $[0, 1]$. The mathematical expression of the curve relating d with k is a piecewise-defined function whose graph is shown in Figure 6.6. More in detail, when $d < S$, k is 0 and the robot stops as required by the SSM constraint. Otherwise, when $d > \nu S$, where $\nu > 1$ is a design parameter needed to guarantee continuity of k, the robot can move at full speed to improve the production time, i.e., νS represents the warning distance. Finally, when d assumes a value between S and νS a polynomial of order at least 3 should be used to avoid acceleration discontinuities. Obviously, ν is another design parameter that change the size of the scaled speed mode zone.

Practically, the trajectory is scaled computing Equation 6.5 and using a

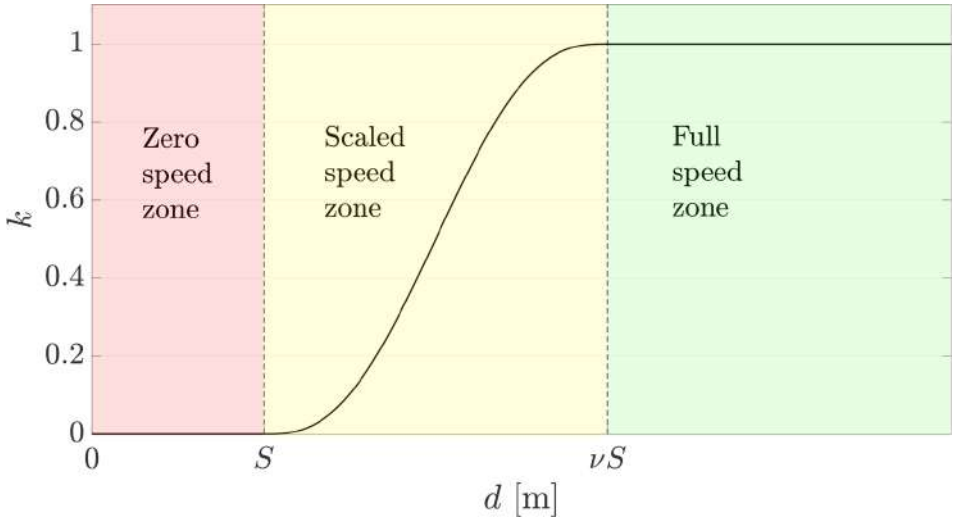


Figure 6.6.: Relation between d and k; S is the minimum separation distance computed in real time as in (6.4).

scaled time τ_h, i.e.,

$$q_h = p_5(\tau_h; \mathcal{T}) \quad \tau_{h+1} = \tau_h + kT_c, \tag{6.8}$$

where q_h is the actual joint command at time t_h. Obviously, the joint command q_h, as well as the scaled time τ_h, are generated with sampling time T_c.

This approach effectively scales the joints velocities. In fact, using (6.8), it is

$$\dot{\tau} \approx \frac{\tau_{h+1} - \tau_h}{T_s} = k. \tag{6.9}$$

By time differentiating (6.8), (6.10) demonstrates that the velocity is scaled by the *safety factor* k,

$$\dot{q}_h = p_4(\tau_h; \mathcal{T})k. \tag{6.10}$$

This approach guarantees that the task $\mathcal{T}$ remains the same in position, but, simultaneously, the resulting velocity is scaled according to k. Note that, in case the industrial control interface of the robot allows the user to change online the speed override of any motion instruction, it is sufficient to set a speed override equal to the factor k.

7. Experimental validation

To assess the practical relevance of the proposed method, this Chapter shows some experimental results. The experiments are executed on two different setups. The first lab-scale setup has been hosted by the RoboticsLab [117] of University of Campania "Luigi Vanvitelli" which hosts numerous industrial and collaborative robots as well as adavanced robotic equipment funded by National and European research projects as well as by Regione Campania. The second full-scale setup has been hosted by the Pomigliano D'Arco plant of Leonardo S.p.A. [118], an Italian multinational company specialising in aerospace, defence and security. This setup includes a robotic work cell designed under the Clean Sky Horizon 2020 Programme and named LABOR, according to the specifications of the European research project GA n. 785419. The work cell performs the cooperative assembly of aeronautical panels composed of hybrid (composite-metal) structural parts.

7.1. Lab-scale experiment

7.1.1. Set-up description

This section shows a complete example of a collaborative task conducted @RoboticsLab [117] of University of Campania "Luigi Vanvitelli". The experiment highlights all the features of the proposed methods. The workstation used to execute all the algorithms, except the DT-CNN, is a standard desktop PC running Ubuntu 16.04 LTS with an Intel Core i7-8700K CPU at 3.70GHz, 12GB of RAM and an nVidia GeForce GT 730 GPU with 12GB of memory. The Yolo prediction process runs on another PC running Ubuntu 18.04.2 LTS with an Intel Core i9-7980XE CPU at 2.60GHz, 18 GB of RAM and an nVidia Titan V GPU with 12GB of memory. The application is based on a ROS industrial architecture and the external PC communicates, through an Ethernet cable at 50Hz, the reference joint positions to the robot controller, equipped with the MotoPlus SDK [119].

The collaborative scenario is shown in Figure 4.15. A Yaskawa MOTOMAN

SIA5F industrial robot has been used to perform the simulated assembly-like task. The collaborative workspace is about 4x2 m. The innovative multimodal vision system is the one described in Section 4.6; it combines the Microsoft Kinect v1 depth camera (Table 4.1) and the Optris PI450 thermal camera (Table 4.2).

The demonstration simulates the robot executing a sealing operation, while a human operator executes a manual piece change and a visual inspection.

7.1.2. Software architecture

Figure 7.1 explains how the proposed algorithms have been developed and integrated to each other from a software point of view. The boxes contained into the orange highlighted square represent the contribution to the robot control strategy. On the other hand, the external HDT box represents the vision system architecture which observes the collaborative workspace by merging the depth image and the thermal image in input (C1 and C2) and provides the proper inputs to the control architecture. The novel strategy focused on the minimum protective distance S computation, which is processed by the trajectory scaling algorithm described in Section 6.2 to optimize the collaborative scenario. This algorithm compares the separation distance d and the minimum protective distance S, scaling the pre-planned path, when indispensable. The computed joint configuration is finally sent to the robot controller.

7.1.3. Preliminary configurations

The covariance matrix Q_2 which quantifies the uncertainty on the velocity dynamics of Equation 5.4 has been chosen as

$$Q_2 = \text{diag}(0.02,\ 0.05,\ 0.05)\ \ \text{m}^2/\text{s}^2, \tag{7.1}$$

while the noise covariance has been estimated as

$$N = \text{diag}(0.0009,\ 0.0008,\ 0.001)\ \ \text{m}^2. \tag{7.2}$$

According to Equation 6.4, to compute the time required by the robot system to respond to the operator's presence, T_R, the rate of the ROS topic on which the value of the variable d is published has been observed. The rate of this topic is about 10 Hz.

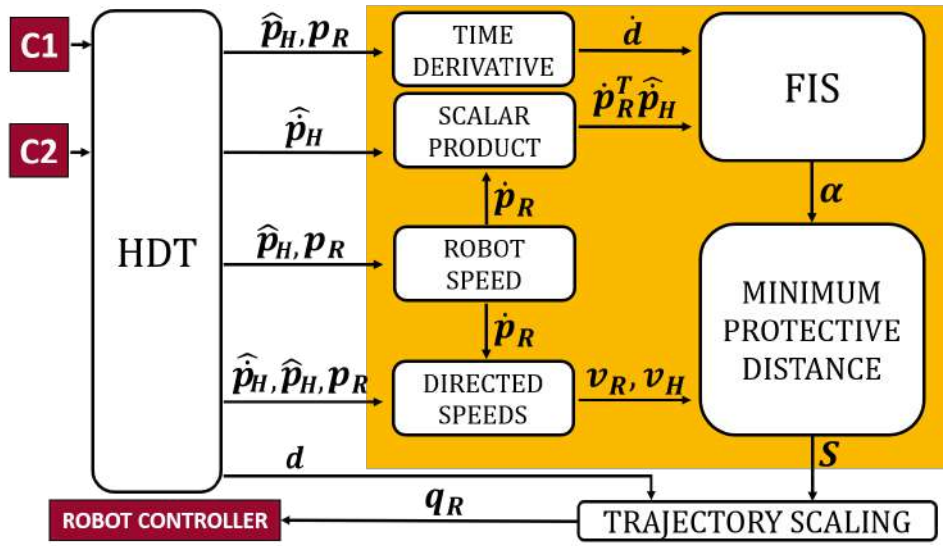


Figure 7.1.: ROS developed architecture.

On the other hand, the response time of the machine which brings the robot to a safe, controlled stop, T_S, has been computed by reproducing a worst case scenario. The robot, completely extended in a vertical position, has been suddenly stopped during a pre-planned trajectory executed at the robot maximum speed. The trajectory consists in rotating the second robot joint, L, of a 90° angle, reaching the worst robot configuration (see Figure 7.2).

To compute the Euclidean distance travelled by the robot while braking, B, the same test has been analyzed.

The absolute accuracy of the robot, Z_R, has typical values between 0.2 and 1 mm.

The sensor uncertainty Z_S is related to the depth camera uncertainty and depends on the extrinsic calibration accuracy. A 2D point of pixel coordinates has been taken from the depth image viewer, obtained the z Cartesian coordinate and 2D point of Cartesian coordinates has been calculated from 2D point of pixel coordinates, from z and from the intrinsic parameters of the used depth camera. Finally, the norm of the point in Cartesian coordinates has been computed to establish the value of Z_S.

The intrusion distance C, based on the operator reach, has been chosen to better appreciate the zero speed zone.

Figure 7.2.: Yaskawa Motoman robot worst case configurations.

7.1.4. Task description

The collaborative task scenario simulated at @RoboticsLab laboratory simulates a robot executing a pre-planned path at a given nominal speed, and a human operator who enters the collaborative workspace to perform some inspection checks. During his work, the operator handles different objects, e.g., a bubble level and a notepad. He is close to the robot and he works at different distances, approaching the robot randomly.

Figure 7.3 shows the main snapshots of the task execution. At $t = 12\,\mathrm{s}$ the operator puts a bubble level on the desk, then he closes the toolbox at $t = 18\,\mathrm{s}$ and he takes a notepad to write some inspection measurements at $t = 26\,\mathrm{s}$. The experiment demonstrates that, by reading the temperature of the point belonging to the human cluster which is the closest to the robot, the fuzzy inference system (Figure 7.4) marks it as belonging to an object manipulated by the operator or actually belonging to the human body surface. Indeed, during these phases, the point closest to the robot belongs to the handled tools, thus the robot may proceed at a slightly higher speed. Only at the end of the experiment, the worker puts the tools in place ($t = 41\,\mathrm{s}$) and leaves the collaborative zone ($t = 45\,\mathrm{s}$), thus his body surface is directly the most exposed to the robot motions, i.e., the fuzzy system reduces the robot speed to ensure the worker safety.

7.1.5. Task results

Results are shown in Figure 7.4.

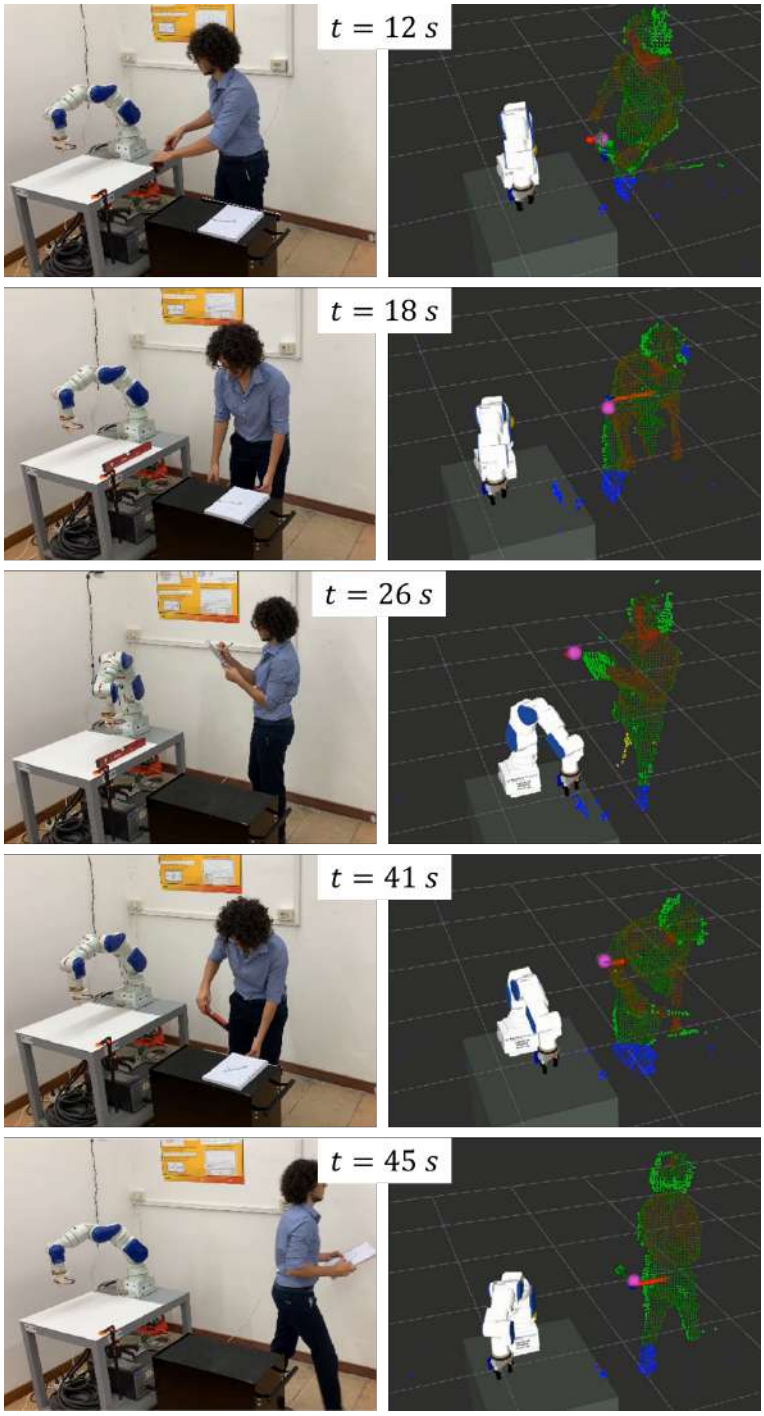


Figure 7.3.: Snapshots of the task execution of the lab-scale experiment.

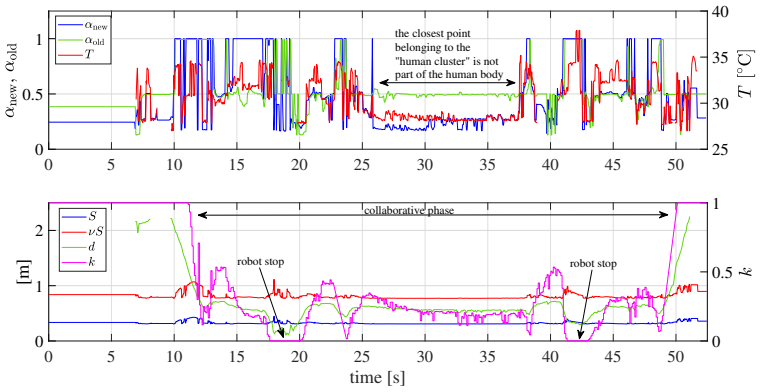


Figure 7.4.: The top plot shows a comparison between α_{old} values (green line) obtained by following [120] and α_{new} values obtained with the proposed fuzzy system (blue line). The bottom plot shows the estimated separation distance d, the protective distances proposed by the research work (S and νS) and the scaling factor k (magenta line – right axis).

The bottom graph represents the robot speed scaling performance. At about 10 s the operator enters the collaborative workspace and the system starts to measure the separation distance d (green line). When d goes below the warning distance νS, with $\nu = 2$, (red line) at about 11 s, the trajectory scaling factor k (magenta line) becomes lower than 1 and the robot reduces its velocity without changing its path. During some intervals of the collaborative phase, d goes below the minimum protective distance S, thus k becomes 0 and the robot stops.

The top graph compares the α_{old} values (green line) obtained through the preliminary fuzzy inference system proposed in [120], and the α values computed with the new algorithm (indicated with α_{new} in the figure (blue line)). The difference is due to the use of the new input, i.e., the temperature value, T (red line), of the point at minimum distance from the robot, which belongs to the human cluster. Until 12 s the operator is inside the collaborative workspace and manipulates the bubble level to perform his task. After that, at about 18 s, he closes the toolbox, thus his body part becomes close to the robot: the elbow, the right arm, the hand, the shoulder are most exposed to risk in the direction of the robot. Therefore, the mean value of the temperature of these points is about 35° C, that the fuzzy inference system associates to the human surface. In this first phase, the α_{new} signal follows the α_{old} one without major changes. The most interesting phase

is the second one. At about 26 s, the operator takes notes on a notepad to perform some inspection checks. From this moment on, the notepad is the object closest to the robot. Note that it clearly belongs to the human cluster, even if it is not part of the operator body surface. The temperature value can be read to realize this difference. Indeed, during this phase the point at minimum distance has got a mean value of about 28° C, which does not represent a human surface temperature value. The proposed fuzzy inference approach adopts this input also for the α_{new} computation, which is lower than the α_{old} estimation because the risk is considered lower. In other words, during similar situations, the minimum protective distance S can be reduced to allow the robot to move faster. The last phase, at about 38 s is similar to the first one, when the closest point of the human cluster belongs to the human surface.

7.1.6. Comparison with the standard SSM

This section describes a test to prove the efficiency of the fuzzy inference approach with respect to the traditional SSM approach into industrial applications. In other words, the test compares the performance obtained by applying S_{ISO} during a collaborative task with the performance obtained by using the novel computation S. The results have been published in [120].

During the test, the robot is executing a pre-planned path at a given nominal speed, while, suddenly, a human operator enters the collaborative workspace to perform some manual operation close to the robot, at different distances.

Figure 7.5 shows the results of the experiment. The graph at the top of the figure shows the distance between the human operator and the robot and it can be compared with S_{ISO} and with the interval $[S,\ \nu S]$, where $\nu = 2$. The bottom plot shows the two inputs of the fuzzy inference system ($\dot{d}$ and $\dot{\boldsymbol{p}}_R^T \hat{\dot{\boldsymbol{p}}}_H$) and the trajectory scale factor k. A video of the experiment is available at `https://youtu.be/RzLZ6RQBPCY`.

The robot executes a planned task, suddenly (at about 16 s) an operator enters into the workspace simulating a manual task. This is visible in the top plot of Figure 7.5, where the human-robot distance decreases. Note that for almost the whole task duration the separation distance robot-operator is below the S_{ISO} signal, this would have caused frequent starts and stops of the robot if the traditional approach had been applied. Instead, through the proposed one, the robot reduces its velocity according to the observed separation distance. This is visible in the k signal of the bottom plot that varies according to d. Notice that k goes to 0 only when the distance d

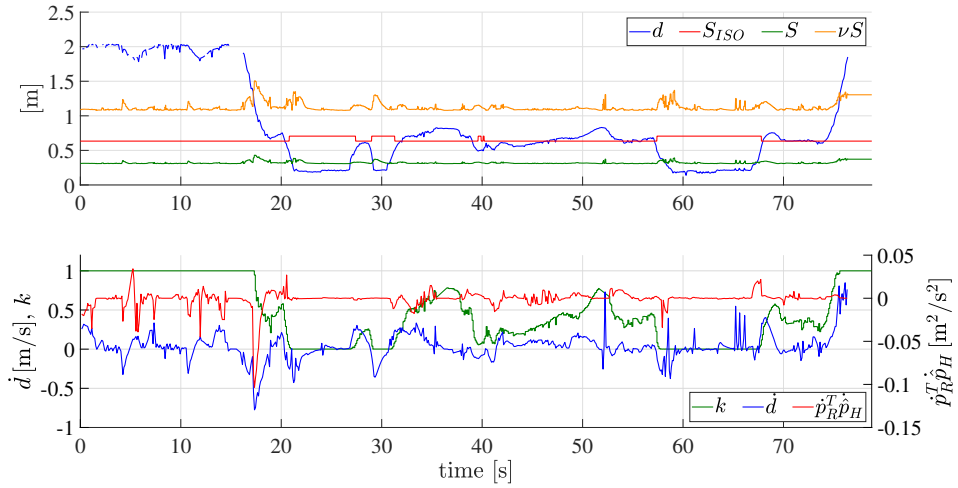


Figure 7.5.: Experiment: an operator enters the shared workspace while the robot is moving. The top plot shows the estimated distance robot-operator (d), the protective distances proposed by the regulation without sensing (S_{ISO}) and the protective distances proposed by the industrial research work (S and νS). The bottom plot shows the trajectory scaling factor k, the time derivative of the distance $\dot{d}$ and the scalar product of velocities.

goes below the protective distance S. Moreover, another property of the proposed solution is that S increases only when the distance decreases (i.e., when $\dot{d} < 0$) and not when the distance increases. This is due to the computation of the directed speed and the fuzzy rules.

The shown experiment and the related video demonstrate how the proposed approach guarantees a safe human-robot coexistence in the collaborative workspace. This is achieved both in accordance with the ISO/TS regulations and minimizing dead times in the production process.

7.2. Full-scale experiment

The aerospace manufacturing is looking at ways to change their production processes in order to improve costs, flexibility and efficiency. The LABOR project, acronym for *Lean robotized AssemBly and cOntrol of composite aeRostructures*, aims at introducing robotic solutions for the assembly line of fuselage panels adopting medium size robots equipped with smart tools, HRC approach and a distributed software architecture. The project started in March 2018 as part of the European Clean Sky 2 research program. Three partners (LOCCIONI, UNICAMPANIA and UNISA) are developing the prototype cell in collaboration with LEONARDO S.P.A. that is the Topic Manager.

