Wearable MIMUs for the Identification of Upper Limbs Motion in an Industrial Context of Human-Robot Interaction

Mattia Antonelli[1][a], Elisa Digo[1][b], Stefano Pastorelli[1][c] and Laura Gastaldi[2][d]
[1]*Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin, Italy*
[2]*Department of Mathematical Sciences "G. L. Lagrange", Politecnico di Torino, Turin, Italy*

Keywords: MIMU, Upper Limb, Motion Prediction, Industry 4.0, Linear Discriminant Analysis, Movement Classification.

Abstract: The automation of human gestures is gaining increasing importance in manufacturing. Indeed, robots support operators by simplifying their tasks in a shared workspace. However, human-robot collaboration can be improved by identifying human actions and then developing adaptive control algorithms for the robot. Accordingly, the aim of this study was to classify industrial tasks based on accelerations signals of human upper limbs. Two magnetic inertial measurement units (MIMUs) on the upper limb of ten healthy young subjects acquired pick and place gestures at three different heights. Peaks were detected from MIMUs accelerations and were adopted to classify gestures through a Linear Discriminant Analysis. The method was applied firstly including two MIMUs and then one at a time. Results demonstrated that the placement of at least one MIMU on the upper arm or forearm is suitable to achieve good recognition performances. Overall, features extracted from MIMUs signals can be used to define and train a prediction algorithm reliable for the context of collaborative robotics.

1 INTRODUCTION

Technological developments of Industry 4.0 are increasingly oriented to the automation of human gestures, supporting operators with robotic systems that can perform or simplify their task in the production process. In this innovative industrial context, collaborative robotics can be considered safe if the human and the robot can coexist in the same workspace. Indeed, the ability of the robot to detect obstacles, even dynamic ones offered by human movements, is crucial. Hence, the machine has to integrate with sensors recording human motion and systems processing these data, to avoid collisions and accidents (Safeea and Neto, 2019).

Once the safety is guaranteed, the collaboration between human and robot could be further improved by identifying human actions, timings and paths and consequently developing adaptive control algorithms for the robot (Lasota, Fong and Shah, 2017; Ajoudani *et al.*, 2018). In this perspective, the prediction of human activities plays a fundamental role in human-machine interaction. Indeed, some literature works have already adopted human motion prediction to improve the performance of robotic systems, by reducing times of tasks execution while maintaining standards of safety (Pellegrinelli *et al.*, 2016; Weitschat *et al.*, 2018).

The operation of human motion prediction requires a reliable tracking of the human trajectory and movement in real-time. The capture of human movement could be carefully performed by using vision devices such as stereophotogrammetric systems and RGB-D cameras (Mainprice and Berenson, 2013; Perez-D'Arpino and Shah, 2015; Pereira and Althoff, 2016; Wang *et al.*, 2017; Scimmi *et al.*, 2019; Melchiorre *et al.*, 2020). However, despite their precision, vision systems have some disadvantages such as encumbrance, high costs, problems of occlusion, and long set-up and

[a] https://orcid.org/0000-0002-4549-1822
[b] https://orcid.org/0000-0002-5760-9541
[c] https://orcid.org/0000-0001-7808-8776
[d] https://orcid.org/0000-0003-3921-3022

Antonelli, M., Digo, E., Pastorelli, S. and Gastaldi, L.
Wearable MIMUs for the Identification of Upper Limbs Motion in an Industrial Context of Human-Robot Interaction.
DOI: 10.5220/0010548304030409
In *Proceedings of the 18th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2021)*, pages 403-409
ISBN: 978-989-758-522-7

calibration times. All these aspects make vision technologies not suitable for an industrial context to assess human-robot interaction.

The recent development of a new generation of magnetic inertial measurement units (MIMUs) based on micro-electro-mechanical systems technology has given a new impetus to motion tracking research (Lopez-Nava and Angelica, 2016; Filippeschi *et al.*, 2017). Indeed, wearable inertial sensors have become a cornerstone in real-time capturing human motion in different contexts such as the rehabilitation field (Balbinot, de Freitas and Côrrea, 2015), sports activities (Hsu *et al.*, 2018) and industrial environment (Safeea and Neto, 2019). Even if they are not excellent in terms of accuracy and precision, MIMUs are cheap, portable, easy to wear, and non-invasive. Moreover, they overcome the typical limitations of optical systems because they do not suffer from occlusion problems, they have a theoretically unlimited working range, and they reduce calibration and computational times. For these reasons, the adoption of wearable MIMUs in an industrial context of human-robot interaction could be deeper investigated.