7.2.1. The LABOR project

The system is shown in Figure 7.6 and consists of a jig that holds the panel to be assembled, two Fanuc M-20iA-35M 6-axis robots mounted on two linear axes in order to reach the whole panel and real time measurements to perform the quality control of the assembled components. The cell automatically performs the referencing of the robot working area on the basis of the recognition of geometric features of the parts to be coupled (edges, holes, etc.) through the use of 3D smart inspection tools. After the "one shot" drilling and countersinking operations, the hole is processed to guarantee a high standard of the process. Installation and sealing of the fastener complete the working cycle. Furthermore, the advanced multimodal perception system proposed in this research work has been integrated to monitor, in real time, the collaborative workspace of the internal robot to allow safe HRC tasks. In fact, the assembly process is composed of very complex sub-processes that can require human intervention, such as the dismantling and consequent deburring of the metal parts. To avoid interrupting the

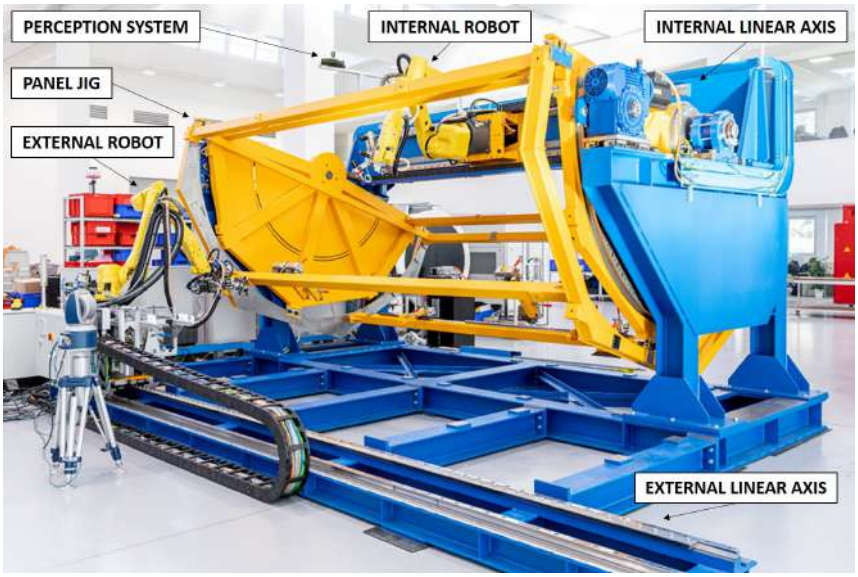


Figure 7.6.: LABOR final work cell design.

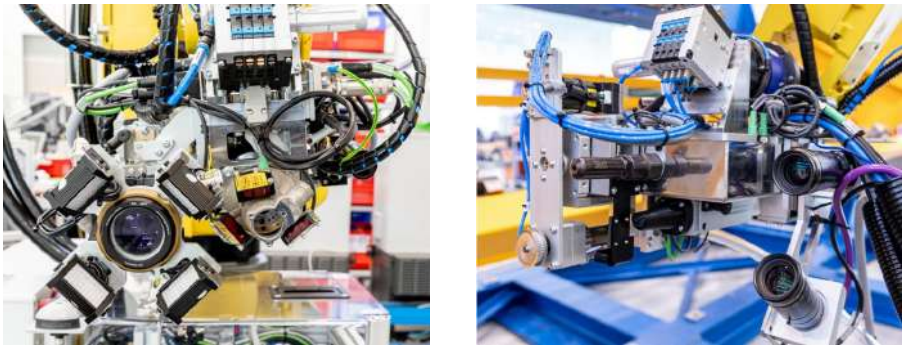
automatic process of the system and therefore stopping the robots, thus lengthening production times, the proposed SSM strategy has been integrated to allow humans to carry out these operations while robots continue their work cycle.

7.2.2. Workcell design

The LABOR system is equipped with a *Fastener warehouse* which handles up to 8 different fasteners (with variable diameters and lengths), a *Drilling tip warehouse* which is a carousel of 15 positions, and a dedicated *HRC workstation*. Safety fences are installed to guarantee the operator accessibility. Four dedicated smart tools have been developed.

The external robot handles two different tools which are shown in Figure 7.7. They can be exchanged through a quick tool change system placed on the robot basement.

- *Drilling and 2D Inspection Tool*: the robot approaches the drill point and performs an orthogonal alignment with the panel skin through the 3 laser sensors installed around the drilling nose. An automatic electrospindle is mounted on a linear axis to control the advancing motion during the drilling task. A vacuum pipe removes powder and chips, while a force sensor controls the robot during the panel stack



(a) (b)

Figure 7.7.: External robot tools: *Drilling and 2d Inspection Tool* (a) and *Fastening and Sealing Tool* (b).

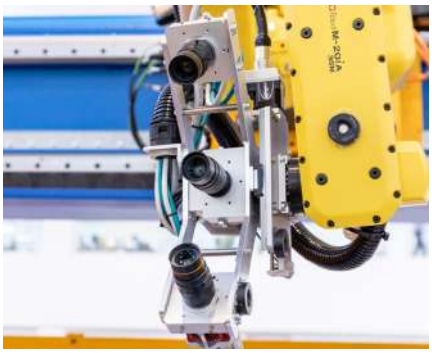
clamping. A lubrication system and two telecentric lens on two 2D cameras have been integrated to perform the hole and countersink inspection, i.e. fastener flushness, locking ring, stem height, hole and countersink diameter measurements and absence of burrs check;

- *Fastening and sealing tool*: the pneumatic gripper has been conceived to hold the fastener end. The fastener is automatically selected from the *Fastener Warehouse* and is brought to the gripper tip through a pneumatic transmission system. The fastener is rotated to execute the sealing application through a pneumatic rotary actuator. The sealant gun has been housed into the tool mechanical structure, and connected to the electrical motor that controls the sealant dispensing. Finally, a structured LED light pattern projector and the 2D camera have been integrated to execute the fastener flushness measurement.

The internal robot handles only one tool that is fixed on the robot flange (see Figure 7.8). The tool is composed of two parts: the counterthrust tool and the 3D internal inspection tool:

- *Counterthrust tool*: the main component is the counterthrust rod that has been connected through to the suction pipe with the cell aspirator;

- *3D inspection tool*: the structured LED light pattern projector and the three cameras composing the tool have been installed on a screw-nut mechanism actuated by the electric motor. The tool executes the referencing of internal robot with respect to internal panel features and the installed fastener measurements, i.e. delamination and fastener sleeve height and diameter measurements.



(a)

(b)

Figure 7.8.: Internal robot tools: *3D Inspection Tool* (a) and *Counterthrust Tool* (b).

The workstation dedicated to the HRC module is composed of:

- a RUBEDOS Viper depth camera;
- an OPTRIS PI450 thermal camera;
- a standard desktop PC running Ubuntu 16.04 LTS with an Intel Core i7-9700K CPU at 3.60GHz and an nVidia GeForce RTX 2080/PCIe/SSE2 GPU installed for the Yolo CNN human prediction.

The HRC-PC has been installed into the LABOR robotic cell, placed near to the cell structure, in front of the linear stage of the external robot, as shown in Figure 7.9 (left). The multimodal perception system composed of both the depth camera and the thermal camera has been installed into the LABOR robotic cell as shown in Figure 7.6. The vision system has been assembled and mounted on a support of one DoF, as shown in Figure 7.9 middle. The set-up has been then mounted on a movable bracket of another one DoF, as shown in Figure 7.9 right. The choice of giving two DoF to the perception structure guarantees more flexibility for camera positioning, so that the collaborative workspace can be observed from different point of views, both for the experimental demonstration and for some future improvements. The whole system has been placed on the linear stage of the internal robot, according to the final simulation results discussed below.

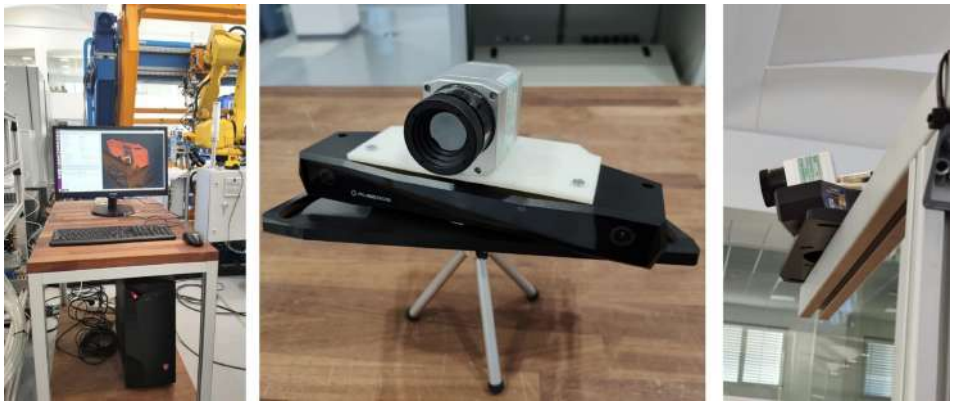


Figure 7.9.: Hardware integration of HRC module into the LABOR workcell: PC for image processing and HRC algorithms (left); multimodal perception system installed on a 2-DoFs movable bracket(middle and right)

7.2.3. Software architecture

From a software point of view, the main concept of the LABOR architecture is the development and the integration of different intelligent modules, as shown in Figure 7.10. Each module is an independent node which manages all the hardware components and it is related to and communicates only with the HMI module, the cell supervisor. Commands and feedbacks are sent from/to the HMI modules through an OPC-UA bus, the communication protocol chosen according to its intrinsic flexibility, adaptability, transparency which have to be fulfilled to satisfy the distributed intelligence approach. The interaction and communication of each module through the network allow to build a more complex system and to achieve the final complete task.

Focusing on the HRC module, it is important to highlight that the Fanuc robot controller allows the user to set on line a register with the desired value of the speed override. This value is set by the external PC running the speed scaling algorithm that computes the speed override k. The external PC communicates with the cell Human Machine Interface (HMI) module via TCP-IP with a frequency at which the scaling algorithm runs, i.e., 30 Hz. Finally, the HMI updates the register of the robot controller in real time via EtherCAT Automation Protocol (EAP).

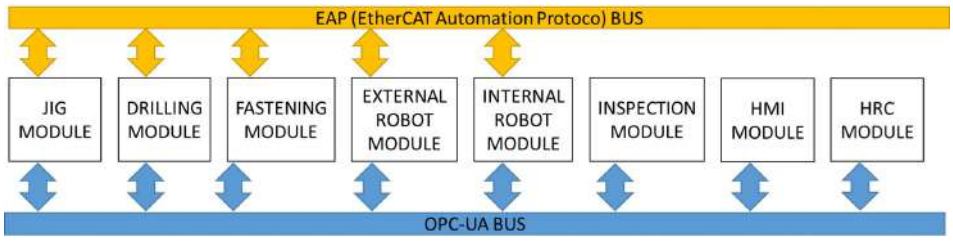


Figure 7.10.: LABOR distributed intelligence software architecture.

7.2.4. Digital twin and simulated environment

In this section, the concepts about the SSM approach to human-robot coexistence are analyzed through the support of a digital model of the LABOR work cell, which includes the robots, the end effectors, the jig and the perception system for workspace monitoring. The simulation environment selected for Labor cell is V-REP [121] that provides a unified framework, combining many powerful internal and external libraries that are useful for robotics simulations. This includes dynamic simulation engines, forward/inverse kinematics tools, collision detection libraries, vision sensor simulations, path planning.

The execution of the robot working cycle and the manual assembly operations inside the collaborative workspace are simulated with the aim to optimize the final DMU. The simulator allows simulating in detail the whole working cycle, by analyzing the robots reachability, the robots path execution, the human detection and tracking and every condition that happens in the work cell. Some tools that help testing the proposed algorithms for the achievement of the final full-scale experiment requirements have been developed, e.g., the robot speed monitoring related to the collaborative SSM tasks.

Perception system configuration and sensors positioning

The interest for the HRI concept is due to the increase of the production efficiency, by introducing perceptive solutions which allow operators to be detected when they enter the cell, safely execute some manual operations and exit the cell, without the complete stop of the machines. In this perspective, the developed simulation environment includes some sensors which could be added inside the simulated scene.

The perception system has been simulated through the integration of different kinds of sensors, i.e., depth cameras, thermal cameras and lidars. If the

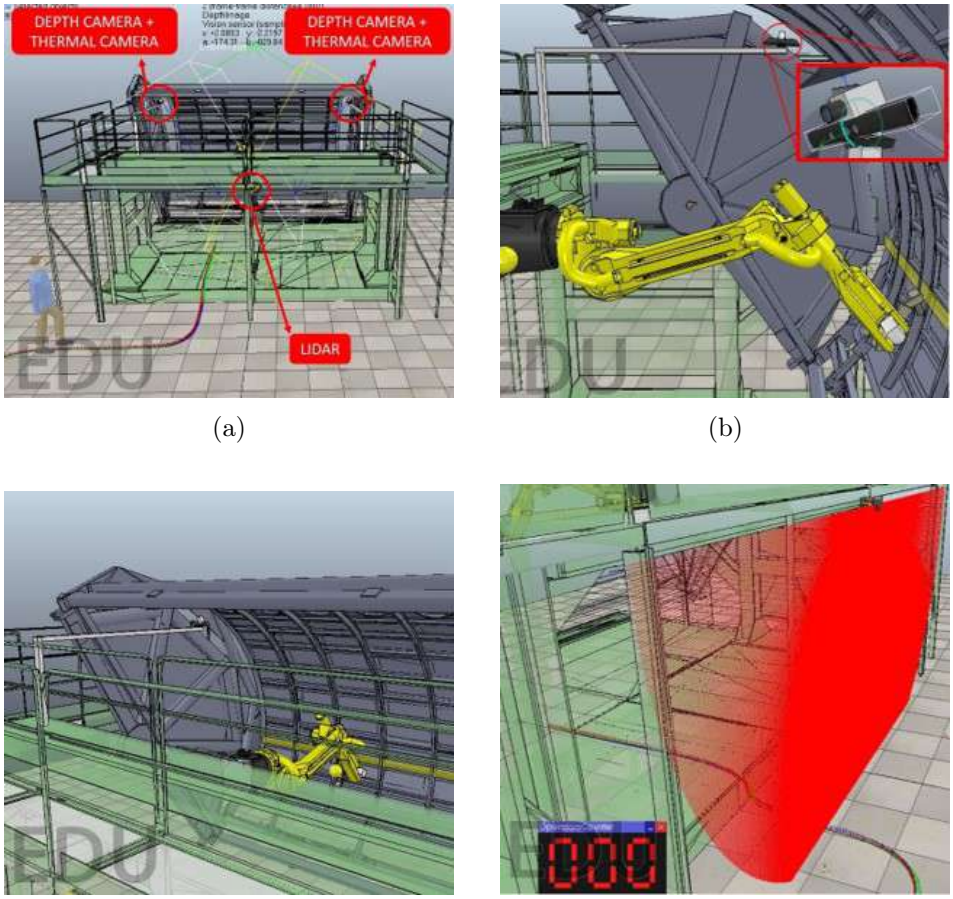
choice of the multimodal perception system represents the innovative and experimental architecture to detect and track humans inside the work cell, proposed by this thesis, the choice of the laser scanner authorizes the use of this perception system because this sensor is standardized by the current ISO regulation. In fact, ISO 10218-2 [8] specifies that non-standardized sensors such as passive infrared protective devices (e.g., infrared thermal imaging systems) are permitted for use as part of a larger safety system, but, without an associated performance standard they are not considered reliable enough as the sole means of protection. Therefore, the final perception system should be composed by two depth cameras, two thermal cameras and a standard lidar (e.g., SICK [122]).

Figure 7.11a shows the simulated LABOR work cell and highlights the final configuration of the perception system and the positioning of its components. The thermal cameras must be mounted above the respective depth sensors because the basic idea of the proposed HDT algorithm is to acquire the two stackable views and combine them to obtain a unique image, with greater information content than the single sensors alone. The simulator enables to simulate the view of the cameras during the execution of the work cycle of the robots. This feature has been used to optimize the components positioning, with the aim of maximizing the volume of sight and minimizing the operator occlusions during the manual assembly operations. At the same time, the final position of the multimodal perception system must ensure that the sensors views are not occluded by other moving elements of the cell when they are active when the operator is inside the LABOR cell, e.g., the rotating panel. In this perspective, in order to observe the whole collaborative workspace, the vision system have been located at a height of about 3.75m from the floor and inclined about 60° downward, as shown in Figure 7.11b. The system have been finally mounted on a suitable beam, which is fixed to the supporting beam of the internal robot linear stage (see Figure 7.11c).

Regarding the lidar sensor, the simulator allows setting its angle view. Assuming to use the laser scanner as a barrier which detects human accesses and counts them, the sensor has been mounted at the top-center of the operator access door and in parallel to the access plane and its angle view has been set to 180°, as shown in the right image of Figure 7.11d.

Manual assembly operations and HDT

In this subsection, an example of manual assembly operation is simulated to integrate the simulated work cell with the collaborative SSM task. An op-



(a) (b)

(c) (d)

Figure 7.11.: LABOR simulated environment: perception system configuration (a); multimodal vision system positioning (b) and its 2-DoFs movable bracket (c); lidar positioning (d).

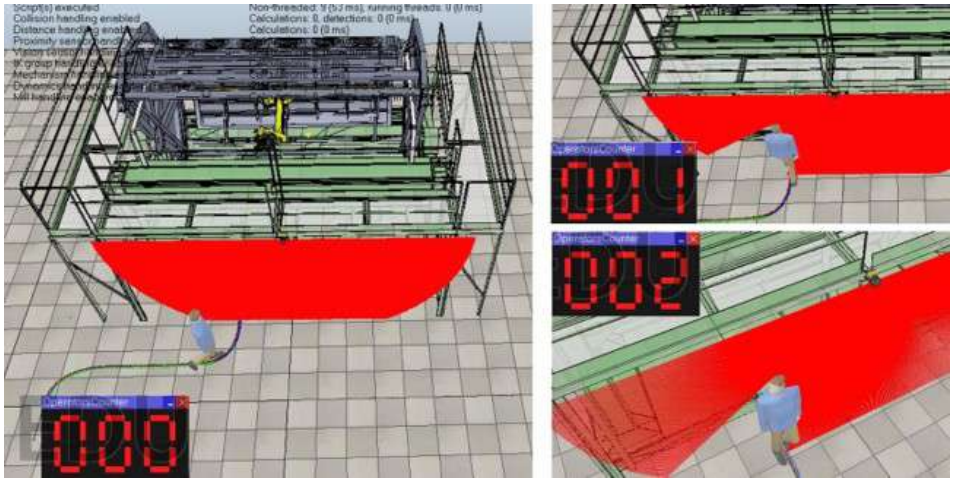


Figure 7.12.: LIDAR sensor integration and operator counter

erator executes a manual operation on the panel and the integrated sensors detect and track him during the task. The simulator makes a digital human available (Bill) to implement simulations with people presence. The mannequin can be animated by manual insertion of joints along the skeleton. A realistic human path simulation has been obtained through a predefined path, composed by some via-points, which lead the operator from the access door till the area where he performs assembly operations. The next main feature of the simulator is to simulate the monitoring system the operator inside the work cell through the perception system proposed by this research work.

The area where operators can enter the cell is monitored through the lidar sensor. A LUA script activates the lidar sensor and produces a signal every time that a human crosses it. Figure 7.12 shows how the Operator Counter can count how many times the operators enter and exit from the LABOR cell or how many operators are currently inside the cell, e.g., by considering the direction of access.

When the operator is inside the work cell, the DTCs perform the HDT. Figure 7.13 shows a temporal sequence in which the operator enters the cell, arrives until the area of interest, makes the manual assembly operations and finally comes back to leave the cell. The simulator allows simulating the view of the cameras according to their specific output information (e.g., depth, temperature) by applying some available filters on the image frames. The top right of each image of Figure 7.13 shows the final camera views. The static background has been subtracted from the cameras views to highlight

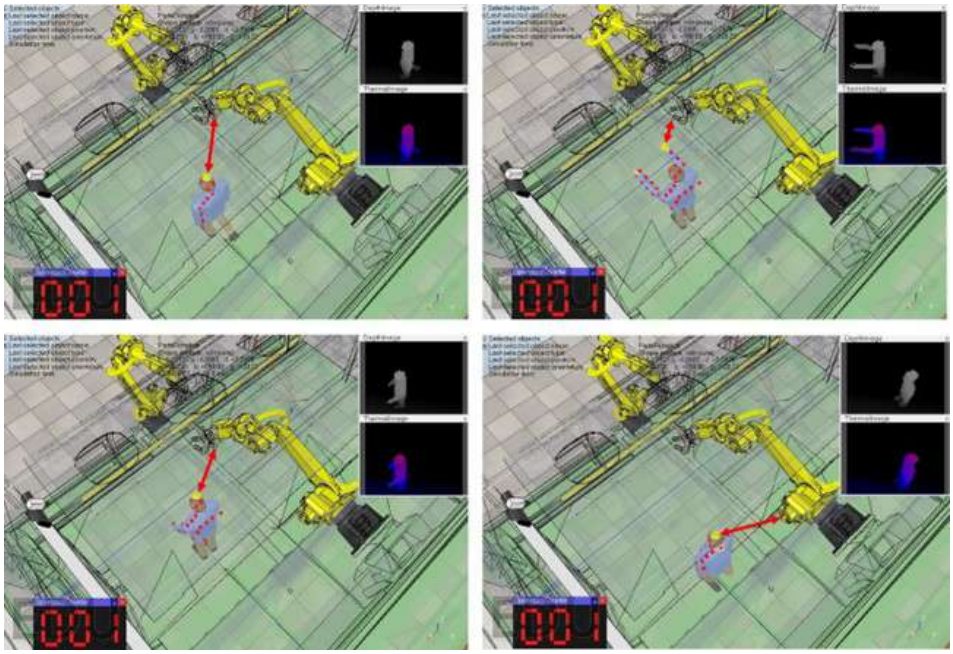


Figure 7.13.: Simulation of HDT during manual assembly operations

only dynamic objects. Moreover, the whole volume of the internal robot has been filtered from the current image frame, as the proposed pipeline. The resulting images have been processed to obtain a point-cloud of red points superimposed on each human surface, which is observed from the proposed multimodal vision system.

Robot path execution and SSM simulation

The final DMU aims at showing a complete test simulation of a typical SSM task in which an operator executes a manual operation on the panel while a predefined robot assembly operation is being performed. In this perspective, this subsection shows the simulation of a drilling task, which is performed inside the collaborative workspace. Five holes are made by the robots on a shear tie of the Side AFT panel, while the operator arrives in front of the stringer joint to make manual operations.

The meshes of every robot link have been imported inside the simulator and simulated angular joints have been identified between the couples of link to build the robot kinematic chain. Then, the internal robot has been equipped with its inspection tool and the external one has been equipped

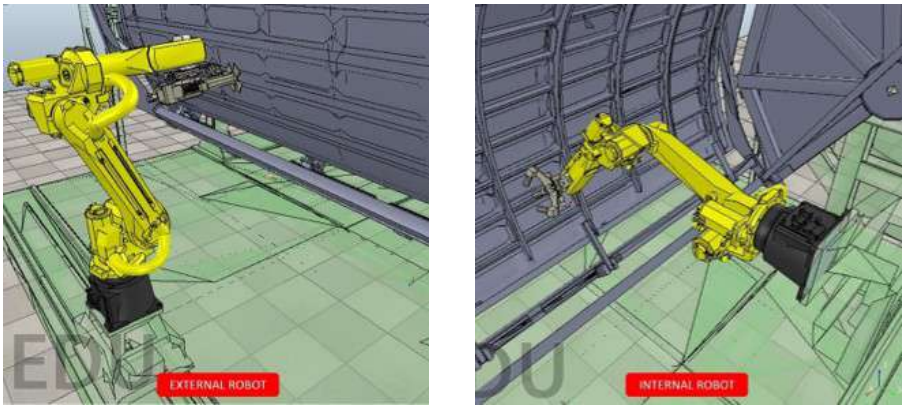
with the drilling tool, as shown in Figure 7.14.

A LUA script has been written to move the robots along a predefined path. The simulator lets the scripts to be interconnected, i.e., to communicate signals from a subsystem to another, e.g., from the perception system to the simulated robot controller. This feature has been used to write two distinct LUA scripts: one for the perception system, which implements the algorithm that computes the minimum distance points between the closest human and the internal robot, and another script for the robot controller, which implements the algorithm that solves the SSM collaborative tasks. In other words, the output of the first script becomes the input of the second one, so that the robot speed is finally modulated.

To analyze the performances and compare the expected results with the simulated one, the computed human point which is closest to the robot surface is identified with a yellow sphere, as shown in Figure 7.13. Note that the point- cloud based approach for HDT allows to more precisely detect and track the points at minimum distance. The density of the point cloud has been chosen to reduce the computational load of the simulation.

Figure 7.15 shows some snapshots of the simulated task: while the human operator executes his manual task, the robots execute a planned trajectory. Note that the drilling trajectory does not follow cycle time considerations but it simply represents an example to visualize the effect of the speed-scaling algorithm. While the external robot follows the instructions of the internal robot and executes the fastener insertion, the internal robot executes the inspection task. Its speed is modulated through the SSM algorithm, which takes into account the computed minimum distance, as shown in the graph.

The simulator allows also displaying of time graphs. For example, Figure 105 reports the robot velocity being modulated by considering the computed minimum distance, which decreases when the operator is closest to the robot. The dotted lines are the two fixed thresholds which are selected to compute the scale-factor: the green line represents the warning distance and the red line represents the security distance. When the operator is closer than the safety distance, the robot stops; when the operator is farther than the warning distance, the robot moves at nominal speed.



(a) (b)

Figure 7.14.: Robots equipped with the end effectors: external robot equipped with the drilling tool (a) and internal robot equipped with the inspection tool (b).

7.2.5. Optimization of the collaborative work cycle

To fully exploit the HRC module advantages, the work cycle of the SIDE FWD panel, which needs the combination of both robot and manual activities, has been redefined with respect to a fully-automated work cycle. Table 7.1 describes the main assembly operations to be executed on a general aircraft panel and compares them with and without HRC application.

The general idea is that the human worker and the internal robot can work simultaneously on different panel areas, as shown in the simulation of Figure 7.16. This approach ensures the human safety by executing the collaborative algorithm described in this thesis work. Simulated analyses estimate that the time to manually assemble the panel is around 15 h. LABOR reduces the working time of about 40% with a standard fully-automated working cycle, while the percentage rises by using the HRC module (48%).

7.2.6. Safety: risk assessment and risk mitigation

The contents of this section refer to part of the document presented to Leonardo S.p.a. to request the execution of the full-scale experiment on the LABOR robotic work cell. The analysis is based on the current regulations about risk reduction and risk mitigation detailed in Appendix B. The document aims at comprehensively evaluating the risks to the health and safety of workers present during the test, as specified in the art. 2 of

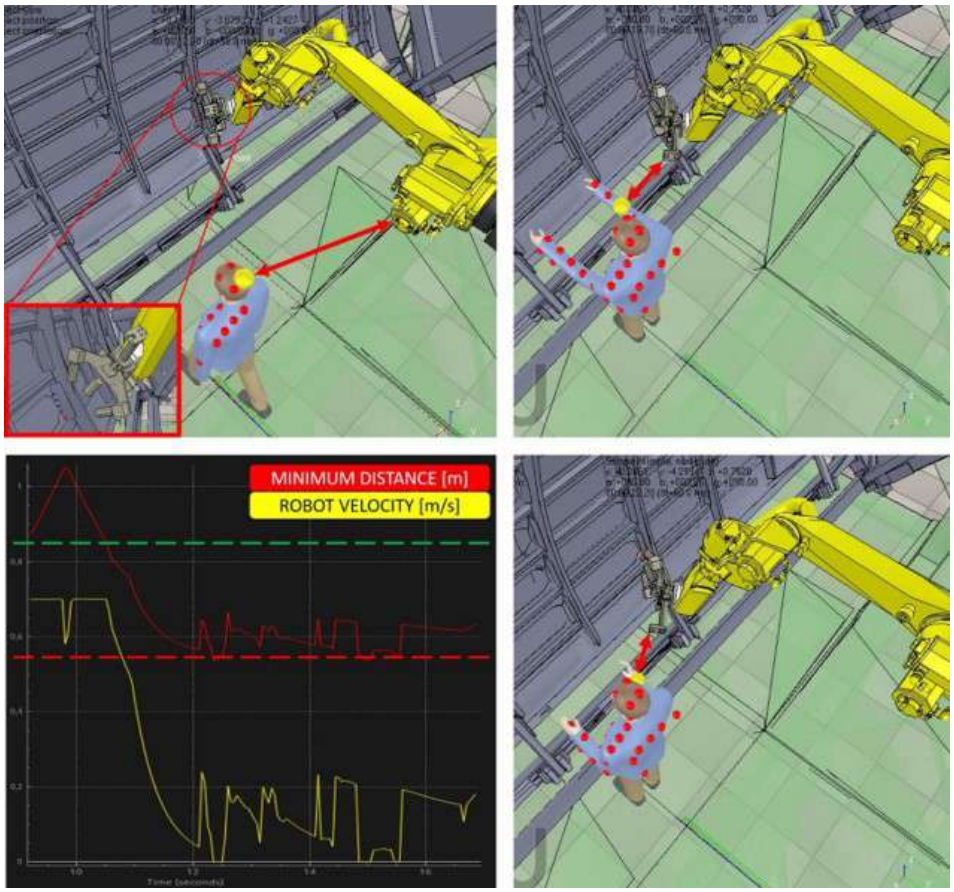


Figure 7.15.: Internal robot executing inspection task

Table 7.1.: General aircraft panel assembly operations without and with HRC

	FULLY-AUTOMATED		COLLABORATIVE	
OP ID	HUMAN	ROBOT	HUMAN	ROBOT
OP1	Assembly on SFRP skin	NA	Assembly on SFRP skin	NA
OP2	NA	Drilling	De-assembling	Drilling
OP3	NA	Countersinking	Cleaning and deburring	Countersinking
OP4	NA	Inspection	Re-assembling	Inspection
OP5	De-assembling	NA	De-assembling	Sealing
OP6	Cleaning and deburring	NA	Cleaning and deburring	Riveting
OP7	Re-assembling	NA	De-assembling	Inspection
OP8	NA	Sealing		
OP19	NA	Riveting		
OP10	NA	Inspection		

Legislative Decree no. 81/08, TUSL. The document presents the analytical description of the test procedure and the equipment used. This is a test of a prototype machine, in which some components are being tested. In particular, the software implemented relating to the HRC experimentation (robot speed modulation algorithms and operator-robot separation distance modulation algorithms), as well as the used vision system, do not comply with the minimum level of performance required by the UNI standard EN ISO 10218-2. Nonetheless, a risk assessment is provided which documents the risks to which workers may be exposed during the execution of the experimental test and the appropriate prevention and protection measures adopted within the regulatory framework on the safety of HRC applications are identified.

Description of the human-robot coexistence phases in the assembly cycle of the SIDE FWD panel

The identification of the risks required a careful analysis of the experimental procedure being tested, which has been collected in this subsection. The operations described are related to the assemby of the SIDE-FWD panel and they are performed simultaneously by the internal robot and the operator in the collaborative area. The reference work cycle is reported in Appendix

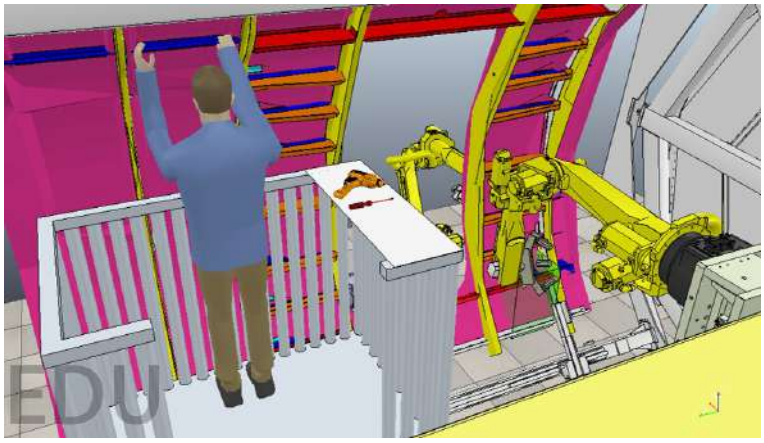


Figure 7.16.: Example of a manual assembly operation: the operator removes the blue element from the panel for the execution of deburring operations.

C.

Note: to avoid mechanical and electrical risks relating to the movement of tools, the drilling and riveting operations of the reference work cycle have been replaced with a waiting interval of the same duration. This choice does not cause variations or invalidate the result of the ongoing experimentation, but only guarantees greater safety for the operators involved. In other words, the experiment proposes to carry out a *dry test*, in which the riveting and drilling operations are simulated.

Figure 7.17 shows the work areas involved in the testing procedure. The yellow areas are assembled automatically by the internal robot while the blue areas are manually processed by the human operator.

More in detail, the test is based on the following assumptions:

1. The robot carries out the referencing operation in area 2 and subsequently carries out the complete cycle (operation no. 1040, Appendix C) according to the production instructions. At the same time, the operator enters the collaborative area and begins the climbing maneuver necessary to reach the work areas number 11 and 12. Subsequently, he performs the operations n° 20 - 60, then start applying the sealant (op. n° 70).

2. The robot moves towards the work area number 4 and, after having

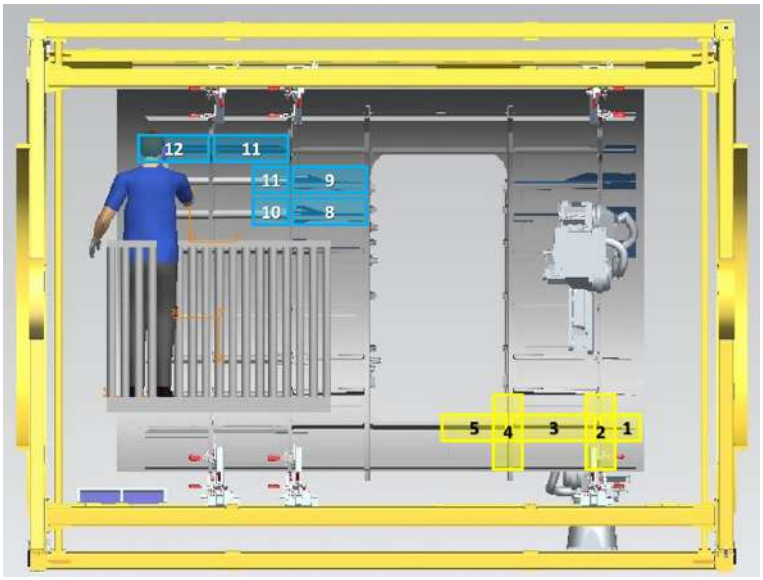


Figure 7.17.: Work areas considered for the experimental test procedure: the yellow areas assembled by the internal robot and the blue areas manually processed by the operator.

performed the referencing (op. n° 1050), repeats the same operations carried out in area 2 (op. n° 1060). The operator, simultaneously, finishes the manual assembly operations in area number 11 and 12 (from op. n° 70 - 90) and then slightly moves the scaffold platform (op. n° 100) to position himself in front of the processing areas number 9 and 11. Then he performs the application of the cleclos, the removal of the pop rivets and the clecos and begins the removal of the pax-door-intercostal support (op. n° 110 - 140).

3. While the operator completes operation no. 140, the robot guide is moved from the current position and brought to the pre-established position for processing work area number 3 (op. no. 1070). Once the referencing is finished, the robot begins the operation. n° 1080 while the operator carries out the deburring operation (op. n° 150) and begins applying the sealant (op. n° 160).

4. Once the operation n° 1080 is completed, the robot starts and completes the referencing operation in work area 1 (op. n° 1090) and then it starts the operation n° 1100 while the operator is completing the application of the sealant. When the latter has finished, operations from n° 160 to n° 180 are carried out during which the operator completes

processing of work areas number 9 and 11.

5. During the execution of the last manual operation in areas number 9 and 11, the robot starts the operation n° 1110 moving to the work area number 5, then approaching the scaffolding. In the meantime, the operator moves with the platform towards the processing areas number 8 and 10.

6. The operator completes the manual processing cycle in areas 8 and 10 (operations n° from 200 to 280) while the robot completes operation n° 1120 and is stopped, introducing a passive time of 122 seconds into the cycle.

7. The operator begins the exit maneuver of the scaffolding from the collaboration area, then exits the platform and transports the scaffolding outside the cell.

General protection measures

All general measures to protect the health and safety of workers have been observed, namely:

- An assessment of all risks to the health and safety of the workers involved during the execution of the collaborative operation has been carried out, as described below.

- Risks have been eliminated (where possible) or reduced to a minimum in relation to the knowledge acquired in the design phase.

- An ergonomic assessment has been carried out based on the OWAS method (not present in this thesis because it is outside the scope of this work) and the ergonomic principles expressed in the UNI EN ISO 6385:2017 regulation have been respected for the postures assumed by workers during the work procedure, for the choice of the equipment used and for the definition of the methods of work and production, in order to reduce the health effects of work.

- Risk reduction at source has been implemented as far as possible.

- Risk prevention and mitigation measures have been considered in accordance with the possible dangers highlighted and detailed in this section.

- The number of workers who are, or may be, exposed to risk has been

limited to a minimum.

- Adequate information and training of workers and managers on safety measures have been carried out.
- Adequate instructions have been given to all workers present during the entire execution of the experimental test.

Risk analisys and risk assessment

The strategy for risk estimation and mitigation is indicated in ISO 12100, in ISO 13849-1 and RIA TR R15.306, as described in Appendix B.

The analysis of the process preceded the adoption of procedures to reduce the estimated risks, referring to the hierarchy of measures defined in the standards (Table 7.2).

Table 7.2.: ISO EN 12100:2010 - Hierarchy of risk reduction process measures.

Risk level	**Risk mitigation**
Negligible	Hazard elimination or risk reduction by design
Low	Risk reduction through protection and possible complementary protective measures
Medium	Reducing risk by making information available for use
High	Residual risk

The approach chosen for risk assessment is the one based on the *risk matrix*, that is the most suitable for this context. Although this is a subjective evaluation, it appears to be the only one applicable to the experimental process in question since no historical data is available. The only data evaluated are those collected during the experimental phases of the implemented algorithms. The selected risk matrix uses four levels for each risk factor, illustrated in Table 7.3.

The meanings of the severity levels reported in Table 7.3 are described below:

- *Catastrophic*: irreversible injuries that make it very difficult, or impossible, to continue working (death or disability or permanent illness);
- *Severe*: generally irreversible injuries or serious and debilitating ill-

Table 7.3.: Selected risk matrix.

Exposure	**Catastrophic**	**Severe**	**Moderate**	**Mild**
Very probable	High	High	High	Medium
Probable	High	High	Medium	Low
Improbable	Medium	Medium	Low	Negligible
Impossible	Low	Low	Negligible	Negligible

nesses, with slight difficulty in continuing work after recovery;

- *Moderate*: more serious scratches, bruises and cuts, requiring medical attention from professionals (more care than first aid measures) but with the possibility of returning to the same job;
- *Mild*: no lessons or mild lessons (scratches, bruises) that can be treated with first aid measures and require little or no loss of work time.

The meanings of the probability of harm occurring reported in Table 7.3 are described below. They considered: the frequency and duration of exposure to a hazard; the staff who carry out the activities; the history of the machine/activity; the working environment; the reliability of safety functions; the possibility of neutralizing or circumventing protection measures; the ability to maintain protective measures; the ability to avoid harm.

- *Very probable*: it is almost certain to happen;
- *Probable*: it can happen;
- *Improbable*: it is not likely to happen;
- *Impossible*: the probability of this happening is close to zero.

Once severity and probability were assessed, an initial risk level could be derived from the selected risk matrix. The risk matrix allows to correlate risk factors to risk levels, as illustrated in Table 7.3. For example, the table shows how the combination of "Serious" and "Probable" gives a risk level of "High". The result of combining the risk factors of severity and probability generates low to high levels of risk. Since the risk assessment process using the risk matrix is subjective, the risk levels are also subjective.

Hazard identification produced a list of hazards, dangerous situations and dangerous events. Table 7.4 describes the identified possible hazards that

can occur during the execution of the entire experimental test, while Table 7.5 assigns the risk level to each one, related to the risk matrix of Table 7.3. Note that the identified hazards have been distinguished in the following macro-categories:

1. Operator: Positioning of the scaffold;
2. Robot: Movement of the linear stage in the direction of approaching to the scaffolding;
3. Operator: assembly execution (operation execution on the scaffold);
4. Risks due to failure to comply with ergonomic principles;
5. Risks due to failures or malfunctions of the vision system;
6. Risks due to failures or malfunctions of the control system;
7. Risks due to non-compliance with general requirements for collaborative operations;
8. Transversal risks.

Table 7.4.: Hazard identification.

ID	Hazard	Consequences
1.1	Ergonomic nature: pushing/pulling scaffolding in the collaborative area	Fatigue, bending, twist
1.2	Mechanical nature: impact of the scaffolding with hooks present on the walking surface	Impact, stumble, fatigue, unease
1.3	Mechanical nature: fell from the stairs while ascending/descending the scaffolding	Slipping, stumble, fall
2.1	Mechanical nature: sudden acceleration without stopping at the target position	Robot-scaffold collision, crushing, drag, cut, entrapment, shearing
3.1	Mechanical nature: slippery platform/wet shoes	Slipping, stumble, fall
3.2	Mechanical nature: the scaffold is not stable	Operator falls, scaffold overturning

Continued on next page

Table 7.4 Hazard identification (continued from previous page)

ID	Hazard	Consequences
3.3	Ergonomic nature: the workpiece is not blocked/tightened, has fallen from the platform or panel	Panic, human errors, involuntary and incorrect behaviors
3.4	Ergonomic nature: accidental fall of clecos/pop rivets/instrumentation from the scaffold	Panic, human errors, involuntary and incorrect behaviors
3.5	Ergonomic nature: contact with rough surfaces, with sharp edges and corners, with protruding parts	Cuts, injuries, bruises
3.6	Thermal nature: objects or materials at high temperatures/contact with sealant	Burns, injuries
3.7	Generated by noise: suction system/rivet recovery system/hissing of pneumatic components/worn parts	Unease, stress, lack of balance
3.8	Generated by vibrations: instability of the scaffold caused by the movement of the other components of the system	Unease, lumbar disorders, neurological disorders, osteoarticular disorders, vascular disorders
4.1	Incorrect positioning of emergency stop devices	Inability to instantly activate the emergency protective stop function
4.2	Unnecessary effort when carrying out assembly operations/Flicker/Glare/Shadow/Lighting of the elements to be assembled	Unease, fatigue, musculoskeletal disorders, stress
4.3	Combination of hazards: repetitive activity + effort + high ambient temperature	Dehydration, loss of consciousness, heat stroke
4.4	Distraction/failure to observe the operator supervising his colleague on the scaffolding	Loss of direct visibility of the collaboration area, inability to instantly activate the emergency protective shutdown function

Continued on next page

Table 7.4 Hazard identification (continued from previous page)

ID	Hazard	Consequences
4.5	Distraction/failure to observe the operator on the scaffold during manual movement of the platform	Panic, loss of direct visibility of moving parts, unease
5.1	Mechanical nature: breakage of the camera interface support	Incorrect estimated measurements, absent measurements
5.2	Mechanical nature: breakage of the vision system support bracket	Incorrect estimated measurements, absent measurements
5.3	Electrical nature: data communication is slow or lost	Incorrect estimated measurements, absent measurements
6.1	Sudden shutdown of one or more software modules	Absence of signals/data for correct monitoring of the robot's speed
6.2	Malfunction of the human operator recognition system: production of false positives	Incorrect estimated measurements, unnecessary robot slowdown, unnecessary increase in production times
6.3	Malfunction of the human operator recognition system: production of false negatives	Incorrect estimated measurements, robot speed greater than the maximum allowed for the current scenario
6.4	Malfunction of the separation distance monitoring module: errors due to background subtraction/clustering procedure affecting the estimation of minimum distance points	Incorrect estimated measurements, unreliable robot speed estimate
6.5	Speed monitoring module malfunction: the estimated speed is higher than the maximum allowable	Incorrect estimated measurements, robot speed higher than the maximum allowed
7.1	Lack of visual indication when the collaborative operating mode is active in the robotic cell	Less operator attention, failure to comply with the general principles of the collaborative operations
7.2	Failure to activate the speed limits determined by the risk analysis	Robot speed higher than the maximum allowed

Continued on next page

Table 7.4 Hazard identification (continued from previous page)

ID	Hazard	Consequences
8.1	Work-related stress	Psychological stress
8.2	Resulting mental stress from the fear of a collision with the robot	Psychological stress

Risk mitigation

The LABOR work cell has been designed in compliance with the Machinery Directive. For all the details relating to the analysis of risks linked to the general use of the machine (full-automated tasks) and the technical solutions implemented for their mitigation, i.e. access barrier, emergency circuit or manual mode limitations, please refer to the use and maintenance manual and to the final documentation of the system (it is not reported because it is considered outside the thesis work goal).