Two previous studies have been conducted with the intent of improving the human-robot collaboration by collecting and analyzing typical industrial gestures of pick and place at different heights. The upper limbs motion of ten healthy young subjects has been acquired with both a stereophotogrammetric and an inertial system. The first work has promoted the creation of a database collecting spatial and inertial variables derived from a sensor fusion procedure (Digo, Antonelli, Pastorelli, *et al.*, 2020). Since results have highlighted that the obtained database was congruent, complementary, and suitable for features identifications, the study has been amplified. Indeed, the second work has developed a recognition algorithm enabling the selection of the most representative features of upper limbs movement during pick and place gestures. Results have revealed that the recognition algorithm provided a good balance between precision and recall and that all tested features can be selected for the pick and place detection (Digo, Antonelli, Cornagliotto, *et al.*, 2020).

However, these two studies have involved the use of an optical marker-based system, which is unsuitable for an industrial context of human-robot interaction. Accordingly, the present work has concentrated only on features collected tracking the human upper limbs movement with MIMUs. Ten healthy young subjects have executed pick and place gestures at three different heights. Two inertial sensors on the upper arm and forearm of participants have been considered for data analysis. In detail, the aim was to adopt MIMUs to guarantee the same classification performances obtained with markers trajectories optimizing the experimental set-up and reducing the computational times.

2 MATERIALS & METHODS

2.1 Participants

Ten healthy young subjects (6 males, 4 females) with no musculoskeletal or neurological diseases were recruited for the experiment. All involved participants were right-handed. Mean and standard deviation values of subjects' anthropometric data were estimated (Table 1). The study was approved by the Local Institutional Review Board. All procedures were conformed to the Helsinki Declaration. Participants gave their written informed consent before the experiment.

Table 1: Anthropometric data of participants.

	Mean (St. Dev)
Age (years)	24.7 (2.1)
BMI (kg/m^2)	22.3 (3.0)
Upper arm length (cm)	27.8 (3.2)
Forearm length (cm)	27.9 (1.5)
Trunk length (cm)	49.1 (5.2)
Acromions distance (cm)	35.9 (3.6)

2.2 Instruments

The instrumentation adopted for the present study was composed of an inertial measurement system. In detail, four MTx MIMUs (Xsens, The Netherlands) were used for the test. Each of them contained a tri-axial accelerometer (range ± 5 G), a tri-axial gyroscope (range ± 1200 dps) and a tri-axial magnetometer (range $\pm 75\ \mu$T). Three sensors (Figure 1A) were positioned on the participants' upper body: right forearm (RFA), right upper arm (RUA) and thorax (THX). All MIMUs on participants were fixed by aligning their local reference systems with the relative anatomical reference systems of the segments on which they were fixed. Another MIMU (TAB) was fixed on a table with the horizontal x-axis pointing towards the participants, the horizontal y-axis directed towards the right side of subjects, and the vertical z-axis pointing upward (Figure 1B). The four sensors were mutually linked into a chain through cables and the TAB-MIMU was also connected to the control unit called Xbus Master. The communication between MIMUs and a PC was guaranteed via Bluetooth. Data were acquired through the Xsens proprietary software MT Manager with a sampling frequency of 50 Hz.

2.3 Protocol

The test was conducted in a laboratory. The setting was composed of a table on which the silhouettes of right and left human hands were drawn, with thumbs 32 cm apart. In addition, a cross was marked between the hands' silhouettes. Subsequently, three coloured boxes of the same size were placed on the right side of the table at different heights: a white box on the table, a black one at a height of 18 cm from the table, and a red one at a height of 28 cm from the table (Figure 1B).

Subjects were first asked to sit at the table. Then, a calibration procedure was performed asking participants to stand still for 10 s in a seated neutral position with hands on silhouettes. Finally, subjects performed pick and place tasks composed of 7 operations: 1) start with hands in neutral position; 2) pick the box according to the colour specified by the experimenter; 3) place the box correspondingly to the cross marked on the table; 4) return with hands in neutral position; 5) pick the same box; 6) replace the box in its initial position; 7) return with hands in neutral position. During these operations, subjects were asked not to move the trunk as much as possible, in order to focus the analysis only on the right upper limb.