A description of the specific PPE that have been adopted in the work cell during the execution of the HRC task has been presented below.

1. *Protective measures against impacts, slips and blows*
    - Provide for the use of a protective helmet
    - Provide for the use of safety footwear (EN 345)
2. *Protective measures against punctures, cuts and abrasions*
    - Provide for the use of protective gloves (EN 388-4342)
3. *Prevention and protection measures against noise*
    - Implement all possible noise reduction measures
    - Provide for the use of earplugs for hearing protection SNR 25 dB(A)
4. *Prevention and protection measures against inhalation of dust and fumes*
    - Make sure that the operator is not hit by the flow of dusty air generated by the suction and/or the ventilation system of the work cell

Table 7.5.: Preliminary risk assessment

ID	Severity	Probability	Risk Level
1.1	Mild	Probable	Low
1.2	Mild	Very probable	Medium
1.3	Severe	Probable	High
2.1	Catastrophic	Improbable	Medium
3.1	Moderate	Probable	Medium
3.2	Catastrophic	Improbable	Medium
3.3	Mild	Probable	Low
3.4	Mild	Probable	Low
3.5	Severe	Very probable	High
3.6	Moderate	Very probable	High
3.7	Moderate	Very probable	High
3.8	Mild	Very probable	Medium
4.1	Catastrophic	Improbable	Medium
4.2	Mild	Probable	Low
4.3	Severe	Improbable	Medium
4.4	Catastrophic	Improbable	Medium
4.5	Catastrophic	Improbable	Medium
5.1	Catastrophic	Impossible	Low
5.2	Catastrophic	Impossible	Low
5.3	Catastrophic	Improbable	Medium
6.1	Catastrophic	Improbable	Medium
6.2	Mild	Improbable	Negligible
6.3	Catastrophic	Improbable	Medium
6.4	Catastrophic	Improbable	Medium
6.5	Catastrophic	Improbable	Medium
7.1	Catastrophic	Impossible	Low
7.2	Catastrophic	Improbable	Medium
8.1	Mild	Probable	Low
8.2	Mild	Probable	Low

- Provide for frequent cleaning of the work environment, using vacuum cleaners equipped with filters to avoid the reintroduction of fine dust into the work environment
- Provide for the use of protective glasses
- Provide for the use of FFP2 airway protective masks

5. *Prevention measures against vibrations*
   - Install additional mechanical stops to reduce the possibility of transmission of vibrations from the work cell movement system to the scaffold on which the operator works during the execution of the manual assembly operations
6. *Transversal prevention measures*
   - Training operators about the operations to be carried out, regarding coexistence, risk exposure and use of active and passive PPE.

For the HRC experimental demonstration on the SIDE FWD panel, redundant emergency stop buttons have been provided: the devices used are buttons that can be easily operated with the palm of the hand and positioned at each control station of the operators involved.

As well as the safety-related parts of the control circuit that perform the emergency stop function already present in the machine as general safety systems, the redundant buttons dedicated to the HRC have also been designed in compliance with ISO 13849-1. In particular, they were selected in accordance with the requirements of ISO 10218-2 which requires that these parts comply with $PL_r = d$ in category 3.

In accordance with ISO 10218-2, the robotic system has a single emergency stop function, which can be activated by any button installed in the work cell and acts appropriately on the entire system. Therefore, the emergency stop occurs without creating additional dangers or requiring further interventions, therefore it implements a stop function of the two robots, simultaneously with the stop of the linear axes and the jig. It should be noted that the tools installed on the robots have been designed and built in compliance with ISO 10218-2 and that the safety measures regarding their emergency stop functions do not entail additional dangers for the operator.

Table 7.6 presents the risk mitigation measures applied to each hazard iden-

tified in Table 7.4, while Table 7.7 presents the assessment of the residual risk associated with them.

In conclusion, the identified possible risks have been assessed with the risk matrix method and the residual assessment of Table 7.7 showed how the adoption of adequate mitigation actions made many of them negligible.

Some risks, however, continue to be present after the residual assessment: among them stand out those due to the incorrect positioning of the emergency stop devices, the risks due to failure to comply with the procedures given to the workers and the risks caused by accidental malfunctions of the experimental algorithms that implement the SSM collaborative pipeline. For this class of risks, unfortunately, no mitigation measures have been found sufficient to bring them to a negligible or low risk level.

The residual risk associated with the dangers indicated above, is however admissible according to the standards as it presents an estimated Medium/Low risk level and, even if it cannot be reduced by additional protective equipment, must be avoided through staff training/information.

7.2.7. Description of the test procedure

The employees involved in carrying out the experimental test with their related tasks are described below.

- Operator A executes the manual assembly on the scaffold;
- Operator B monitors the collaborative work cell;
- Operator C moves the scaffold guided by operator A;
- Operator D executes the software verification from the HRC workstation.

The experimental procedure involves the following steps:

1. Deactivation of all safety functions intended for non-collaborative operation;
2. Operator D checks the correct functioning of the emergency stop buttons (information obtained from the status of the visual traffic light signaling systems of the LABOR work cell);
3. Operator D verifies the correct functioning of the vision system (in-

Table 7.6.: Applied risk reduction measures

ID	Risk reduction methods
1.1	Lubricate or replace scaffold wheels if worn
1.2	Use of an operator who supervises the first during the phase of entry of the scaffold into the work area
1.3	Gloves, non-slip footwear and helmet
2.1	Adoption of software limit switches (safety line) on the movement of the internal robot slide
3.1	Anti-slip gloves and footwear and helmet
3.2	Further mechanical stops
3.3	-
3.4	-
3.5	Gloves and Safety glasses/visors
3.6	Gloves
3.7	Hearing protection
3.8	Further mechanical stops
4.1	-
4.2	-
4.3	-
4.4	Use of two operators who supervise their colleague on the scaffold
4.5	Use of two operators who supervise their colleague on the scaffold
5.1	Interface check before starting collaborative mode
5.2	Bracket check before starting collaborative mode
5.3	Real-time control: emergency stop if the data update frequency is less than the minimum required
6.1	Real-time control: emergency stop in case of sudden shutdown of a software module
6.2	-
6.3	-
6.4	-
6.5	Real-time control: emergency stop in case of estimated speed greater than the defined maximum
7.1	Operator training and visual indication of the collaborative operating mode
7.2	Real-time control: emergency stop if speed limits are not defined
8.1	Training and information on operations to be carried out, risks and preventive and protective measures
8.2	Training of operators on operations, preventive and protective measures

Table 7.7.: Final risk assessment

ID	Severity	Probability	Risk Level
1.1	Mild	Improbable	Negligible
1.2	Mild	Improbable	Negligible
1.3	Mild	Probable	Low
2.1	Catastrophic	Impossible	Low
3.1	Moderate	Improbable	Low
3.2	Catastrophic	Impossible	Low
3.3	Mild	Probable	Low
3.4	Mild	Probable	Low
3.5	Moderate	Impossible	Negligible
3.6	Mild	Impossible	Negligible
3.7	Moderate	Impossible	Negligible
3.8	Mild	Impossible	Negligible
4.1	Catastrophic	Improbable	Medium
4.2	Mild	Probable	Low
4.3	Severe	Improbable	Medium
4.4	Catastrophic	Impossible	Low
4.5	Catastrophic	Impossible	Low
5.1	Catastrophic	Impossible	Low
5.2	Catastrophic	Impossible	Low
5.3	Mild	Impossible	Negligible
6.1	Mild	Impossible	Negligible
6.2	Mild	Improbable	Negligible
6.3	Catastrophic	Improbable	Medium
6.4	Catastrophic	Improbable	Medium
6.5	Mild	Impossible	Negligible
7.1	Mild	Impossible	Negligible
7.2	Mild	Impossible	Negligible
8.1	Mild	Impossible	Negligible
8.2	Mild	Impossible	Negligible

formation obtained from the status of the traffic light visual signaling systems of the LABOR work cell);

4. The jig is manually rotated until the pre-established position for the task is reached;

5. Operator A wears PPE (protective helmet, non-slip footwear, protective gloves, earplugs, protective glasses);

6. Operator B positions himself outside the cell while holding the emergency button;

7. The robots move to the configurations required for processing area 2 (see Figure 7.17) and begin their cycle;

8. Operator D activates the collaborative software architecture (information obtained from the state of the traffic light visual signaling systems of the LABOR work cell);

9. Operator A and operator C transport the scaffold inside the cell, carrying out the entry maneuver indicated in the user manual;

10. Operator C exits the LABOR cell;

11. Operator A activates the mechanical protection stops of the scaffold;

12. Operator A climbs onto the platform of the scaffold and begins the climbing maneuver necessary to climb over the jig beams and reach work areas number 11 and 12 (see Figure 7.17);

13. Operator A carries out the operations in the sequence and contents indicated in the assembly cycle (see Section *Description of the human-robot coexistence phases in the assembly cycle of the SIDE FWD panel* above) except for the operation of the drilling and riveting tools (dry test);

14. At the end of the collaborative procedure, operator D stops the robots;

15. Operator A begins the exit maneuver of the scaffold from the collaboration area, then exits the platform;

16. Operator C enters the cell and transports the scaffold out of the cell together with operator A.

Figure 7.18.: The LABOR robotic cell during a cooperative assembly operation: the internal robot is performing a quality inspection task while the human worker is dismantling a metal part for deburring after hole drilling.

7.2.8. Test results

The HRC demonstration has been performed on the full-scale robotic cell reported in Figure 7.18 (left), where the internal robot is performing a quality inspection task on the left side of the panel, while a human worker is dismantling a metal part on the right side of the panel after its drilling for deburring purposes. Figure 7.18 (top-right) shows the real-time digital twin of the LABOR cell in the Gazebo simulation environment (which is ROS compatible) in which it is possible to distinguish the thermal point cloud and the two spheres at minimum distance, the yellow one which represents the point at minimum human-robot distance belonging to the human operator surface, and the purple sphere that belongs to the robot mesh. Finally, Figure 7.18 (bottom-right) shows the DT-CNN frame in which there is the bounding box around the recognize human worker.

The results of the robot speed modulation algorithm, on the other hand, is reported in Figure 7.19. The robot moves at scaled speed when $S < d < \nu S$ and only for a few moments it moves at zero speed. In fact, the red zone at about 80 s highlights the time interval when the robot is completely stopped since the separation distance d is below S.

Figure 7.2.8 shows what would have happened if the minimum protective

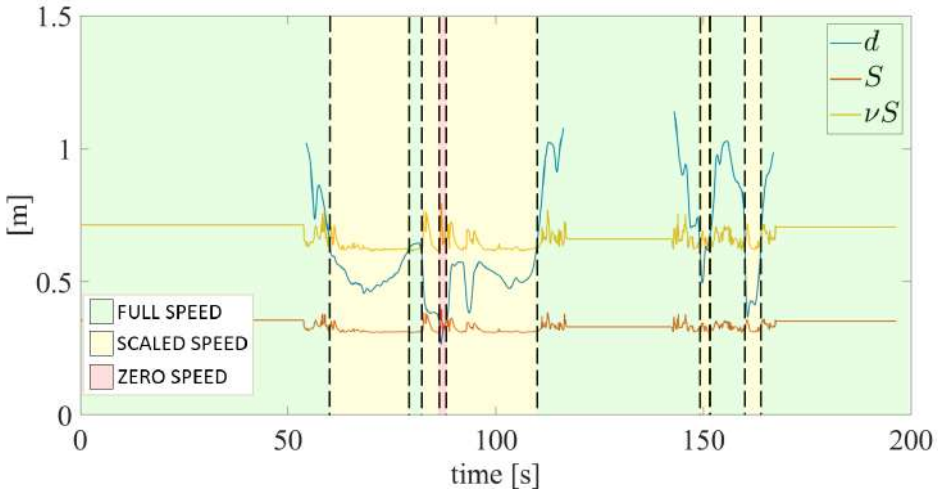


Figure 7.19.: Proposed method applied to LABOR work cell: $S(t)$ is computed online as in (6.4). The plot shows the comparison between the separation distance d (blue line) and the time-variant range $[S, \nu S]$. The robot mostly moves at full or scaled speed, while the complete stop area is minimal.

distance S had been a constant value calculated according to the standard Equation 3.3): the duration of robot stops is significantly larger, with a clear drawback in terms of robot work cycle time. Note that the robot moves at its nominal speed only when the human worker is not inside the collaborative area, i.e., the blue line representing the separation distance d is absent.

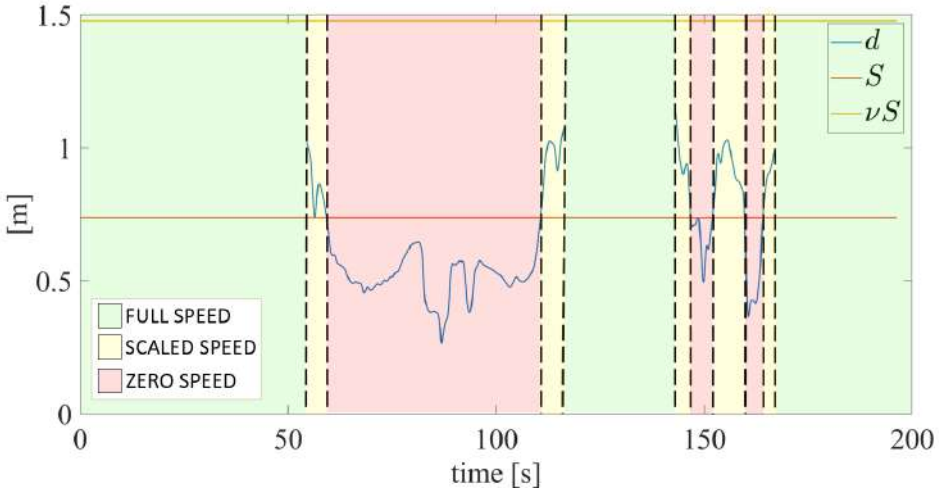


Figure 7.20.: Standard method: S is a constant value computed as in (3.3). The plot shows the comparison between the separation distance d (blue line) and the fixed range $[S, \nu S]$. The zero speed zones are very large, causing considerable increase of the production time. The robot moves at full speed only when the human operator is outside the collaborative area (green zone).

8. Conclusions

8.1. Brief description of the innovative prototype

An innovative smart monitoring technology is proposed for collaborative robotic work cells robustly prevents human co-workers from colliding with a moving robot without any hard safeguards. The application of the technology in collaborative tasks requires a detailed risk assessment prior the implementation on the factory floor. The technology is compliant with ISO 10218-1/2 and TS 15066 which form the current basis for safety in a collaborative robotic work cell. The proposed technology implements the SSM scenario, which controls the robot speed based on the separation distance from the worker. The proposed multi-modal perception system is composed of two cameras: a depth camera and a thermal camera. The key and innovative aspect of the technology is the developed sensor fusion algorithm which improves the robustness of the human tracking and the robot speed control algorithm, which avoids unnecessary robot slowdowns, thus increasing productivity, while preserving human safety. More in detail, the software pipeline starts with the identification of the human worker which is closest to the robot, thus computes the separation distance between him and the robot. The operator speed and direction of motion are estimated through an intention estimation strategy. The real-time data are collected to compute the minimum protective distance allowed by the ISOs. Thrugh a thermal point-cloud the software distinguishes points belonging to the body surface of the human operator closest to the robot from any tools held by the worker, e.g., a drilling tool or a screwdriver. After that, a robust and reliable speed scaling algorithm is executed: it is optimized to take on different behaviours during less risky situations or more dangerous ones for always guaranteeing both operator safety and minimum production time. The robot speed change is smooth for psychological advantages to operators: it changes from 0% to 100% of nominal velocity, not pre-defined speed levels. The basic idea of the proposed human detection algorithm is the development of an innovative CNN which merges spatial and thermal data to robustly identify human workers. Unlike the standard pre-trained RGB-CNNs as well as single-source CNNs, which can confuse objects with

shape or temperature similar to humans and wrongly classify them into the Human Class, the proposed CNN produces an almost zero percentage of false positives. Consequently, the speed scaling algorithm does not generate unnecessary robot stops, therefore maximizing the production time. The control system approach proposed here adopts a fuzzy inference logic that scales the nominal robot velocity according to some heuristic considerations which estimate the level of danger for the worker, thus implementing the risk assessment in real-time.

8.2. Summary of contributions

One of Industry 5.0's most challenging research areas is human-robot interaction and their aspirations to cooperate in collaborative workspaces. The goal of this research work is to make the existing regulations better in order to increase production time and ensure worker safety in industrial shared workspaces. The research activity focused on the design and development of a novel SSM collaborative pipeline. The work starts from the study of innovative methods for monitoring the robot's workspace in order to robustly detect the presence of the operator in real time and calculate his distance from the robot. The human-robot separation distance has been then used to modulate the robot speed on pre-programmed path without necessarily interrupting the execution of the task when the actual scenario is not properly dangerous for the human worker. In fact, the robot speed control is based on a risk assessment evaluation which is made in real time to properly analyze the hazard level to which the human operator is exposed to into the current scenario.

The main contributions which have been obtained during this research work are listed below:

1. Study of the SoA on workspace monitoring algorithms: approaches based on skeleton identification (skeleton-based), CNN-based approaches based on RGB images, point-cloud based approaches;

2. Study of the main safety standards for collaborative robotics applications: ISO 10218 1/2, ISO TS 15066, ISO 13855;

3. Design and development of a novel HDT algorithm: simulation in a VRep environment and programming in a ROS architecture of an innovative sensor fusion technique through the use of a depth camera and a thermographic camera; implementation of a new CNN trained with images obtained by combining spatial data and thermal data,

using the Yolo framework;

4. Study of the SoA algorithms for the minimum human-robot distance monitoring: identification of the main limitations of the case studies and design of possible operational solutions according to the requirements of the LABOR work cell;

5. Self-adaptive separation distance computation algorithm for mono/multi operators scenarios: simulations in the VRep environment and programming in the ROS environment of point-cloud based algorithms for the identification of the pair of points at minimum distance, one belonging to the robot and the other belonging to the closest operator;

6. Design and development of robot speed control algorithms: simulation in VRep environment and programming in ROS architecture of algorithms for the robot speed modulation (pre-programmed trajectories) based on sensory information, prediction of position and speed of human operators collaborating with robots on the basis of previous positions (Kalman filter integration) and integration of a Fuzzy logic control system for the optimization of cycle time, based on risk analysis;

7. Integration of the developed technologies: experimental validation of the developed algorithms at the @RobotisLab laboratory;

8. Research of instrumentation with industrial requirements suitable for the implementation of the developed algorithm: design and configuration of the multimodal vision system in the LABOR work cell with annexed integration of the PILZ safety instrumentation for the control of the mobile guards installed around the cell;

9. Study and implementation of calibration procedures for multi-modal vision systems in industrial environments: intrinsic and extrinsic calibration of the thermal sensor and depth sensor installed in the LABOR work cell;

10. Fine tuning of the HDT algorithm: optimization and configuration of sensor fusion techniques and of the new DT-CNN in an industrial environment;

11. Integration of robot speed scaling algorithms into the Fanuc robot controller of the LABOR work cell: smooth robot speed change for psychological advantages to operators (0% to 100% of nominal velocity, not pre-defined speed levels);

12. Final experimental validation of the proposed HRC module in the LABOR work cell: risk assessment and implementation of the experimental tests envisaged by the GA of the research project.

8.3. Final remarks

The maturity level of the technology has been assessed within the European project LABOR [123] with a target Technology Readiness Level (TRL) 7 – System prototype demonstration in operational environment. This project belongs to the H2020 Clean Sky 2 private-public partnership for research and demonstration of advanced technologies for regional aircrafts with the LEONARDO Aircraft [118] leadership.

Brief videos showing the capabilities of the prototype:

- *A multimodal approach to human safety in collaborative robotic workcells*: https://www.youtube.com/watch?v=VgJT5ohOHcI
- *A multimodal approach to human safety in collaborative robotic workcells*: https://www.youtube.com/watch?v=cVWxaciNy_g
- *A fuzzy inference approach to control robot speed in human-robot shared workspaces*: https://www.youtube.com/watch?v=RzLZ6RQBPCY
- A *Multimodal Perception System for Detection of Human Operators in Robotic Work Cells*: https://www.youtube.com/watch?v=BrcvKmSiR9Q

8.4. Challenges and future directions

8.4.1. Safety certification and compliance

The full integration of the advanced vision system into a CE robotized application plays a pivotal role in enhancing the capabilities of robotic platforms, enabling them to perceive and interact with human operators and to allow HRC tasks. As these collaborative systems become increasingly integrated into various industries, ensuring their certification and usability is of paramount importance. As explained in Appendix A, the safety certification framework for both the proposed vision system and the SSM algorithm encompasses a comprehensive set of guidelines and standards established by regulatory bodies, including but not limited to the ISO and the RIA. Compliance with these standards is imperative to ensure that the proposed

hardware and software solution operates safely in different environments and scenarios.

Section 7.2 describes the evaluation of the potential hazards associated with the specific application demo. It represents a preliminary risk assessment made up with the aim of being able to perform the experimental test on the real cell and extract the results set out above. The general safety certification can be accomplished by expanding this preliminary risk assessment and by considering factors such as system malfunctions, environmental uncertainties and every kind of hazards related to human-robot interaction, not only relative to the specific application. Every identified risks have to inform the development of safety mechanisms and protocols to mitigate and manage these potential hazards effectively. Note that the compliance of the proposed vision system extends beyond the physical hardware to include robust software safety measures. This means that fail-safe mechanisms, real-time monitoring and redundancy in critical components -which have not been completely implemented now- have to be integrated. The actual software has to be integrated with extra routines to detect anomalies, trigger emergency shutdown procedures when necessary and provide feedback to the operator or central control system, especially when something unexpected happens.

In conclusion, the safety certification and compliance of the proposed vision system are integral aspects of its development and deployment within the field of robotics. By adhering to established standards, conducting thorough risk assessments, and implementing hardware and software safety measures, the vision system aims to enhance the overall safety profile of robotic platforms, fostering their integration into many industries while prioritizing the well-being of human operators.

8.4.2. From case study to manufacturing

The strategy allows to distinguish the cases of the presence of a mannequin or shapes similar to human body with respect to the real human one. This feature makes the network usable for different purposes such as video surveillance or for general manufacturing HRC tasks. In addition, the algorithm could also have robust results in a clothing store of the future, if applied to a robotic moving platform: in a common service store there are operators, mannequins, clothes and many forms that recall the human features which can be robustly distinguished by this technology.

As possible future developments, the performance of the sensor fusion al-

gorithm could be improved by increasing the frequency of publication of the merged image video stream. To increase the frequency of publication, the sensor fusion algorithm, as well as the depth-thermal image mapping algorithm could be written in CUDA programming language and then run on the GPU. This can significantly increase the software performance.

In addition, the networks trained on a thousand image samples could significantly increase the performance for human detection by increasing the number of samples of the training dataset and doing the network retraining.

Possible improvements could be to increase the spatial coverage of the detection to monitor a larger collaborative workspace. The immediate solution could be to install a greater number of depth and thermal cameras with different angles and FoV and, then, merge the images into a one, resulting image to be processed. In addition, the sensor fusion algorithm could be extended by merging the use and fusion of other types of sensors into the vision system, by using the third channel (B) which now is empty. It would then be interesting for the application not only to carry out the human detection but also to make a prediction on the type of task or action of the human starting from his movements. Predicting a certain human behavior could favor the algorithms development that safeguard the security and that helps humans in work or daily activities.

Appendices

A. Safety standards for robotic systems

Standards are defined norms that have been approved by experts and released by an organization. They cover a large range of subjects, including sustainability, IT security, environmental management, and also robotics safety.

A.1. Organizations

Organizations issuing standards, technical specifications, norms and codes relevant to safety of robotic systems can be classified into three categories: international, national and local. Obviously, the most well-known organizations are those that operate on an international scale. It should be mentioned that standards released by national organizations may have international applicability It is even more true in emerging topics where national institutions have not considered yet. Note that standardization work at both national and international level and is the result of the work of experts in the corresponding field from different backgrounds. For example, for the international standardization work of the ISO/TC/299/WG3 group in charge of standards associated with industrial robots, the group consists of:

- industrial robot integrators;
- manufacturers of industrial robots;
- companies using industrial robots;
- academics/researchers;
- government agencies;
- prevention organizations.

The resulting international standard represents the best practices that may be used in that field and it is reached through consensus amongst experts from different nations.

International
The International Organization for Standardization (ISO) is an federation dating back to 1946 with a membership of more 150 national standard organizations. It has published tens of thousands of standards in numerous fields and has several hundreds of technical committees working on revising and publishing new standards, notably the ISO/TC/299 on robotics mentioned above. This technical committee also relies on the input from more grounded organizations, such as the Robotic Industries Association (RIA), which publishes guidelines designed for robotics applications based on existing standards.

National
Globally, there are over a hundred national standards organizations. They both publish their own standards and contribute to the development of international standards, such as those published by ISO. Among these national organizations, the US voluntary standards and conformity assessment system is administered and coordinated by the American National Standards Institute (ANSI), a private, non-profit organization. In addition, the Standards Council of Canada, a federal company that encourages voluntary standards in Canada, has accredited the CSA Group. Standards from CSA Group also have an international reach, being used notably in China, but their primary goal is to modify international standards to fit Canadian reality.

A.1.1. Machinery Directive (2006/42/EC)

Whether one uses HRC or not, all machinery must be compliant to the Machinery Directive (2006/42/EC). Every EU members have taken the Machinery Directive into national legislation to provide a common European level of protection for the safety and health of workers who use equipment in industrial environments. As a consequence, all systems made in the EU or imported into the EU must adhere to European safety and technical standards. The conformity is documented by the machine's CE-marking documents.

Since robots can be potentially harmful for humans, to fulfill the Machine Directive (2006/42/EC), international standards have been developed for designing and implementing safe industrial robotic systems. These stan-

dards can by classified into type A, B or C.

- Type A is basic safety standard and it is valid for all types of machines;
- Type B is generic safety standard and it valid for a spectrum of machines;
- Type C is machine safety standard which describes detailed safety requirements for specific machines.

A robotic system is compliant with the CE-marking if the risk assessment have been made, according to both ISO 12100 [124] (Type A) and ISO 13849-1 [125] (Type B) and the related necessary safety measures have been implemented, according to ISO 10218-1 [7] (Type C), ISO 10218-2 [8] (Type C) and ISO 15066 [4].

Note that the following sections mainly provide details on collaborative robotic applications. A general overview of international standards for safety and risk assessment in robotic applications is also provided, but further details on this last topic will be covered extensively in Appendix B.
Note that the following sections mainly provide details on collaborative robotic applications. A general overview of international standards for safety and risk assessment in robotic applications is also provided, but further details on this last topic will be covered extensively in Appendix B.

A.1.2. ISO 12100 (Type A)

ISO 12100 [124] is an international standard that plays a basic role in ensuring system safety because it lists general principles of machine safety. The document provides a comprehensive framework for assessing and mitigating risks associated with machinery, and this extends seamlessly to robotic systems. It outlines a structured methodology for risk assessment, which encompasses the identification of potential hazards, their evaluation in terms of severity and probability, and the subsequent design and implementation of safety measures. These principles are based on knowledge and experience of the design, use, incidents, accidents and risks associated with machinery [126]. Additionally, ISO 12100 emphasizes the importance of collaboration and communication between all stakeholders involved in the lifecycle of a robotic system, whether related to mechanical, electrical, or software aspects. This approach encourages transparency and information exchange, leading to a more comprehensive understanding of the potential risks and a more effective implementation of safety measures.

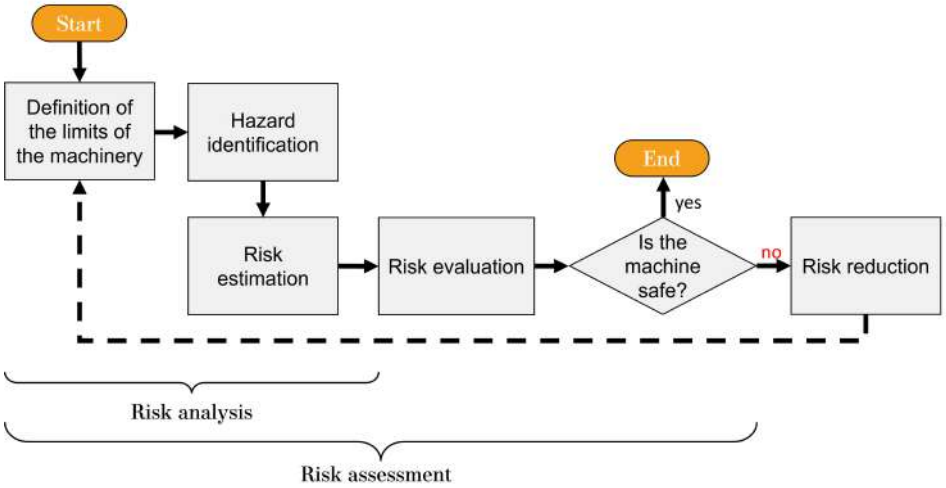


Figure A.1.: Risk assessment evaluation and reduction according to ISO 12100.

More specifically, ISO 12100 defines an iterative risk assessment process mandatory for the setup of any robot installation, as shown in Figure A.1. It starts with a functional and geometric design of the machine, from which spatial and usage boundaries are defined, e.g., reachable workspace, tasks and operator qualifications. Next, all tasks of the process and every hazards related to them are identified. Then, ISO 13849-1 defines a method, called *risk tree*, discussed in next Chapter, which is used to estimate risks due to three factors: potential injury severity, probability and/or frequency of exposure to the hazard and possibility to avoid or limit the hazard. Subsequently, it has to be decided upon if all hazards have been adequately addressed in the machine design. If not, the machine design is modified to reduce the specific risk and the risk assessment process is repeated.

ISO 12100 specifies that risk reduction need to be carried out in the following priority order:

1. Risk reduction through safe design;
2. Risk reduction through safety devices;
3. Risk reduction through information.

A.1.3. ISO 13849-1 (Type B)

ISO 13849-1 [125] is another crucial standard in the realm of safety for machinery which works hand in hand with ISO 12100 to enhance both safety and functionality. When applied in tandem, ISO 12100 and ISO 13849-1 serve as a powerful duo, offering a holistic approach to safety in robotic systems. By adhering to both standards, manufacturers and integrators can not only develop highly functional and efficient applications but also ensure that the systems are designed with safety as a top priority, safeguarding human operators and the surrounding environment.

More in detail, ISO 13849-1 delves into the design of control systems with safety functions, which include the design of software. In the context of robotics, where control systems are central, it provides guidelines for designing, implementing, and validating these control systems to achieve the desired so-called Performance Level (PL). The PL of a machine can be evaluated through the so-called *risk tree* method proposed into the document itself and detailed in Chapter B.

A.1.4. ISO 10218-1/2 (Type C)

ISO 10218 standard describes the new concepts of industrial robot safety. Recognizing the increasing prevalence of robots in lots of industrial applications, ISO 10218 provides a comprehensive framework for ensuring the safe design, operation and interaction of these machines with human operators. The standard delineates safety principles and requirements that address both the mechanical and electrical aspects of industrial robots, with a particular focus on minimizing risks associated with HRI. The document plays a crucial role in establishing guidelines to guarantee the safety of such interactions, emphasizing features like safety-rated soft axis and speed limiting to prevent harm to humans in close proximity to the robots. ISO 10218 is critical in promoting the development and integration of industrial robots across various industries while ensuring that the safety of workers remains a paramount concern.

The standard is divided in two parts: ISO 10218-1 [7] presents guide lines for the assurance of safety in design and construction of the robot and for the safeguarding of human operators during robot integration, installation, functional testing, programming, operation, maintenance and repair; ISO 10218-2 [8] has been recently modified and it allows operator to cooperate with robots by using limited speed and power and additional safeguard installation, although the topic is not complete and obviously discussed. The

standard specifies four basic interaction levels for HRC, which are deeply discussed in Section A, and summerized below:

- Level 1: Safety-rated Monitored Stop: robot is not moving when human enters the shared workspace;
- Level 2: Hand-Guiding: robot can be moved by direct input including a dead man's switch;
- Speed and Separation Monitoring: given a minimum distance and a maximum velocity, the robot is allowed to move near the human operator;
- Power and Force Limiting: by inherent design or control - collisions are limited by force and pressure given to the human body.

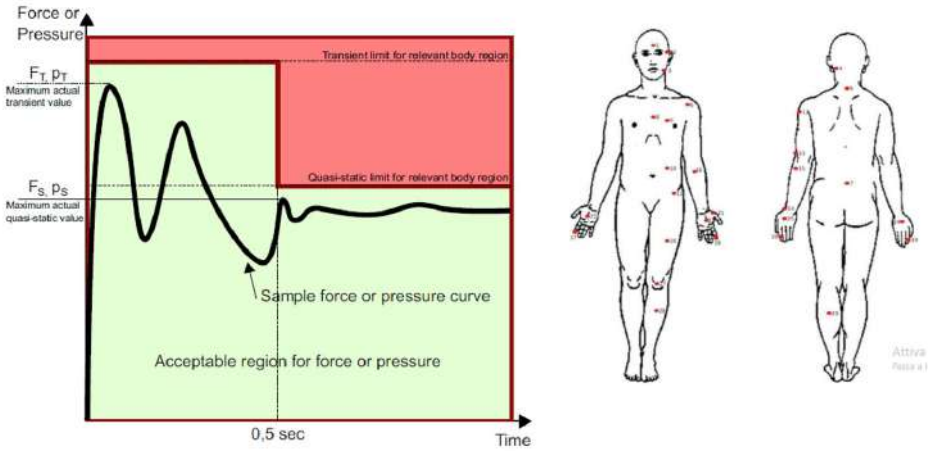
Additionally, the maximum velocity is limited to 0.25 m/s, the maximum affordable force to 150 N and the maximum dynamic power is bounded by 80 W, when collaborative robots are deployed.

A.1.5. Technical specification ISO/TS 15066

The application of collaborative principles can be difficult without guidance more detailed than that given in the two parts of ISO 10218. As a result, the responsible ISO working group has developed a so-called "Technical Specification" document, ISO/TS 15066 [4] which has become a supplement to ISO10218-1/-2 and provides additional detail on the safety requirements on collaborative robots and applications. Here, acceptance of collaborative systems is defined by measurements of impact in case of collision.

The document distinguishes two possible types of contact between moving parts of the robot system and areas on a person's body:

1. Transient contact: This is also referred to as "dynamic impact" and describes a situation in which a person's body part is impacted by a moving part of the robot system and can recoil or retract from the robot without clamping or trapping the contacted body area, thus making for a short duration of the actual contact. Transient contact depends on the combination of the inertia of the robot, the inertia of the person's body part, and the relative speed of the two actors.
2. Quasi-static contact: this includes clamping or crushing situations in which a person's body part is trapped between a moving part of the



(a) (b)

Figure A.2.: ISO/TS 15066 details: a graphical representation of acceptable and unacceptable forces or pressures (a) and the reference body model for contact areas identification.

robot system and another fixed or moving part of the work cell. In such a situation, the robot system would apply a pressure or force to the trapped body part for an extended time interval until the condition can be alleviated.

The robot speed must be limited so that the robot system is able to comply with the protective limits associated with the exposed body area, as shown in Figure A.2a. For the purposes of this specification, a body model including 29 specific body areas categorized in 12 body regions has been created. Figure A.2b shows the contact areas in the body model, while Table A.1 shows an overview of the specific body areas, classified into general body regions, where force and pressure limits for a transient or quasi–static impact are illustrated.

In general, the robot system shall be designed to adequately reduce risks to an operator by not exceeding the applicable threshold limit values for quasi-static and transient contacts. Risk reduction associated with transient contact could involve limiting the speed of moving parts (such as the robot, tooling, or workpiece) and an appropriate design of the physical characteristics such as the surface area of the moving part that could contact the operator. On the other hand, risk reduction associated with quasi-static could include speed limits and physical characteristics, similar to transient

Body region		Specific body area	Quasi-static contact		Transient contact	
			Maximum permissible pressure [a] p_s N/cm^2	Maximum permissible force [b] N	Maximum permissible pressure multiplier [c] P_T	Maximum permissible force multiplier [c] F_T
Skull and forehead [d]	*1*	*Middle of forehead*	*130*	*130*	*not applicable*	*not applicable*
	2	*Temple*	*110*		*not applicable*	
Face [d]	*3*	*Masticatory muscle*	*110*	*65*	*not applicable*	*not applicable*
Neck	4	Neck muscle	140	150	2	2
	5	Seventh neck muscle	210		2	
Back and shoulders	6	Shoulder joint	160	210	2	2
	7	Fifth lumbar vertebra	210		2	2
Chest	8	Sternum	120	140	2	2
	9	Pectoral muscle	170		2	

Figure A.3.: Biomechanical limits.

contact, plus design characteristics of the parts of the robot system that involve the possible trapping or clamping of an operator or body area.

B. Risk assessment in HRC

Related techniques for robot hazard assessment encompass a spectrum of methodologies and tools designed to ensure the safe deployment and operation of robotic systems in various settings. These techniques often involve risk analysis, which evaluates potential hazards, identifies their severity, and assesses the likelihood of occurrence. Hazard identification, task analysis, and the use of safety standards are common components of this process. Additionally, simulation and modeling techniques allow engineers to predict and analyze potential risks in a controlled environment before physical deployment. Furthermore, advances in artificial intelligence and sensor technologies have enabled real-time risk assessment and mitigation through features like collision detection and emergency shutdown mechanisms. The multifaceted approach to robot hazard assessment is crucial in guaranteeing the safety of both human operators and the surrounding environment in an increasingly automated world.

B.1. Related robot hazards

When robots were first used in industry, neither manufacturers nor users asked for any standards regarding safety. Robots were not designed for specific tasks, thus the primary feature of a robot design was its motion flexibility, which increased the possibility of injury to some degree. As time went on, the topic of robot safety attracted a lot of attention. Robots have a wide range of motion, may move continuously in up to n axes, can be flexibly programmed for various velocities and motions on each individual axis, and can interact with humans, other machines, and structures.

Many accidents, including fatal ones, have occurred in production factories over the years when robots have been used. The study [127] explores why, in spite of control measures, robot accidents that injure operators continue to happen and suggests remedial measures for risk assessments. In order to reconstruct the mechanism of damage, the job being done, and the bodily location of the harm, the study gathered 369 operator-injured robot accidents that occurred in Korea from 2009 to 2019 (10 years). According to

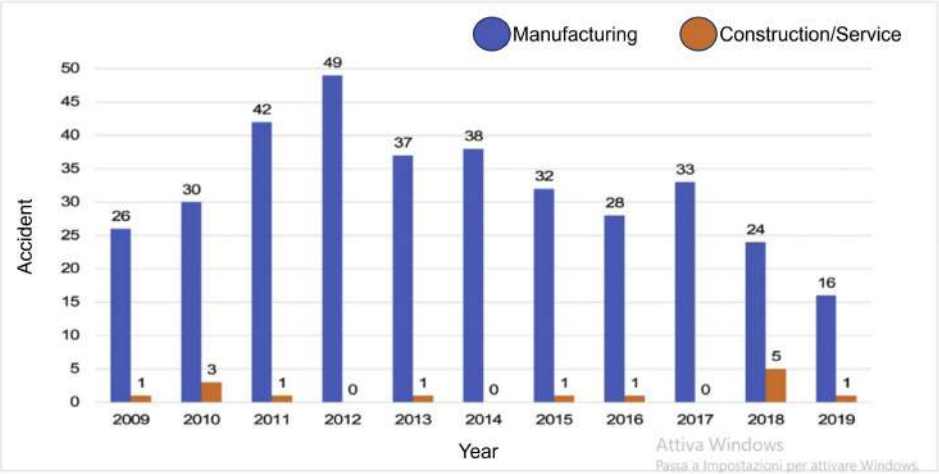


Figure B.1.: Annual robot accidents by industry (2009-2019), [127]

the statistical data collected in Figure B.1, the number of yearly robot accidents ranged from 27 to 49, the highest occurring in 2012 and the lowest in 2007. More specifically, more than 95% of the robot related accidents (355 cases) occurred in manufacturing businesses, while the remaining 14 were reported from the service and construction sectors. The primary reason for the high percentage of accident cases in manufacturing settings corresponds to the statistical fact that manufacturing businesses utilize 89% of industrial robots in total, as compared to 11% of service industries' utilization. The work also classified the 369 robot-related accidents by accident type. The result is that two types were predominantly responsible for 325 (88%) of the 369 accident cases: the first type, with 191 (52%) cases, corresponded to "jammed," "caught in," or "crushed" cases, and the second type, with 134 (36%) of the cases, corresponded to "collision" or "impact" cases. The first type occurred when a part of the body was stuck either between a fixed structure and robot arms or between fixed and moving parts. It also occurred when a body part was placed between more than two moving parts of industrial robots. In the "collision" or "impact" cases, most of them related to workers' irresponsible approaches to the robot or the sudden start-up of the robot when closely working near the machinery sites. Another investigation to be highlighted was concerned with injured body parts to understand the parts most exposed to hazards. The conclusion of the work was that the body parts that are most susceptible to injury when they meet with an accident. Hands and fingers are the most vulnerable, as seen in 113 cases (31%), followed by head/face and neck/shoulder/chest, amounting to 90 cases (24%) and 83 cases (22%), respectively. Other parts, such as the

abdomen and back, legs, and arms are also among the affected ones. Additionally, apart from operators' external injuries by robots, chronic injuries within robotic work environments such as noise-induced hearing loss, skin hypersensitivity, white knuckle due to vibration, etc. need to be analyzed as well.