A metronome set to 45 bpm was adopted to ensure that each of the seven operations was executed by all subjects at the same pace. Each participant performed 15 consecutive gestures of pick and place, 5 for every box. The sequence of boxes to be picked and placed was randomized and voice-scanned by the experimenters during the test.

2.4 Signal Processing and Data Analysis

Signal processing and data analysis were conducted with Matlab® (MathWorks, USA) and SPSS® (IBM, USA).

The robotic multibody approach was applied, by modelling the upper body of participants in rigid links connected by joints (Gastaldi, Lisco and Pastorelli, 2015). In detail, three body segments were identified: right forearm, right upper arm and trunk. All signals obtained from MIMUs during the registered movements were filtered with a second-order Butterworth low-pass filter with a cut-off frequency of 2 Hz. Subsequently, accelerations of the MIMU on the thorax were used to verify that the movement principally involved only the right upper limb and not the trunk.

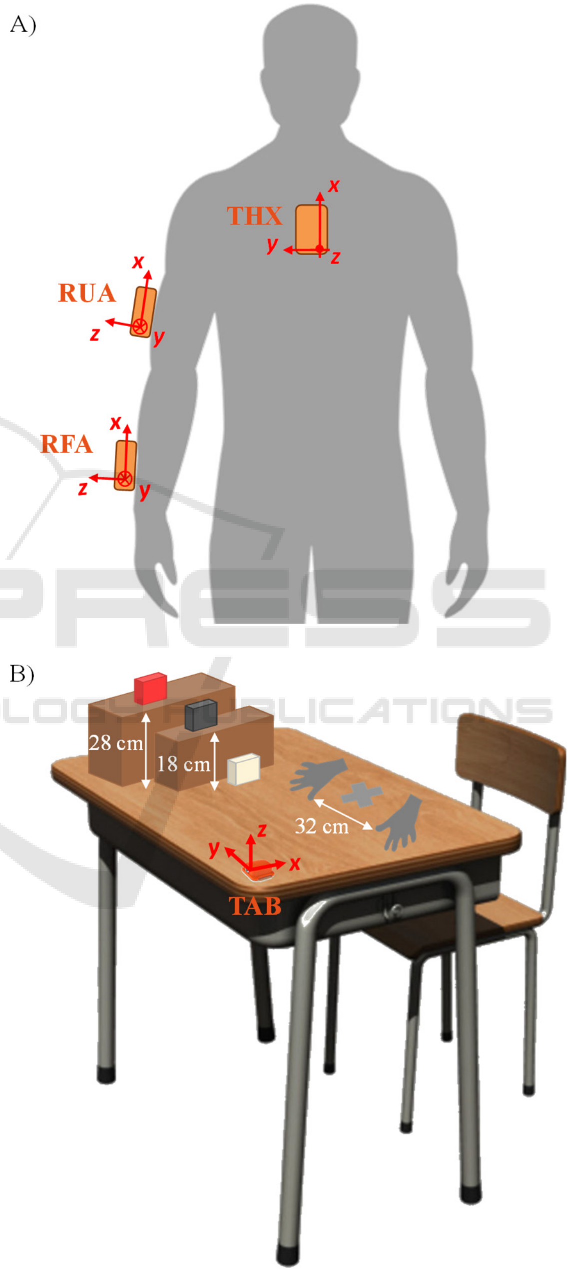

Figure 1: A) Positioning of three MIMUs on participants' upper body and their local reference systems; B) Experimental setting with table, boxes, hands silhouettes, cross and TAB-MIMU.

As a result, only accelerations along all axes of MIMUs on the forearm (x-RFA, y-RFA, z-RFA) and upper arm (x-RUA, y-RUA, z-RUA) were considered for all subjects.