In summary, robot accident origins can be categorized into three main classes: engineering, human behavior and environmental conditions.

- **Engineering**:

  The first category contains the failure of robot mechanism (electrical, mechanical), sensors, robot controller and associated equipment (electrical, mechanical, software). The consequences if these failures are abrupt motions, runaways, arm high uncontrolled speed, acceleration, force, energy ejections, etc.

- **Human behavior**:

  The second category contains human error factors which may originate from inadequate safety training, incorrect ergonomic workplace or equipment design, high task cognitive load, inadequate task distribution, etc. The consequences can be: loss in situational awareness, attention and hazard perception, unauthorized entry into dangerous work space, erroneous robot operation and task performance, etc.

- **Environmental conditions**:

  The third category is related to the conditions required for a normal robot and convenient human operations. This implies ambient temperature, humidity, lighting, noise and vibration levels, as well as ergonomic factors consideration in equipment and workstation design.

B.2. Standardized risk assessment

Risk assessment method is much used and discussed in some safety standard of robot, as presented in Section A, i.e. the international ISO 12100 [124], ISO 13849-1 [125] and the national ANSI B11.TR3-2000 [128] and another safety standard for industrial robots, written in conjunction with the RIA, the ANSI/RIA R15.06 [129]. The global risk assessment procedure for isolated industrial robotic system is depicted in Figure B.2. Note that particularities pertaining to collaborative and mobiles robots will be addressed in

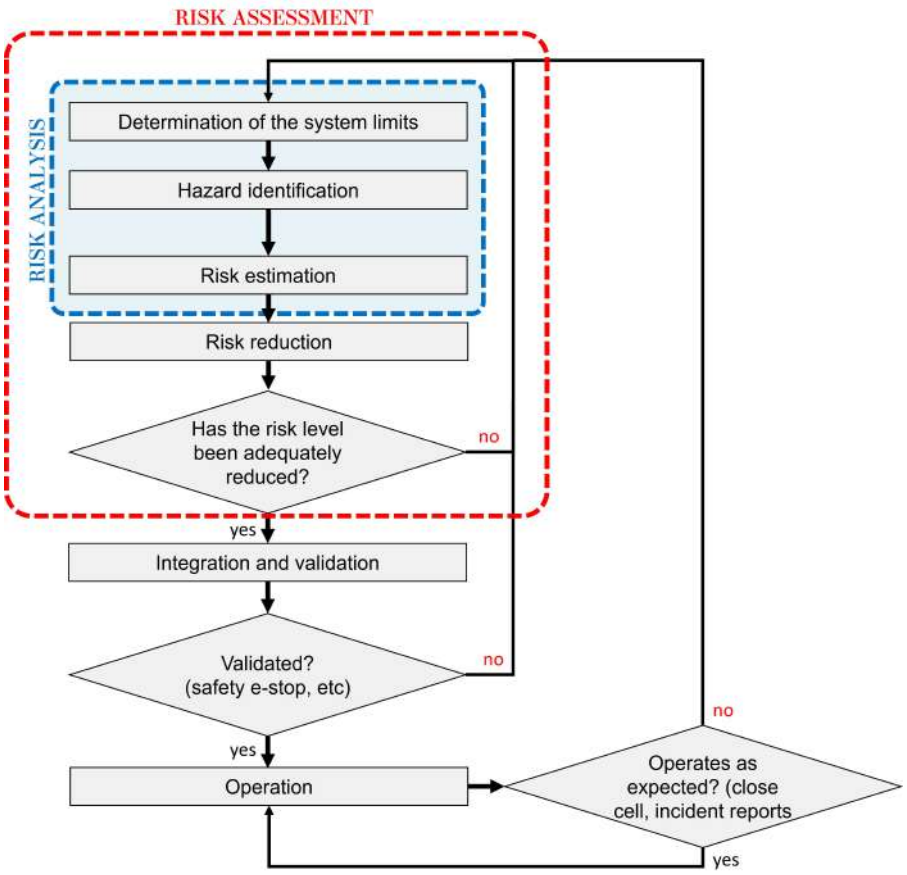


Figure B.2.: Risk assessment procedure (adapted from ISO 12100 [124]) with subsequent integration, validation and monitoring.

the subsequent sections.

The aim of risk assessment is to collect and produce information about the machine hazards to design and update the specification safety design. The required information for machinery risk assessment is to identify both the planned and also unplanned use of the machine, as well as its functions and structures. The risk assessment technique contains the following several steps:

1. Definition of the robot application field and identification of all the limits of the robotic system application in terms of use, space and time throughout its life cycle;

2. Identification of every potential sources of harm, known as *hazards*,

for each robot task: operational methods analysis, ways of interaction with human workers and the mechanisms failure probability rate estimations;

3. Estimation of risk category for each identified hazard in terms of probability, frequency and severity of the occurrence of an injury or damage;

4. Risk evaluation in order to decide if the projected risk is acceptable or not;

5. Adopt risk reduction measures when it is not acceptable, by installing the corresponding protection systems or applying standard procedures.

B.2.1. Risk analysis

The risk analysis is the first exercise done by the integrators to start the risk assessment procedure, identifying every potential hazards of the system and estimating their level of danger. As shown in the Figure B.2, the risk analysis consists of the first three steps of the risk assessment procedure. Note that the risk analysis should not take any risk mitigation strategies into account. This means that throughout the risk assessment, the operator should always be regarded as unqualified and the workspace as unprotected. This manner, all potential dangers may be quickly detected, preventing the omission of any prospective danger. In a way, this state represents the worst possible use of the system.

Determination of the system limits

This step amounts to determine the intended use of the machine and the environment in which it is to perform these functions. It is therefore essential to have proceeded, before the risk identification stage, to a functional design process of the robotic system. For instance, this stage includes defining the layout and place where the robot will be installed, the surrounding objects, the parts handled, the number of operators and their training, the tasks that the robot will perform, etc. This means that the step identifies the robot application areas and all constraints related to the intended use, i.e. layout, time, dynamics, kinematics, mechanical constraints, software requirements, etc.

Hazards identification

Identification of every potential sources of hazards is an iterative process, thus the analysis is conducted again after implementing the risk-reduction measures (mitigation) to validate the desire outcome has been reached. Note that the approach is unique for the industrial robotic system application to be evaluated: the hazard identification, as well as the next step of risk analysis, is specific to a particular machine and installation. Therefore, it is necessary to carry out a new task/hazard identification when the environment, tasks or operators change (as part of a machine move, for example).

Table B.1 provides a non-exhaustive list of possible causes of hazards with isolated industrial robotic systems (as opposed to collaborative and mobile robots, which will be discussed later). This table takes material damages into account in addition to injuries to the human body. It is important to remember, nevertheless, that international standards only make reference to the latter when the word "harm" is used. The ISO 10218-2 Annex A contains a list of significant hazards.

Risk estimation

After completing the potential hazards identification, the last step of the risk analysis is to select the criteria quantification to determine the risk level of each hazard. According to ISO 12100:2010, it is necessary to estimate two aspects or parameters for each risk that has been identified.

1. the severity of harm and;
2. the probability of that harm.

The latter normally comprises three subparameters:

1. the exposure of the person(s) to the hazard;
2. the occurrence of a hazardous event;
3. the possibilities to avoid or limit the harm.

To rank the probability and severity of the injury, there are several risk estimating tools available. Not every one of those instruments, though, covers the same number of parameters. For the risk level estimation, two equivalent and very similar methods are presented below: the first method is described in the international standard ISO 13849-1 [125] while the second method is described into the national standard RIA TR R15.306 [130].

Table B.1.: Non-exhaustive list of potential causes of hazards involving industrial robotic systems, [8].

Categories	**Examples**
Workers related	Unqualified operator incapable of controlling the robot Human error Unauthorized access to the robot workspace Not following the manufacturer's instructions
Control	Interference Software error Bad programming
Power	Contact between various electrical cords in the robot system Leaks in an hydraulic system
Mechanical	Deficient/broken part
Environment	Sources of electromagnetic interferences

ISO 13849-1

As anticipated in Section A, ISO 13849-1 delves into the design of control systems with safety functions. In the context of robotics, where control systems are central, it provides guidelines for designing, implementing, and validating these control systems to achieve the desired so-called Performance Level (PL). The PL of a machine can be evaluated through the so-called *risk tree* method proposed into the document itself. The result of the risk assessment should be a well-argued and documented report, summarized in a so-called Required Performance Level (PLr) of the machine.

PL is the discrete level used to specify the ability of safety-related parts of a control system to perform a safety function under predictable conditions. PLr, otherwise, is the mandatory performance level of a machine to reduce risk. It is important to note that the PL must be greater than or equal to the PLr because the more risk a machine poses, the higher level of performance

it needs:

$$PL \geq PLr \tag{B.1}$$

Performance levels are shown Figure B.3 and range from PLa to PLe, where a is the lowest and e is the highest. A small, slow mobile robot would likely require a lower performance level (closer to PLa), while a big, fast industrial robot would likely require a higher performance level (closer to PLe). The level of performance depends on the architecture of the control system, the reliability of the adopted components, the fault detection capability, as well as the resistance to the most common faults in multi-channel control systems.

For each safety function performed by a safety-related part of the control system it is necessary to determine the related PLr, which often follows these three parameters according to ISO 13849-1:

1. Severity of injury:
   - S1: Mild (normally reversible injury)
   - S2: Severe (normally irreversible injury or death)
2. Frequency and/or exposure to the hazard:
   - F1: Rare or infrequent (the exposure time is short)
   - F2: Frequent to continuous (the exposure time is long)
3. Possibility of avoiding risks or limiting damage:
   - P1: Possible under specific conditions
   - P2: Hardly possible

Using the corresponding parameters from Figure B.3, it is possible to determine the PLr of a machine. For example, if the severity of the injury is mild, start at S1. If the frequency and/or exposure to the hazard is frequent or continuous, let's continue for F2. Finally, if the possibility of avoiding the danger or limiting the damage is low, P2 is reached. This brings us to PLc, which means having to operate at this level or a higher one, therefore between PLc and PLe.

Note that Figure B.3 illustrates both the PL and the Safety Integrity Level

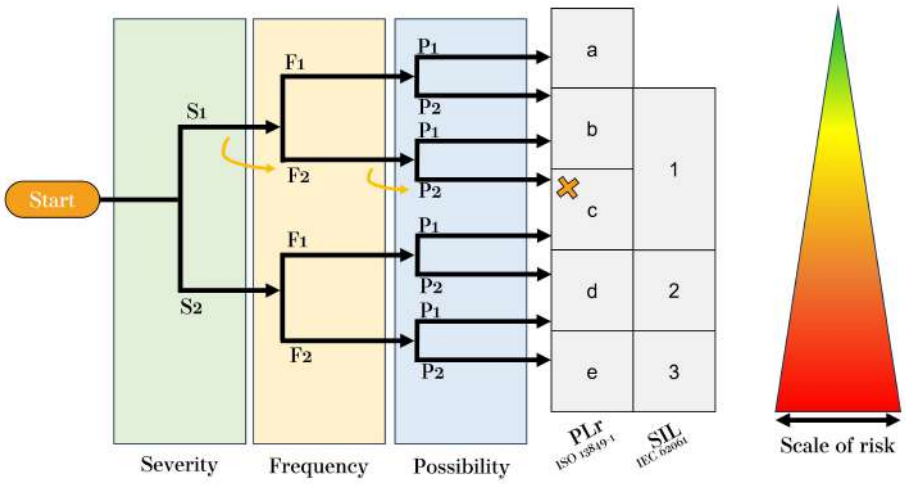


Figure B.3.: Risk tree method of ISO 13849 for Performance Level estimation.

(SIL). These metrics assign different security specifications, but they use similar thresholds to define the performance that should be applied to the robotic system according to its level of risk. The metrics are used by different government agencies: ISO 13849-1 defines performance levels, while IEC 62061 defines safety and integrity levels.

RIA TR R15.306

Similarly to what is proposed in the international standard, also RIA TR R15.306 proposes a national procedure for risk evaluation, as illustrated by the chart in Figure B.4. It comprises three parameters, namely severity, exposure and avoidance:

1. Severity of injury:
    - S1 - minor (bruises, etc.)
    - S2 - moderate (broken bone, short hospitalization, etc.)
    - S3 - serious (death, chronicle disease, amputation, etc.)
2. Exposure to the hazard:
    - E0 - prevented (not used in the initial analysis, since we ignore risk mitigation measures at this stage)

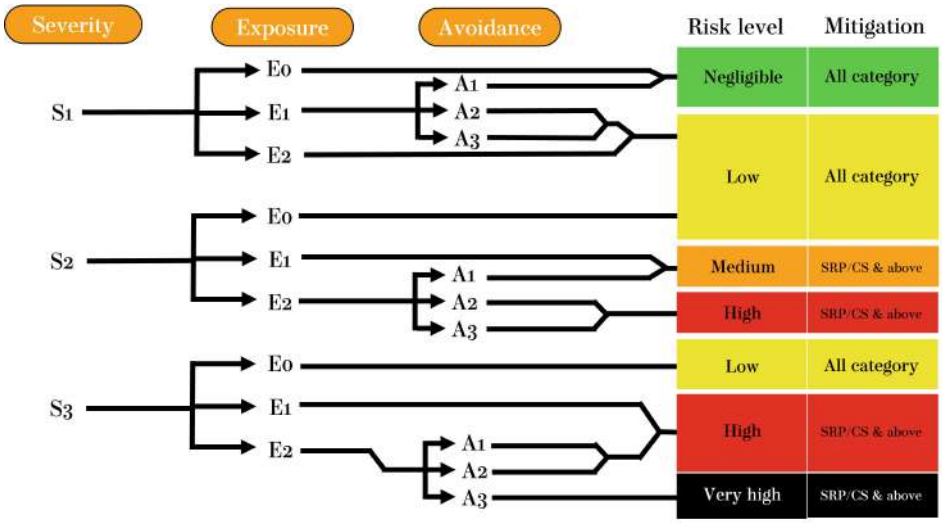


Figure B.4.: Chart to evaluate the risk level for each dangerous phenomenon (adapted from RIA TR R15.306 [130], which is a supplement of ANSI/RIA R15.06-2012 [129]).

- E1 - low (less than 1 time par day)
- E2 - high (more than 1 time par day)

3. Avoidance of the hazard:
   - A1 - probable (sufficient space, under robot speed limitation, early warning)
   - A2 - improbable (insufficient space, but under robot speed limitation, obstructed exit)
   - A3 - impossible (insufficient space, caged operator)

B.2.2. Risk reduction

After risk analysis has been made, a standard approach for risk mitigation is the one shown in Figure B.5 and proposed in RIA TR R15.306-2016. The method requires to apply all necessary measures in a hierarchical order, where the primary step should be always the hazards elimination by the work cell redesign, while the next steps should involve the incorporation of safeguarding technologies, training, warning procedures, and personnel safety equipment definition. Preventive and corrective measures must be

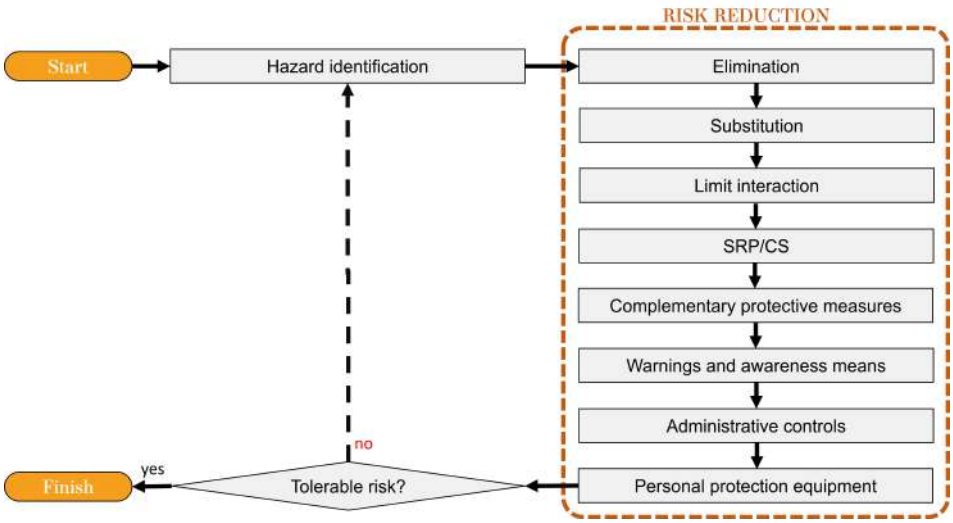


Figure B.5.: Generalized risk reduction algorithm stated in the robot safety standard, [129]

put in place and the risk index will then be reevaluated accordingly.

Some examples of devices belong to the hierarchical categories listed below are displayed in Figure B.6. Only a brief summary of some mitigation measures will be given here and it is mandatory to refer to the relevant standards/technical specifications/guidelines for more information. Note that the list below is applicable to a common industrial robotic system, while more specific preventive solutions for collaborative scenario -which extend the industrial ones- will be detailed in Section B.6.

1. **Elimination:** changing the architecture of the industrial robot system in terms of hardware, software, process or layout. It is important to note that this step changes the whole system design in order to intrinsically remove the risk. For example, removing obstacles that could be the cause of the jamming risk can eliminate the risk of jamming without the need of guards and barriers.

2. **Substitution**: changing materials handled, and replacing the robot by another less powerful, slower or with a smaller workspace.

3. **Interaction limitation**: limitating physical interactions between the human operator and the industrial robot within the latter's workspace.

4. **Safeguarding and safety-related part of a control system (SRP/CS)**:

The simplest and most popular way to reduce the risk of someone getting into close contact with an industrial robot is to surround it with a perimeter fence, which acts as a barrier between the robot and anyone it might injure entering its workspace. A barrier is required if the potentially harmful mechanism (which isn't always a robot) is less than 2.5 meters above ground, according to ISO 13857. In this case, the same standard suggests rigid panels with specific specifications based on the system's measurements and workspace. For example, the minimal height of the panels is 1.8 m, regardless of the system. These fences may have doors for maintenance, thus they must be equipped with sensors or a locking mechanism to detect intrusion into the robot workspace.

5. **Complementary protective measures**: Complementary precautions that are frequently used include handrails, mechanical blocks, extra padding, e-stop software functions and safe access to the robot.

6. **Warnings and awareness means**: Rotating beacons, alarms, warning panels are among the measures used to increase the operator's awareness of the potential dangerous phenomenon.

7. **Administrative controls**: Organizational-type measures are also essential to reduce risks involving robotic systems. Indeed, some risks cannot be completely eliminated, thus a proper training of the staff is critical to increase awareness. The information shared includes the nature of the risks, existing protection methods, proper safe ways to approach the robot, etc. Other measures include compliance with the manufacturer's instructions, regular inspections and preventive maintenance of the robot, rewarding workers for safe behavior, etc.

8. **Personal protection equipment**: Common examples include glasses, helmet, boots, etc.

B.3. Integration, Validation and Monitoring

Following the completion of the risk assessment, which consists of both the risk analysis and the risk mitigation, the last step is to re-analyse the possible hazards by quantifying the risk parameters previously indicated, but this time accounting for the risk-reduction measures implemented. The computation of the new risk levels will enable the implementation of the robotic system integration. The validation step occurs after the system integration. ANSI/RIA R15.08 [131] suggests the following approach for

(a) (b) (c)

Figure B.6.: Mitigation measures that can be applied to industrial robot cell: emergency stop (a), safety light curtain (b) and lockout procedure (c).

any robotic system:

- visual inspection;
- practical tests;
- measurement;
- observation during operation;
- review of application-specific schematics, circuit diagrams and design material;
- review of task-based risk assessment;
- review of specifications and information for use.

Doing safety stop tests, for example, can be useful as a validation step to make sure that the robot stop time and the stop safety function's response time have both been considered when determining where to place virtual barriers, which are objects whose presence is detected by proximity sensors near the robot workspace.

Users must then continuously monitor the system in order to reevaluate the risk using new data obtained from experience feedback (incidents, close calls, etc.).

B.4. Hazards inherent to collaborative tasks

Beyond any potential sources of hazards detailed in Section B.2, there are other possible hazards that are more specific in HRC tasks. Obviously, the close proximity of human operators to the robot is the main source for many of them, but some others are linked to the task itself. A non-exhaustive list is detailed in Table B.2, as well as corresponding mitigation measures:

Table B.2.: Non-exhaustive list of potential causes of hazards involving collaborative robotic systems.

Risk type	**Action**	**Mitigation**
Physical risks	collisions, crushing, jamming, repetitive impacts, tool used by the robot	lightweight robot, rounded surfaces, safe speed limitation, safe force and power limitation, training
Psycho-social risks	isolation, pace difficult to follow by the operator, work transformation	improving the working conditions of workers
Risks of musculoskeletal disorders	high repetitivity, excessive efforts, high precision required, inadequate posture that may be required for extended periods of time	arranging workstations to respect the comfort zones, using appropriate handling techniques, optimizing lighting, choosing the right tools

B.5. Risk assessment and mitigation measures for HRC

Risk assessment for collaborative applications follows the same procedure to that previously shown in Figure B.2, referred to isolated industrial robots. The difference is about the more interactive actions that need to be taken into account during the use case identification, i.e. the first stage of the risk assessment procedure, which then have repercussions in small changes to the subsequent steps of the risk assessment, as specified in ISO/TS 15066 [4].

B.5.1. Use case identification

As mentioned in Section B.2, the information to record in this stage includes the task to execute, the lifetime phase of the machinery in which it takes place, the qualification level of the personnel involved and the frequency of occurrence of the particular use case. An example of a use case overview for a collaborative task is shown in Table B.3.