A method to identify all pick and place gestures from MIMUs accelerations was implemented. In each of the six signals of each participant, a pick and place gesture of a box was recognized as a double peak. Accordingly, for each participant, 15 pairs of peaks were identified as corresponding to 15 performed gestures. In Figure 2, as an example, the acceleration signal along the x-axis for the RUA MIMU is reported. The amplitude of each pair of consecutive peaks was averaged calculating p_i, with $i = 1 \div 15$ (Figure 2). Values of p_i estimated for all signals and all participants were collected in a single matrix of 150 rows (corresponding to 15 pick and place gestures performed by 10 subjects) and 6 columns (corresponding to MIMUs accelerations).

Starting from this matrix containing peaks values, a Linear Discriminant Analysis (LDA) was implemented and repeated considering (a) the whole matrix, (b) only RFA-MIMU accelerations and (c) only RUA-MIMU accelerations. Observations were divided into two groups, one for the training (TR) and one for the test (TT) of the algorithm. Three splits were considered: (i) 100% TR – 100% TT, (ii) 66% TR – 33% TT, (iii) 33% TR – 66% TT. In all cases, the two groups were defined randomly picking the same balanced number of observations from the three gestures categories. Results of LDA were processed into scatterplots, confusion matrices and F1-scores to evaluate the classification performances. Since the three splits produced similar outcomes, only the results of the latter case (iii) are presented.

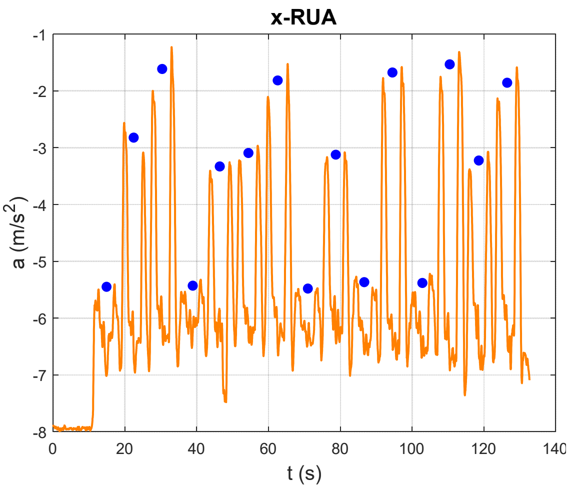

Figure 2: Identification of pick and place gestures from MIMUs signals. Example of subject n°6: x-RUA acceleration (orange) and averaged peaks values (blue dot).

3 RESULTS

In each of the three analyses (all accelerations, only RFA-MIMU, only RUA-MIMU), LDA identified two linear functions of data for the classification of gestures. Considering eigenvalues of both functions, the first one expressed alone at least 98% of data variability in all cases (99.5% for all accelerations, 98.6% for RFA, 99.5% for RUA). Thereby, the second function covered the remaining data variability (0.5% for all accelerations, 1.4% for RFA, 0.5% for RUA). Accordingly, coefficients (Table 2) and values of correlations (Table 3) of the first linear function were reported and discussed for all three cases.

Scatterplots represented in Figure 3 define linear boundaries among classes regions for the three analyses. Figure 4 depicts confusion matrices obtained in all three cases from the classification of pick and place gestures belonging to the test group. Accordingly, Table 4 shows F1-scores (%) estimated from the confusion matrices combining the precision and the recall.

Table 2: Coefficients of the linear function 1 identified from data in all three analyses (all accelerations, only RFA-MIMU, only RUA-MIMU).

	Coefficients		
	All accelerations	RFA MIMU	RUA MIMU
x-RFA	1.802	-2.117	-
y-RFA	-0.442	0.546	-
z-RFA	-1.993	2.936	-
x-RUA	1.734	-	2.173
y-RUA	0.249	-	0.687
z-RUA	-0.447	-	-0.781
const	8.290	-3.409	7.159

Table 3: Values of correlations for each variable with function 1 in all three analyses (all accelerations, only RFA-MIMU, only RUA-MIMU).

	Correlations		
	All accelerations	RFA MIMU	RUA MIMU
x-RFA	0.546	-0.871	-
y-RFA	-0.009	0.025	-
z-RFA	-0.292	0.326	-
x-RUA	0.522	-	0.848
y-RUA	-0.129	-	-0.182
z-RUA	-0.163	-	-0.254

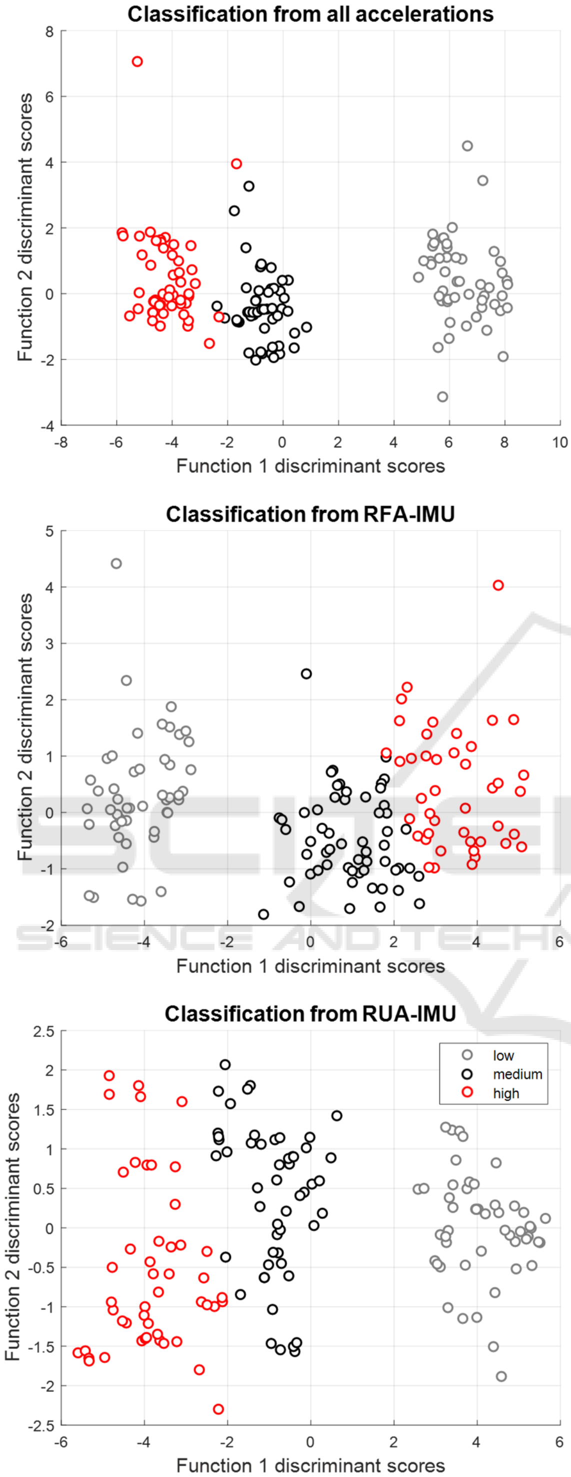

Figure 3: Scatterplots obtained from discriminant scores of functions 1 and 2 for the three cases. Pick and place gestures performed by all subjects are classified as low (grey), medium (black) or high (red) ones.

All accelerations		Predicted Low	Predicted Med	Predicted High	Sum
Actual	Low	33	0	0	33
	Med	0	32	1	33
	High	0	0	34	34
Sum		33	32	35	100

RFA IMU		Predicted Low	Predicted Med	Predicted High	Sum
Actual	Low	33	0	0	33
	Med	0	30	3	33
	High	0	9	25	34
Sum		33	39	28	100

RUA IMU		Predicted Low	Predicted Med	Predicted High	Sum
Actual	Low	33	0	0	33
	Med	0	28	5	33
	High	0	7	27	34
Sum		33	35	32	100

Figure 4: Confusion matrices obtained from the classification procedure in all three analyses.

Table 4: F1-scores (%) estimated for the three gestures (low, medium, and high) of all analyses.

Analyses	F1-scores (%)		
	Low	Medium	High
All accelerations	100	98.5	98.6
RFA-MIMU	100	83.3	80.6
RUA-MIMU	100	82.4	81.8

4 DISCUSSIONS

The aim of the present work was to classify industrial tasks based on MIMUs signals of human upper limbs, to improve the human-robot interaction in a cooperative environment. In detail, pick and place gestures at three different heights were executed by ten healthy young subjects and were recorded through two inertial sensors on the upper arm and forearm. Accelerations peaks were detected for both RFA and RUA MIMUs and were adopted to classify pick and place gestures by means of LDA. Hence, the classification method was applied three times: (i) on all six accelerations, (ii) only on RFA-MIMU accelerations, (iii) only on RUA-MIMU accelerations.