Table B.3.: Example overview of use cases of a collaborative task.

Use case	**Description and frequency**
UC1: Setup and programming	Specially trained operator Rare
UC2: Normal production	Simply trained operator Frequent
UC3: Manual intervention	Simply trained operator Infrequent
UC4: Foreseeable misuse	Untrained operator Rare
UC5: Maintenance	Specially trained operator Infrequent
UC6: Cleaning	Untrained operator Infrequent
UC7: Dismantling	Untrained operator Rare

Examining these use cases shows that there is a risk of both transitory contact with the operator's hand or lower arm and quasi-static contact for use cases UC1–UC5. However, in use cases UC6 and UC7, the robot system is turned off, eliminating any possibility of a collision with the moving robot. A not-isolated robotic cell would expose anyone close to danger. In comparison to conventional robotic systems, this will lead to more stringent safety regulations. As a result, the use case definition step has to specify exactly who, what level of expertise, and how often they might be exposed to risk. Figure B.7 illustrates many possible human operator roles, and Table B.4 provides a thorough explanation of each job, [132].

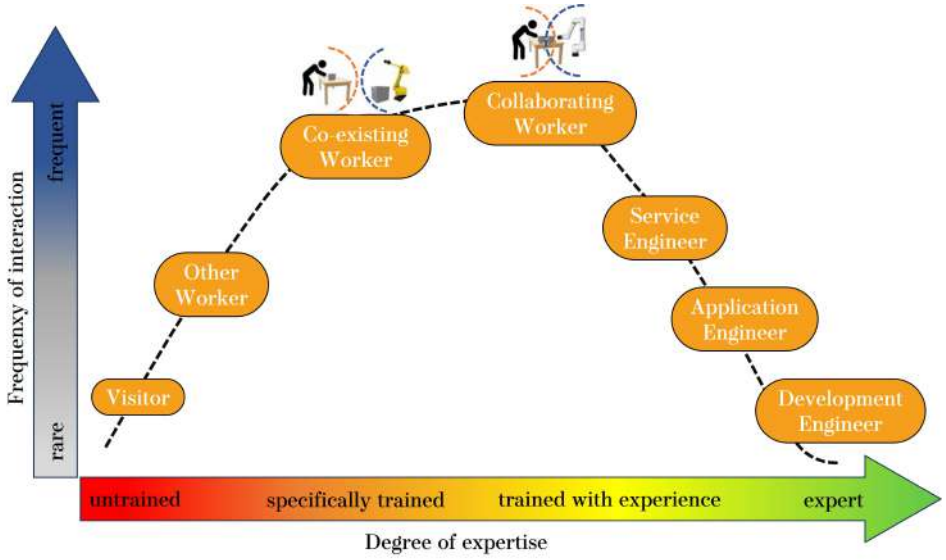


Figure B.7.: Roles of human workers, frequency of exposure and degree of expertise.

Table B.4.: Description of interacting personnel, [132].

Role and description	**Worst unsafe interaction**
Visitor: internal or external people, coming to robot for single visit, uninformed about possible hazards	enters robot workspace although advised not to: collision with robot and hazards occur
Other Worker: employees occasionally coming close to the robot, without any particular assignment involving the robot	enters robot workspace although advised not to: collision with robot and hazards occur (reasonably foreseeable misuse)
Co-existing Worker: employees working physically in an overlapping workspace with robot; assignment does not comprise any interaction with robot	hits or is hit by the robot accidentally, resulting in any of the listed hazards
Collaborating Worker: interacts with robot (manipulator, end effector, work pieces, fixtures, part tray, etc.) in regular operating mode	hits or is hit by the robot due to error in procedure or machine, resulting in any of the listed hazards
Service Engineer: interact with robot, reconfigure, repair and exchange and replace devices, reprogram, recalibrate robot and sensors, execute service routines	safety systems may be disabled, insulation protection may be removed; hit by robot resulting in any of the listed hazards; plus additional hazards such as electrical shock
Application Engineer: same as service engineer plus collaborating worker plus installation, introduce new production processes, de-	interacts with robot when safety systems may be disabled; hit by robot resulting in any of the listed hazards; plus additional

B.5.2. Hazard identification

Regarding the hazard identification in collaborative tasks, the contact situations can be distinguished in quasi-static(constrained) contact and transient(unconstrained) contact, as described in Section A. Three basic categories of hazards related to a HRC task have been detailed below: mechanical or electronic hazards, ergonomic hazards, and cognitive hazards. Every interaction level has its own set of risks, which are influenced by the task specification, operator responsibility, human-robot cooperation distance and human-robot collaboration procedure. Physical and cognitive factors should be taken into account, e.g. when a task demands a lot of mental and physical energy to complete, there is a higher chance of error.

Mechanical and electrical hazards

- **Tasks**: welding, painting, cutting, assembling, drilling, milling;
- **Description**: crushing, trapping, collision, stored energy, electrical shock, burn, poisoning, pressure, shearing, cutting, severing;
- **Causes**: failure of robot parts, human error, failure of control, software failure, firmware failure, safeguarding failure, incorrect work planning, incorrect task design, incorrect task sharing, incorrect time process scheduling, inadequate installation, inadequate usage, incorrect robot trajectory planning, physical obstacles against robot operation during collaboration, hazard from fast worker approach speed and robot's slow reaction time, non-provision of transition from collaborative operation to manual system in case of system malfunction;
- **Consequences**: robot sudden movements, unintended movement of associated machines or parts.

Ergonomic hazards

- **Tasks**: assembling, loading, moving, handling parts;
- **Description**: strain/pain, physical fatigue, hearing loss, visual loss, wrong risk protection;
- **Causes**: ergonomic design deficiency for operation and maintenance, excessive physical load, inadequate task plan design, insufficient work cell design, poor GUI design, incorrect work conditions, wrong task

distribution, inefficient work planning, failure of robot parts or other machinery, faulty installation, incorrect usage of the system, not optimized spatial arrangements, safety features insufficiency, potential hazards from the industrial process (e.g., temperature, loose parts etc.);

- **Consequences**: erroneous task performance, risk taking behavior, elevated noise level and long term exposure, effect on the hearing and balance, poor awareness, poor perception of acoustic signals, insufficient lighting, visual awareness loss, high hazard exposure, hight risk likelihood.

Cognitive hazards

- **Tasks**: any collaborative activity;
- **Description**: fear/anxiety, mental fatigue, stress
- **Causes**: mental stress to operator due to collaborative industrial process, personnel hazard perception, excessive task cognitive load, poor control panel design, poorly designed user interface, bad work conditions, incorrect task distribution;
- **Consequences**: unsafe behavior, erroneous work, task misunderstanding, misuse of the system, recognition of hazards and hazardous situations is obscured, erroneous work, unsafe behavior.

B.5.3. Risk estimation

As explained in Section B.2, the risk related to the considered hazard is a function of severity of harm, that can result from the considered hazard, and the probability of the occurrence of that harm, that means: exposure of person(s) to the hazard, the occurrence of the hazardous event and the possibility to avoid or limit the harm.

Currently, the risk of injury in human-robot collisions cannot be accurately estimated using standardized quantification values and factors. This is because basic statistics from robots are lacking, and the knowledge of the human body's mechanisms underlying systemic harm under all load circumstances is incomplete. A great deal of research has been done recently, mostly in the automotive industry, to provide injury criteria for the classification of collision injury effects. There is still a long way to go before these

low-level injuries (mainly soft-tissue injuries) can be systematically classified to assess the severity of the associated low-level hazards. Currently, initial experimental measurements can only be performed under extreme circumstances, such as with crash test dummies. Thus, while they offer a useful foundation for defining test method design, they do not describe the necessary harm criteria or constraints for human-robot interaction.

When it comes to HRC activities, the degree of an accident can vary from no injuries to fatalities. Less serious injuries like fractures can arise from manipulator clamping or direct compression loads, while more catastrophic injuries like cuts or abrasions might result from contact with sharp or abrasive surfaces. The existing statistical results can be divided into two categories: collisions and clipping locations, which are body portions compressed between robot parts or between the robot and specific external objects. Fingers, hands, head and chest are the most common body parts in potential accidents, as described in Section B.1.

B.5.4. Risk mitigation

As explained in Section B.2, the categories that should be applied for risk mitigation to achieve an intrinsically safe design of the system were elimination, substitution and interaction limitation. Most of them are obviously not applicable to HRC tasks according to the definition of the collaborative task itself, thus they will translate to lower energy consumption, compliant material surfaces for robots, modified tasks, etc. As a result, different solutions should be considered for different interaction level to reduce the risk. Possible solutions to reduce risk in each type of collaborative operation are listed below.

1. **Level 1: Safety-rated Monitored Stop**
    - Safety button to authorize the operator to enter and exit from collaborative working area;
    - camera or other sensors to monitor the working area;
    - training and maintenance.
2. **Level 2: Hand-Guiding**
    - Trajectory planning;
    - emergency stop;

- controls close to the end effector;
- training and maintenance.

3. **Level 3: Speed and Separation Monitoring**
   - Sensor monitor the distance and speed;
   - protective stop if minimum distance or speed limit is reached;
   - consider braking distance in minimum separation distance;
   - training and maintenance.
4. **Level 4: Power and Force Limiting**
   - Sensor monitor the end-effector speed, torque and force;
   - end effector brakes when contact with operator;
   - low inertial;
   - avoid sharp edges;
   - soft skin where can contact with operator.

Avoiding contact with the shoulders and over them is mandatory, as well as it must be avoided any robot movements above them. Note that considering it may not always be realistic, experts on the standardization committee working on ISO 10218 update are proposing to replace the verb "shall" with "should," still strongly encouraging to keep the robot's movements below head level.

For other contacts, ISO/TS 15066 contains specifications on the onset of pain, as shown in Figure B.8, as well as transient contact speed limits. An example of risk mitigation in a PFL collaborative scenario is illustrated in Figure B.9: first of all (1) it is needed to remove the wall which represents a pinch and a crush point, then (2) reduce robot system inertia or mass and (3) also its velocity. Finally, to reduce the risk of potential injuries, (4) it is suggested to modify the robot posture such that contact surface area is increased and (5) moved away from sensitive upper body parts.

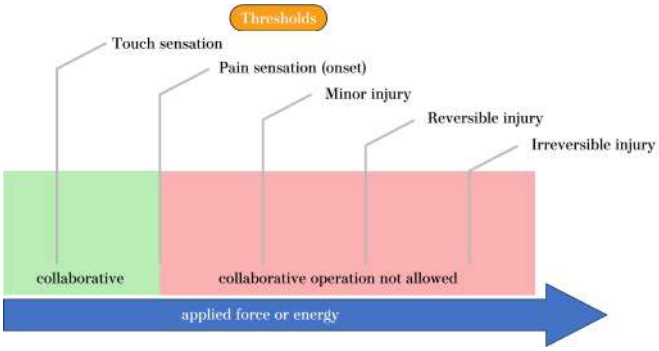
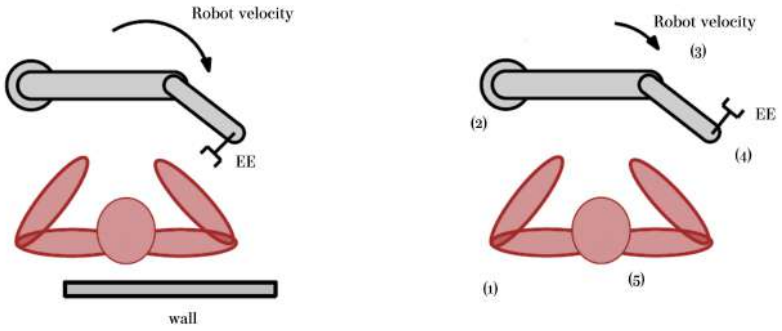


Figure B.8.: Investigation of the onset of pain in relation to collaborative operation (inspired by ISO/TS 15066).



(a) (b)

Figure B.9.: Risk reduction in a PFL collaborative operation inspired by ISO/TS 15066): before risk mitigation (a), after risk mitigation (b).

B.6. Preventive solutions for HRC

As specified in Section B.2, alongside preventative solutions for industrial robotic system, there are some specific devices adopted in case of collaborative scenarios. They can be described in five classes of safeguarding systems, which are the most efficient standardized levels, as seen in the general risk reduction procedure.

1. **Devices for presence detection**
   Many robotics safety measures for human presence identification include laser scanners (Figure **??**a), light curtains (Figure **??**b) and pressure-sensitive mats (Figure **??**c). They were created especially to enable interactive tasks where human operators and robots work together in a common workspace. These devices are frequently employed to detect human movement within a robot's workspace and into potentially dangerous areas, allowing the robot to stop moving as needed. These sensor classes —ultrasonic detectors, passive and active infrared sensors, capacitive and pressure sensing units, etc.— are commonly used for human detection. Robot grippers can also be fitted with force, capacitive, radar, photo-electric transducers, cameras, range finders, and other sensors to improve their "awareness" of their surroundings and to regulate their own operating conditions.

2. **Interlocked fences**
   Non-sensor safety devices, such as interlocked barrier guards and fixed barriers (fences) are typically placed around robots to limit the robot workspace as shown in Figure **??**d.

3. **Warning system**
   Audio and video alarms, i.e. flashing, lamps, warnings and awareness barriers, belong to this category.

4. **Personal protection equipment**
   Foot switches, teach pendant equipped with enabling switches and emergency stop are example of devices belong to this category, as well as special protective clothes more specific for the collaborative task to be done or some wearable equipment's protective.

5. **Safety controller circuit**
   Every safety devices are connected to the safety controller of the system. The control system can be equipped by programmable safety controller (PLS) or by a modulator with direct or remote monitoring of integrated safety systems or by safety relays.

Discussing of safety of a collaborative scenario completely depends on the capability of the robot to manage unexpected events occurrence, as failures or unforeseen environment changes. Fault handling and fault tolerant control should be considered as essential functionalities to be developed both for hardware and software point of view, during safety interaction between humans and robots, as described in [133]. A model of failure classification has been provided in the book [134]. It should be noted that there is more complexity involved in the implementation of HRC tasks. It is important to clearly define of the types of faults that can affect the robot to acceptable levels of robot reliability in HRI, and it should be considered during development and utilization.

In practice, it is never possible to completely handle every potential dangerous event. The robotic system should identify any occurrences or malfunctions of the collaborative scenario and determine their type and location. For safe interaction and high tolerance collision prevention, an appropriate programming of the behavior robotic system, for example, with robust control strategies, can be assured. It requires an appropriate analytic combination (e.g., FMECA, FTA) and evaluations (e.g., through stochastic modeling or experiments) to attained reliability in order to determine whether the modified procedures are adequate and suitable.

Appendix C

ROBOT				HUMAN WORKER			
OP	DESCRIPTION	TIME [s]	AREA	OP	DESCRIPTION	TIME [s]	AREA
ROTATION 5 SECTION 5				ROTATION 5 SECTIONS 1,2			
1030	referencing working area 2	60	2	10	Elevator positioning	60	Section 1, Area 11-12
1040	Perform complete cycle of drilling, countersinking, hole inspection, head change, sealing, fastener installation, fastener inspection and flushness according to Manufacturing Instruction	579	2	20	Clecos application	32	Section 1, Area 11-12
				30	Pop rivets removal	48	Section 1, Area 11-12
				40	Clecos removal (4)	28	Section 1, Area 11-12
				50	Pax door intercostal support removal	80	Section 1, Area 11-12
				60	Execute deburring	390	Section 1, Area 11-12
1050	referencing working area 4	60	4	70	Sealant application	317	Section 1, Area 11-12
1060	Perform complete cycle of drilling, countersinking, hole inspection, head change, sealing, fastener installation, fastener inspection and flushness according to Manufacturing Instruction	579	4	80	Sealed package installation	80	Section 1, Area 11-12
				90	Clecos application	120	Section 1, Area 11-12
				100	Elevator positioning (little lateral and down movement)	25	Section 2, Area 9-11
				110	Clecos application	24	Section 2, Area 9-11
				120	Pop rivets removal	24	Section 2, Area 9-11
				130	Clecos removal (3)	21	Section 2, Area 9-11
				140	Pax door intercostal support removal	40	Section 2, Area 9-11
1070	referencing working area 3	60	3	150	Eseguire deburring	195	Section 2, Area 9-11
1080	drilling, countersinking, hole inspection according to Manufacturing Instruction	201	3				

				160	Sealant application	240	Section 2, Area 9-11
1090	referencing working area 1	60	1				
1100	drilling, countersinking, hole inspection according to Manufacturing Instruction	201	1	170	Sealed package installation	40	Section 2, Area 9-11
				180	Clecos application	64	Section 2, Area 9-11
1110	referencing working area 5	60	5	190	Elevator positioning (little lateral and down movement)	25	Section 2, Area 8-10
				200	Clecos application	24	Section 2, Area 8-10
1120	Perform complete cycle of drilling, countersinking, hole inspection, head change, sealing, fastener installation, fastener inspection and flushness according to Manufacturing Instruction	579	5	210	Pop rivets removal	24	Section 2, Area 8-10
				220	Clecos removal (3)	21	Section 2, Area 8-10
				230	Pax door intercostal support removal	40	Section 2, Area 8-10
				240	Execute deburring	195	Section 2, Area 8-10
				250	Sealant application	240	Section 2, Area 8-10
				260	Sealed package installation	40	Section 2, Area 8-10
PASSIVE TIME= 122				270	Clecos application	64	Section 2, Area 8-10
				280	Elevator exit	60	Section 2, Area 8-10

Bibliography

[1] European Commission. Directorate General for Research and Innovation. *Industry 5.0: towards a sustainable, human centric and resilient European industry.* Publications Office, 2021.

[2] J. Krüger, T.K. Lien, and A. Verl. Cooperation of human and machines in assembly lines. *CIRP Annals*, 58(2):628–646, 2009.

[3] Vaibhav V. Unhelkar, Ho Chit Siu, and Julie A. Shah. Comparative performance of human and mobile robotic assistants in collaborative fetch-and-deliver tasks. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, March 2014.

[4] UNI. ISO/TS 15066 - Robots and robotic devices - collaborative robots. Standard, International Organization for Standardization, 2016.

[5] Jeremy A. Marvel. Performance metrics of speed and separation monitoring in shared workspaces. *IEEE Transactions on Automation Science and Engineering*, 10(2):405–414, April 2013.

[6] Jeremy A. Marvel and Rick Norcross. Implementing speed and separation monitoring in collaborative robot workcells. *Robotics and Computer-Integrated Manufacturing*, 44:144–155, April 2017.

[7] UNI. ISO 10218-1 - Robots and robotic devices - Safety requirements for industrial robots. Part 1: Robots. Standard, International Organization for Standardization, 2011.

[8] UNI. ISO 10218-2 - Robots and robotic devices - Safety requirements for industrial robots. Part 2: Robot system and integration. Standard, International Organization for Standardization, 2011.

[9] Abdelfetah Hentout, Mustapha Aouache, Abderraouf Maoudj, and Isma Akli. Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017. *Advanced Robotics*, 33(15-16):764–799, July 2019.

[10] Carole S. Franklin, Elena G. Dominguez, Jeff D. Fryman, and Mark L. Lewandowski. Collaborative robotics: New era of human–robot cooperation in the workplace. *Journal of Safety Research*, 74:153–160, September 2020.

[11] Jonas Schmidtler, Verena Knott, Christin Hölzel, and Klaus Bengler. Human centered assistance applications for the working environment of the future. *Occupational Ergonomics*, 12(3):83–95, September 2015.

[12] Balasubramaniyan Chandrasekaran and James M. Conrad. Human-robot collaboration: A survey. In *SoutheastCon 2015*. IEEE, April 2015.

[13] Michael A. Goodrich and Alan C. Schultz. Human-robot interaction: A survey. *Foundations and Trends® in Human-Computer Interaction*, 1(3):203–275, 2007.

[14] Tobias Kopp, Marco Baumgartner, and Steffen Kinkel. Success factors for introducing industrial human-robot interaction in practice: an empirically driven framework. *The International Journal of Advanced Manufacturing Technology*, 112(3-4):685–704, December 2020.

[15] Roland Behrens, Jose Saenz, Christian Vogel, and Norbert Elkmann. Upcoming technologies and fundamentals for safeguarding all forms of human-robot collaboration. In *8th International Conference Safety of Industrial Automated Systems - SIAS 2015*, 11 2015.

[16] Nana Wang, Yi Zeng, and Jie Geng. A brief review on safety strategies of physical human-robot interaction. *ITM Web of Conferences*, 25:01015, 2019.

[17] Marco Faber, Jennifer Bützler, and Christopher M. Schlick. Human-robot cooperation in future production systems: Analysis of requirements for designing an ergonomic work system. *Procedia Manufacturing*, 3:510–517, 2015.

[18] Alessandro De Luca and Fabrizio Flacco. Integrated control for phri: Collision avoidance, detection, reaction and collaboration. In *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 288–295, 2012.

[19] Andrea Cherubini, Robin Passama, Arnaud Meline, Andre Crosnier, and Philippe Fraisse. Multimodal control for human-robot cooperation. In *2013 IEEE/RSJ International Conference on Intelligent*

Robots and Systems. IEEE, November 2013.

[20] Nils Mandischer, Marius Gürtler, Carlo Weidemann, Elodie Hüsing, Stefan-Octavian Bezrucav, Daniel Gossen, Vincent Brünjes, Mathias Hüsing, and Burkhard Corves. Toward adaptive human–robot collaboration for the inclusion of people with disabilities in manual labor tasks. *Electronics*, 12(5):1118, February 2023.

[21] Hongyi Liu and Lihui Wang. Gesture recognition for human-robot collaboration: A review. *International Journal of Industrial Ergonomics*, 68:355–367, November 2018.