All three analyses provided a linear function expressing almost all the data variability. Considering its coefficients (Table 2), the highest absolute values are referred to x and z accelerations for all cases. This aspect could be caused by the boxes positioning during the experiment. Starting from these coefficients, the correlation values between the accelerations and the first discriminant function were considered for each analysis (Table 3). In all cases, the most relevant variables in the classification process are peaks of x-RFA and x-RUA signals, testifying that the movement was principally developed along their x-axis.

Considering only the RFA-MIMU, peaks of y-acceleration could be excluded from the classification process, due to its lowest correlation. In this way, the computational time could be reduced in the perspective of an almost real-time application.

According to classification results for the three cases, observations were distributed in the plane obtained from discriminant scores of functions (Figure 3). The three classes occupy spatially well-defined regions. Moreover, it is easy to notice that the 'low' region is better separated from the other two due to the greater distance of the low box placement from the medium and the high ones. This aspect leads to a few misclassifications between medium and high gestures of pick and place. Indeed, observing the first column of all confusion matrices (Figure 4), the classification of low gestures is always correct. On the contrary, the second and third columns highlight some wrong identifications of medium and high gestures.

Considering the confusion matrix including all accelerations (Figure 4), the precision is equal to 99%. Taking into account only one sensor, the precision of the classification drops to 88%, both for RFA-MIMU and RUA-MIMU. F1-scores calculated for each case starting from the relative confusion matrix (Table 4) are greater than 80%. It means that the algorithm based on these signals provided a very good balance between precision and recall for all three movements. Since the F1-scores concerning all accelerations are so high, the usage of signals recorded by MIMUs placed on the upper arm and forearm is suitable to identify industrial gestures of pick and place. The F1-scores obtained using signals provided by only one MIMU can be considered good for both adopted sensors. For this reason, the usage of only one of the two mentioned MIMUs guarantees a high classification accuracy, but also it allows to lighten the set-up. This choice can lead to various advantages: the reduction of the encumbrance, the rise of subject comfort in movements, the decrease of subject preparation time, the reduction of the number of data to elaborate and the increase in the algorithm computational speed. These results could be exploited in human-robot collaborative tasks, in which robots cooperate with operators by recognizing their gestures.

5 CONCLUSIONS

In the field of collaborative robotics, detection and identification of gestures play a fundamental role in an environment where humans and robots coexist and perform tasks together. Over the year different instrumentations have been chosen to track human movements and to develop prediction operations reliable in human-robot interaction.

This study aimed to overcome the shortcoming encountered with the use of motion capture tools unsuited to the industrial world. Starting from signals acquired by wearable devices easy to adopt in the industrial field, the work was intended to assess the performance of LDA classification of typical industrial pick and please gestures.

The conducted evaluation showed excellent results in terms of classification precision. Indeed, a few gestures misclassifications were committed likely because of the proximity of boxes involved in the movements. Thus, the use of only MIMUs for tracking human movement can be considered suitable for collaborative prediction procedures. In detail, the placement of at least one inertial unit on the upper arm or forearm is adequate to achieve good recognition results.

Future plans are first to validate the obtained results by applying LDA to data captured with a stereophotogrammetric system. Moreover, other classification methods such as Convolutional Neural

Networks could be implemented to verify the reproducibility of the results. Other acceleration features in addition to peaks, such as punctual values of the jerk, means or periodicities, could be explored. Then, also angular velocities and orientations could be taken into account for the procedure of gesture recognition. Starting from the features extracted from MIMUs signals, a prediction algorithm of human motion can be defined and trained for an industrial context of human-robot collaboration. The prediction operation can contribute to defining a work environment with the robot adapting to the human.

REFERENCES

Ajoudani, A. *et al.* (2018) 'Progress and prospects of the human–robot collaboration', *Autonomous Robots*. Springer US, 42(5), pp. 957–975. doi: 10.1007/s10514-017-9677-2.

Balbinot, A., de Freitas, J. C. R. and Côrrea, D. S. (2015) 'Use of inertial sensors as devices for upper limb motor monitoring exercises for motor rehabilitation', *Health and Technology*, 5(2), pp. 91–102. doi: 10.1007/s12553-015-0110-6.