[22] Panagiota Tsarouchi, Athanasios Athanasatos, Sotiris Makris, Xenofon Chatzigeorgiou, and George Chryssolouris. High level robot programming using body and hand gestures. *Procedia CIRP*, 55:1–5, 2016.

[23] Alexander Mörtl, Martin Lawitzky, Ayse Kucukyilmaz, Metin Sezgin, Cagatay Basdogan, and Sandra Hirche. The role of roles: Physical cooperation between humans and robots. *The International Journal of Robotics Research*, 31(13):1656–1674, August 2012.

[24] Jim Mainprice, Emrah Sisbot, Thierry Siméon, and Rachid Alami. Planning safe and legible hand-over motions for human-robot interaction. In *IARP2010*, 01 2010.

[25] SONEHARA Mitsuharu FUJII Masakazu, MURAKAMI Hiroki. Study on application of a human-robot collaborative system using hand-guiding in a production line. 2016.

[26] Jochen Heinzmann and Alexander Zelinsky. Quantitative safety guarantees for physical human-robot interaction. *The International Journal of Robotics Research*, 22(7-8):479–504, July 2003.

[27] Sami Haddadin, Alin Albu-Schaffer, Alessandro De Luca, and Gerd Hirzinger. Collision detection and reaction: A contribution to safe physical human-robot interaction. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, September 2008.

[28] Bhanoday Vemula, Björn Matthias, and Aftab Ahmad. A design metric for safety assessment of industrial robot design suitable for power- and force-limited collaborative operation. *International Journal of Intelligent Robotics and Applications*, 2(2):226–234, April 2018.

[29] A. Cirillo, F. Ficuciello, C. Natale, S. Pirozzi, and L. Villani. A conformable force/tactile skin for physical human–robot interaction. *IEEE Robotics and Automation Letters*, 1(1):41–48, January 2016.

[30] Milad Geravand, Fabrizio Flacco, and Alessandro De Luca. Human-robot physical interaction and collaboration using an industrial robot with a closed control architecture. In *2013 IEEE International Conference on Robotics and Automation*. IEEE, May 2013.

[31] Mazen Ghandour, Hui Liu, Norbert Stoll, and Kerstin Thurow. Human robot interaction for hybrid collision avoidance system for indoor mobile robots. *Advances in Science, Technology and Engineering Systems Journal*, 2(3):650–657, June 2017.

[32] B. Povse, D. Koritnik, R. Kamnik, T. Bajd, and M. Munih. Industrial robot and human operator collision. In *2010 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, October 2010.

[33] B. Povse, S. Haddadin, R. Belder, D. Koritnik, and T. Bajd. A tool for the evaluation of human lower arm injury: approach, experimental validation and application to safe robotics. *Robotica*, 34(11):2499–2515, April 2015.

[34] S. Robla-Gomez, Victor M. Becerra, J. R. Llata, E. Gonzalez-Sarabia, C. Torre-Ferrero, and J. Perez-Oria. Working together: A review on safe human-robot collaboration in industrial environments. *IEEE Access*, 5:26754–26773, 2017.

[35] K. Suita, Y. Yamada, N. Tsuchida, K. Imai, H. Ikeda, and N. Sugimoto. A failure-to-safety "kyozon" system with simple contact detection and stop capabilities for safe human-autonomous robot coexistence. In *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, volume 3, pages 3089–3096 vol.3, 1995.

[36] Hun-Ok Lim and K. Tanie. Collision-tolerant control of human-friendly robot with viscoelastic trunk. *IEEE/ASME Transactions on Mechatronics*, 4(4):417–427, 1999.

[37] John Ulmen and Mark Cutkosky. A robust, low-cost and low-noise artificial skin for human-friendly robots. In *2010 IEEE International Conference on Robotics and Automation*. IEEE, May 2010.

[38] B. Vanderborght, A. Albu-Schaeffer, A. Bicchi, E. Burdet, D.G. Caldwell, R. Carloni, M. Catalano, O. Eiberger, W. Friedl, G. Ganesh,

M. Garabini, M. Grebenstein, G. Grioli, S. Haddadin, H. Hoppner, A. Jafari, M. Laffranchi, D. Lefeber, F. Petit, S. Stramigioli, N. Tsagarakis, M. Van Damme, R. Van Ham, L.C. Visser, and S. Wolf. Variable impedance actuators: A review. *Robotics and Autonomous Systems*, 61(12):1601–1614, December 2013.

[39] B. Gombert, G. Hirzinger, G. Plank, and M. Schedl. Modular concepts for a new generation of light weight robots. In *Proceedings of IECON'94 - 20th Annual Conference of IEEE Industrial Electronics*. IEEE, 1994.

[40] G. Hirzinger. Mechatronics for a new robot generation. *IEEE/ASME Transactions on Mechatronics*, 1(2):149–157, 1996.

[41] G. Hirzinger, J. Butterfass, M. Fischer, M. Grebenstein, M. Hahnle, H. Liu, I. Schaefer, and N. Sporer. A mechatronics approach to the design of light-weight arms and multifingered hands. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*. IEEE, 2000.

[42] G. Hirzinger, A. Albu-Schaffer, M. Hahnle, I. Schaefer, and N. Sporer. On a new generation of torque controlled light-weight robots. In *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164)*. IEEE, 2001.

[43] Rainer Bischoff, Johannes Kurth, Günter Schreiber, Ralf Koeppe, Alin Albu-Schäeffer, Alexander Beyer, Oliver Eiberger, Sami Haddadin, Andreas Stemmer, Gerhard Grunwald, and Gerd Hirzinger. The kuka-dlr lightweight robot arm - a new reference platform for robotics research and manufacturing. volume 2, pages 1–8, 01 2010.

[44] G. Hirzinger, N. Sporer, A. Albu-Schaffer, M. Hahnle, R. Krenn, A. Pascucci, and M. Schedl. DLR's torque-controlled light weight robot III-are we reaching the technological limits now? In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*. IEEE, 2002.

[45] John O'Neill, Jason Lu, Rodney Dockter, and Timothy Kowalewski. Stretchable, flexible, scalable smart skin sensors for robotic position and force estimation. *Sensors*, 18(4):953, March 2018.

[46] Alessandro De Luca and Fabrizio Flacco. Integrated control for pHRI: Collision avoidance, detection, reaction and collaboration. In *2012 4th*

IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob). IEEE, June 2012.

[47] A. de Luca and R. Mattone. Sensorless robot collision detection and hybrid force/motion control. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. IEEE, 2005.

[48] Emanuele Magrini, Fabrizio Flacco, and Alessandro De Luca. Control of generalized contact motion and force in physical human-robot interaction. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2015.

[49] T. Arai, R. Kato, and M. Fujita. Assessment of operator stress induced by robot collaboration in assembly. *CIRP Annals*, 59(1):5–8, 2010.

[50] Dana Kulić and Elizabeth A. Croft. Real-time safety for human–robot interaction. *Robotics and Autonomous Systems*, 54(1):1–12, January 2006.

[51] Koji Ikuta, Hideki Ishii, and Makoto Nokata. Safety evaluation method of design and control for human-care robots. *The International Journal of Robotics Research*, 22(5):281–297, May 2003.

[52] Oussama Khatib. Real-time obstacle avoidance for manipulators and mobile robots. In *Autonomous Robot Vehicles*, pages 396–404. Springer New York, 1986.

[53] T. Tsuji and M. Kaneko. Noncontact impedance control for redundant manipulators. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 29(2):184–193, March 1999.

[54] Dana Kulić and Elizabeth Croft. Pre-collision safety strategies for human-robot interaction. *Autonomous Robots*, 22(2):149–164, October 2006.

[55] Roberto Rossi, Matteo Parigi Polverini, Andrea Maria Zanchettin, and Paolo Rocco. A pre-collision control strategy for human-robot interaction based on dissipated energy in potential inelastic impacts. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, September 2015.

[56] Anis Meguenani, Vincent Padois, and Philippe Bidaud. Control of robots sharing their workspace with humans: An energetic approach

to safety. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, September 2015.

[57] Jeffrey Too Chuan Tan, Feng Duan, Ye Zhang, Kei Watanabe, Ryu Kato, and Tamio Arai. Human-robot collaboration in cellular manufacturing: Design and development. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, October 2009.

[58] F. Flacco, T. Kroger, A. De Luca, and O. Khatib. A depth space approach to human-robot collision avoidance. In *2012 IEEE International Conference on Robotics and Automation*. IEEE, May 2012.

[59] Giovanni Buizza Avanzini, Nicola Maria Ceriani, Andrea Maria Zanchettin, Paolo Rocco, and Luca Bascetta. Safety control of industrial robots based on a distributed distance sensor. *IEEE Transactions on Control Systems Technology*, 22(6):2127–2140, November 2014.

[60] Andrea Maria Zanchettin, Nicola Maria Ceriani, Paolo Rocco, Hao Ding, and Bjorn Matthias. Safety in human-robot collaborative manufacturing environments: Metrics and control. *IEEE Transactions on Automation Science and Engineering*, 13(2):882–893, April 2016.

[61] Riccardo Maderna, Maria Pozzi, Andrea Maria Zanchettin, Paolo Rocco, and Domenico Prattichizzo. Flexible scheduling and tactile communication for human–robot collaboration. *Robotics and Computer-Integrated Manufacturing*, 73:102233, February 2022.

[62] Yuqian Lu, Juvenal Sastre Adrados, Saahil Shivneel Chand, and Lihui Wang. Humans are not machines—anthropocentric human–machine symbiosis for ultra-flexible smart manufacturing. *Engineering*, 7(6):734–737, June 2021.

[63] UNI. IEC 60204-1:2016 - Safety of machinery: Electrical equipment of machines - Part 1: General requirements. Standard, International Organization for Standardization, 2016.

[64] UNI. ISO 13855:2010 - Safety of machinery: Positioning of safeguards with respect to the approach speeds of parts of the human body. Standard, International Organization for Standardization, 2010.

[65] V. Hayward. Fast collision detection scheme by recursive decomposition of a manipulator workspace. In *Proceedings. 1986 IEEE International Conference on Robotics and Automation*. Institute of Electrical

and Electronics Engineers, 1986.

[66] A.P. del Pobil, M.A. Serna, and J. Llovet. A new representation for collision avoidance and detection. In *Proceedings 1992 IEEE International Conference on Robotics and Automation*. IEEE Comput. Soc. Press, 1992.

[67] Lucian Balan and Gary Bone. Real-time 3d collision avoidance method for safe human and robot coexistence. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, October 2006.

[68] Juan Antonio Corrales Ramón, Fernando Torres, and Francisco A. Candelas-Herías. Safe human-robot interaction based on a hierarchy of bounding volumes, 2010-05.

[69] J.A. Corrales, F.A. Candelas, and F. Torres. Safe human–robot interaction based on dynamic sphere-swept line bounding volumes. *Robotics and Computer-Integrated Manufacturing*, 27(1):177–185, February 2011.

[70] Alessandro Albini and Giorgio Cannata. Pressure distribution classification and segmentation of human hands in contact with the robot body. *The International Journal of Robotics Research*, 39(6):668–687, March 2020.

[71] J.T. Feddema and J.L. Novak. Whole arm obstacle avoidance for teleoperated robots. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*. IEEE Comput. Soc. Press, 1994.

[72] J.L. Novak and I.T. Feddema. A capacitance-based proximity sensor for whole arm obstacle avoidance. In *Proceedings 1992 IEEE International Conference on Robotics and Automation*. IEEE Comput. Soc. Press, 1992.

[73] Alwin Hoffmann, Alexander Poeppel, Andreas Schierl, and Wolfgang Reif. Environment-aware proximity detection with capacitive sensors for human-robot-interaction. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, October 2016.

[74] Urzelai Joseba Uribe 1, Juan Pedro. Fuzzy controller for obstacle-avoidance with a non-holonomous mobile robot. *Mathware and Soft*

Computing, 5(2-3):279–289, 1998.

[75] Yong Yu and K. Gupta. Sensor-based roadmaps for motion planning for articulated robots in unknown environments: some experiments with an eye-in-hand system. In *Proceedings 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human and Environment Friendly Robots with High Intelligence and Emotional Quotients (Cat. No.99CH36289)*. IEEE, 1999.

[76] Nicola Maria Ceriani, Giovanni Buizza Avanzini, Andrea Maria Zanchettin, Luca Bascetta, and Paolo Rocco. Optimal placement of spots in distributed proximity sensors for safe human-robot interaction. In *2013 IEEE International Conference on Robotics and Automation*. IEEE, May 2013.

[77] A.-J. Baerveldt. Cooperation between man and robot: interface and safety. In *Proceedings IEEE International Workshop on Robot and Human Communication*. IEEE, 1992.

[78] Kurt Skifstad and Ramesh Jain. Illumination independent change detection for real world image sequences. *Computer Vision, Graphics, and Image Processing*, 46(3):387–399, June 1989.

[79] Stefan Kuhn, Thorsten Gecks, and Dominik Henrich. Velocity control for safe robot guidance based on fused vision and force/torque data. In *2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*. IEEE, September 2006.

[80] Feng Duan, J. Tan, and T. Arai. Using motion capture data to regenerate operator's motions in a simulator in real time. In *2008 IEEE International Conference on Robotics and Biomimetics*. IEEE, February 2009.

[81] Markus Fischer and Dominik Henrich. 3d collision detection for industrial robots and unknown obstacles using multiple depth images. In *Advances in Robotics Research*, pages 111–122. Springer Berlin Heidelberg, 2009.

[82] Markus Fischer and Dominik Henrich. Surveillance of robots using multiple colour or depth cameras with distributed processing. In *2009 Third ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC)*. IEEE, August 2009.

[83] Stephan Puls, Jrgen Graf, and Heinz Wr. Cognitive robotics in indus-

trial environments. In *Human Machine Interaction - Getting Closer*. InTech, January 2012.

[84] Jungong Han, Ling Shao, Dong Xu, and Jamie Shotton. Enhanced computer vision with microsoft kinect sensor: A review. *IEEE Transactions on Cybernetics*, 43(5):1318–1334, October 2013.

[85] Fabrizio Flacco, Torsten Kroeger, Alessandro De Luca, and Oussama Khatib. A depth space approach for evaluating distance to objects. *Journal of Intelligent and Robotic Systems*, 80(S1):7–22, October 2014.

[86] Herbert Kaplan. *Practical applications of infrared thermal sensing and imaging equipment*. Tutorial Texts. SPIE Press, Bellingham, WA, 3 edition, June 2007.

[87] Nimmi Arora, Diana Martins, Danielle Ruggerio, Eleni Tousimis, Alexander J. Swistel, Michael P. Osborne, and Rache M. Simmons. Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. *The American Journal of Surgery*, 196(4):523–526, October 2008.

[88] Baopu Li, Haoyang Jin, Qi Zhang, Wei Xia, and Huiyun Li. Indoor human detection using rgb-d images. In *2016 IEEE International Conference on Information and Automation (ICIA)*, pages 1354–1360, 2016.

[89] N. and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893 vol. 1, 2005.

[90] Ming Liang and Xiaolin Hu. Recurrent convolutional neural network for object recognition. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3367–3375, 2015.

[91] Pierre Sermanet, David Eigen, Xiang Zhang, Michael Mathieu, Rob Fergus, and Yann LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks, 2013.

[92] Kaiyang Zhou, Adeline Paiement, and Majid Mirmehdi. Detecting humans in RGB-d data with CNNs. In *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*. IEEE, May 2017.

[93] Anis Koubaa, editor. *Robot operating system (ROS)*. Studies in computational intelligence. Springer International Publishing, Cham, Switzerland, 1 edition, February 2016.

[94] Antonio Gulli and Sujit Pal. *Deep learning with keras*. Packt Publishing, Birmingham, England, April 2023.

[95] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 580–587, 2014.

[96] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788, 2016.

[97] MathWorks, MATLAB. [Online] = https://www.mathworks.com/, note = Accessed: 2023-10-11 .

[98] Wilfried Elmenreich. An introduction to sensor fusion. *Vienna University of Technology, Austria*, 11 2002.

[99] Bo Yang, Zhong liang Jing, and Hai tao Zhao. Review of pixel-level image fusion. *Journal of Shanghai Jiaotong University (Science)*, 15(1):6–12, February 2010.

[100] Thomas Luhmann, Johannes Piechel, and Thorsten Roelfs. Geometric calibration of thermographic cameras. In *Thermal Infrared Remote Sensing*, pages 27–42. Springer Netherlands, 2013.

[101] J. Rangel and S. Soldan. 3d thermal imaging: Fusion of thermography and depth cameras. In *Proceedings of the 2014 International Conference on Quantitative InfraRed Thermography*. QIRT Council, 2014.

[102] James Bowman and Patrick Mihelich. Camera calibration tool, howpublished = http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm, note = Accessed: 2010-09-30 .

[103] Joseph Redmon and Ali Farhadi. Yolo9000: Better, faster, stronger. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6517–6525, 2017.

[104] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *CoRR*, abs/1804.02767, 2018.

[105] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[106] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: Common objects in context. In *Computer Vision – ECCV 2014*, pages 740–755. Springer International Publishing, 2014.

[107] M. Costanzo, G. De Maria, G. Lettera, C. Natale, and D. Perrone. A multimodal perception system for detection of human operators in robotic work cells. In *2019 IEEE International Conference on Systems, Man and Cybernetics*, pages 692–699, Bari, I, October 2019. IEEE.

[108] N. Blodow. Realtime urdf filter, 2012.

[109] Su Il Choi and Byung Kook Kim. Obstacle avoidance control for redundant manipulators using collidability measure. In *Proceedings 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human and Environment Friendly Robots with High Intelligence and Emotional Quotients (Cat. No.99CH36289)*. IEEE, 1999.

[110] Paul Bosscher and Daniel Hedman. Real-time collision avoidance algorithm for robotic manipulators. In *2009 IEEE International Conference on Technologies for Practical Robot Applications*. IEEE, nov 2009.

[111] Greg Welch, Gary Bishop, et al. An introduction to the kalman filter. *University of North Carolina at Chapel Hill, Chapel Hill, NC, United States*, 1995.

[112] Timothy J. Ross. *Fuzzy Logic with Engineering Applications*. John Wiley & Sons, Ltd, jan 2010.

[113] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, jun 1965.

[114] MathWorks. Defuzzification Methods.

[115] Ion Iancu. A mamdani type fuzzy logic controller. In *Fuzzy logic-*

controls, concepts, theories and applications. IntechOpen, 2012.

[116] James J Buckley. Sugeno type controllers are universal controllers. *Fuzzy sets and systems*, 53(3):299–303, 1993.

[117] The Robotics Lab of University of Campania Luigi Vanvitelli, howpublished = https://www.ingegneria.unicampania.it/roboticslab, note = Accessed: 2023-11-22 .

[118] Leonardo company, howpublished = https://www.leonardo.com/en/home, note = Accessed: 2023-11-22 .

[119] Yaskawa: MotoPlus SDK, howpublished = https://www.motoman.com/en-us/products/software/development/motoplus-sdk, note = Accessed: 2023-11-22 .

[120] Angelo Campomaggiore., Marco Costanzo., Gaetano Lettera., and Ciro Natale. A fuzzy inference approach to control robot speed in human-robot shared workspaces. In *Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics - Volume 2: ICINCO,*, pages 78–87, Prague, CZ, July 2019. INSTICC, SciTePress.

[121] V-Rep by Coppelia, howpublished = http://www.coppeliarobotics.com/, note = Accessed: 2023-11-22 .

[122] SICK: Sensor Itelligence, howpublished = https://www.sick.com/it/it/, note = Accessed: 2023-11-22 .

[123] LABOR European project, howpublished = https://www.labor-project.eu/, note = Accessed: 2023-11-22 .

[124] UNI. ISO 12100 - Risk assessment and risk reduction. Standard, International Organization for Standardization, 2010.

[125] UNI. ISO 13849-1 - Safety of machinery. Standard, International Organization for Standardization, 2023.

[126] Gareth Oregan Hoskins, Jared Padayachee, and Glen Bright. Human-robot interaction: The safety challenge (an inegrated frame work for human safety). In *2019 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa (SAUPEC/RobMech/PRASA)*. IEEE, January 2019.

[127] Kangdon Lee, Jaeho Shin, and Jae-Yong Lim. Critical hazard factors in the risk assessments of industrial robots: Causal analysis and case studies. *Safety and Health at Work*, 12(4):496–504, December 2021.

[128] ANSI. ANSI B11.TR3-2000: Risk Assessment and Risk Reduction - A Guide to Estimate, Evaluate and Reduce Risks Associated with Machine Tools. Standard, American National Standards Institute, 2000.

[129] ANSI. ANSI/RIA R15.06-1999 (R2009): Industrial Robots And Robot Systems - Safety Requirements. Standard, American National Standards Institute, 1999.

[130] RIA. RIA TR R15.306-2016: Task-Based Risk Assessment Methodology. Standard, Robotic Industries Association, 2016.

[131] ANSI. ANSI/RIA R15.08-1-2020: Industrial Mobile Robots - Safety Requirements - Part 1: Requirements For The Industrial Mobile Robot. Standard, American National Standards Institute, 2020.

[132] Bjoern Matthias, Susanne Oberer-Treitz, Harald Staab, Erich Schuller, and Steffen Peldschus. Injury risk quantification for industrial robots in collaborative operation with humans. In *ISR 2010 (41st International Symposium on Robotics) and ROBOTIK 2010 (6th German Conference on Robotics)*, pages 1–6, 2010.

[133] J. Carlson, R.R. Murphy, and A. Nelson. Follow-up analysis of mobile robot failures. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004.* IEEE, 2004.

[134] Fabrizio Caccavale and Luigi Villani, editors. *Fault diagnosis and fault tolerance for mechatronic systems: Recent advances.* Springer Tracts in Advanced Robotics. Springer, Berlin, Germany, 2003 edition, October 2002.