Digo, E., Antonelli, M., Cornagliotto, V., *et al.* (2020) 'Collection and Analysis of Human Upper Limbs Motion Features for Collaborative Robotic Applications', *Robotics*, 9(2), p. 33. doi: 10.3390/robotics9020033.

Digo, E., Antonelli, M., Pastorelli, S., *et al.* (2020) 'Upper limbs motion tracking for collaborative robotic applications', in *International Conference on Human Interaction & Emerging Technologies*, pp. 391–397. doi: 10.1007/978-3-030-55307-4_59.

Filippeschi, A. *et al.* (2017) 'Survey of motion tracking methods based on inertial sensors: A focus on upper limb human motion', *Sensors (Switzerland)*, 17(6), pp. 1–40. doi: 10.3390/s17061257.

Gastaldi, L., Lisco, G. and Pastorelli, S. (2015) 'Evaluation of functional methods for human movement modelling', *Acta of Bioengineering and Biomechanics*, 17(4), pp. 31–38. doi: 10.5277/ABB-00151-2014-03.

Hsu, Y. L. *et al.* (2018) 'Human Daily and Sport Activity Recognition Using a Wearable Inertial Sensor Network', *IEEE Access*. IEEE, 6, pp. 31715–31728. doi: 10.1109/ACCESS.2018.2839766.

Lasota, P. A., Fong, T. and Shah, J. A. (2017) 'A Survey of Methods for Safe Human-Robot Interaction', *Foundations and Trends in Robotics*, 5(3), pp. 261–349. doi: 10.1561/2300000052.

Lopez-Nava, I. H. and Angelica, M. M. (2016) 'Wearable Inertial Sensors for Human Motion Analysis: A review', *IEEE Sensors Journal*, 16(22), pp. 7821–7834. doi: 10.1109/JSEN.2016.2609392.

Mainprice, J. and Berenson, D. (2013) 'Human-robot collaborative manipulation planning using early prediction of human motion', in *IEEE International Conference on Intelligent Robots and Systems*. IEEE, pp. 299–306. doi: 10.1109/IROS.2013.6696368.

Melchiorre, M. *et al.* (2020) 'Vision-based control architecture for human–robot hand-over applications', *Asian Journal of Control*, 23(1), pp. 105–117. doi: 10.1002/asjc.2480.

Pellegrinelli, S. *et al.* (2016) 'A probabilistic approach to workspace sharing for human–robot cooperation in assembly tasks', *CIRP Annals - Manufacturing Technology*. CIRP, 65(1), pp. 57–60. doi: 10.1016/j.cirp.2016.04.035.

Pereira, A. and Althoff, M. (2016) 'Overapproximative arm occupancy prediction for human-robot co-existence built from archetypal movements', in *IEEE International Conference on Intelligent Robots and Systems*. IEEE, pp. 1394–1401. doi: 10.1109/IROS.2016.7759228.

Perez-D'Arpino, C. and Shah, J. A. (2015) 'Fast target prediction of human reaching motion for cooperative human-robot manipulation tasks using time series classification', *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 6175–6182. doi: 10.1109/ICRA.2015.7140066.

Safeea, M. and Neto, P. (2019) 'Minimum distance calculation using laser scanner and IMUs for safe human-robot interaction', *Robotics and Computer-Integrated Manufacturing*. Elsevier Ltd, 58, pp. 33–42. doi: 10.1016/j.rcim.2019.01.008.

Scimmi, L. S. *et al.* (2019) 'Experimental Real-Time Setup for Vision Driven Hand-Over with a Collaborative Robot', in *IEEE International Conference on Control, Automation and Diagnosis (ICCAD)*. IEEE, pp. 2–6.

Wang, Y. *et al.* (2017) 'Collision-free trajectory planning in human-robot interaction through hand movement prediction from vision', in *IEEE-RAS International Conference on Humanoid Robots*, pp. 305–310. doi: 10.1109/HUMANOIDS.2017.8246890.

Weitschat, R. *et al.* (2018) 'Safe and efficient human-robot collaboration part I: Estimation of human arm motions', in *Proceedings - IEEE International Conference on Robotics and Automation*. IEEE, pp. 1993–1999. doi: 10.1109/ICRA.2018.8461190.