Upper limbs motion tracking for collaborative robotic applications

Elisa Digo[1], Mattia Antonelli[1], Stefano Pastorelli[1], Laura Gastaldi[1]

[1] Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Turin, Italy
{elisa.digo, stefano.pastorelli, laura.gastaldi}@polito.it
mattia.antonelli9@gmail.com

Abstract. In the perspective of Industry 4.0, the contemporary presence of workers and robots in the same workspace requires the development of human motion prediction algorithms for a safe and efficient interaction. In this context, the purpose of the present study was to perform an operation of sensor fusion, by creating a collection of spatial and inertial variables of human upper limbs kinematics of typical industrial movements. Spatial and inertial data of ten healthy young subjects performing three pick and place gestures at different heights were measured with a stereophotogrammetric system and Inertial Measurement Units, respectively. Elbow and shoulder angles estimated from both instruments according to a multibody approach showed very similar trends. Moreover, two variables of the database were identified as distinctive features able to differentiate among the three gestures of pick and place.

Keywords: Sensor fusion · IMU · Prediction · Industry 4.0 · Pick and place

1 Introduction

Technological developments of Industry 4.0 are strongly leading to automate workers' gestures with the support of robots. Consequently, the concept of interaction between human and robot is gaining an increasingly central role in manufacturing. The team of human and robot cooperating in the same environment with a common plan guarantees, security ensured, a more successful and efficient execution of tasks [1–3]. However, this collaboration cannot be considered optimal unless it occurs on the basis of a prediction operation. In fact, by predicting human activity, the robot can identify actions, paths and timing that will result in a safer and more efficient interaction [4, 5]. Several literature works based on human motion tracking have already addressed the concept of human activity prediction. Some studies have structured their prediction methods on data collected with different motion capture instruments such as stereophotogrammetric systems [6, 7] and RGB-D cameras [8, 9]. Other works used existing databases of human motions to define and train their prediction algorithms [10, 11]. Most of the adopted databases are composed of 3D joints coordinates acquired with optical and stereophotogrammetric systems during movements of the total body [12–15]. Other databases contain both spatial and inertial data, but they do not track upper limbs in a complete way [16] or they do not analyse typical manufacturing gestures [17, 18]. Consequently, the aim of the present work was to perform an operation of sensor fusion, by creating a collection of spatial and inertial variables of typical industrial gestures, like pick and place, performed by ten healthy young subjects.

2 Materials and Methods

For the experimental test of this study, ten subjects performed a sequence of pick and place of three boxes positioned at different heights. Motion tracking of participants' upper body was made with both a stereophotogrammetric system and Inertial Measurement Units (IMUs).

2.1 Participants

Ten healthy young subjects (6 males, 4 females) with no musculoskeletal or neurological disease gave their written informed consent to participate in the study. Subjects anthropometric data were estimated (mean ± standard deviation): age 24.7 ± 2.1 years; BMI 22.3 ± 3.0 kg/m^2; upper arm length 27.8 ± 3.2 cm; arm length 27.9 ± 1.5 cm; trunk length 49.1 ± 5.2 cm; acromions distance 35.9 ± 3.6 cm.

2.2 Instrumentation

The instrumentation involved an inertial system and a stereophotogrammetric system.

Inertial system. A chain of 7 MTx IMUs (Xsens, The Netherlands) was connected to a PC via Bluetooth. Data were acquired through the Xsens proprietary software MT Manager at 50 Hz. A first IMU (TAB) was fixed on a table (Fig. 1.A), the other six IMUs were fixed on participants' upper body (Fig. 1.B): right forearm (RFA); right upper arm (RUA); shoulders (RSH, LSH); sternum (THX) and pelvis (PLV).

Stereophotogrammetric system. It was composed of a self-contained and pre-calibrated V120:Trio bar (OptiTrack, USA) and 17 markers (14 mm of diameter). Data acquisition was made through the software Motive at 120 Hz. A global coordinate reference system was constructed from markers A, B and C on the same table (Fig. 1.A). Fourteen markers were positioned on participants' upper body (Fig. 1.B): styloid processes (WMR, WLR, WML, WLL); elbow condyles (EMR, ELR, EML, ELL); acromions (ACR, ACL); between suprasternal notches (IJ); spinal process of the 8th thoracic vertebra (T8); on inertial sensors RFA (SFA) and RUA (SUA).

2.3 Protocol

Subjects were asked to sit at a table on which the silhouettes of right and left hands were drawn, with thumbs 12 cm apart. For the calibration and temporal synchronization of inertial and stereophotogrammetric systems, participants were asked to hit the table with their right wrist and to stand still for 10 s with hands on silhouettes (neutral position). Subsequently, participants performed pick and place tasks composed of 7 operations: 1) stay in neutral position; 2) pick a box; 3) place the box between hands silhouettes; 4) return to neutral position; 5) pick the same box again; 6) place the box on initial position; 7) return to neutral position. The task was repeated with three boxes of the same size placed on the table at different heights (Fig. 1A): on the table (white), at a height of 18 cm (black) and at a height of 28 cm (red). Each subject performed 15 gestures, 5 for each box, in a random order.

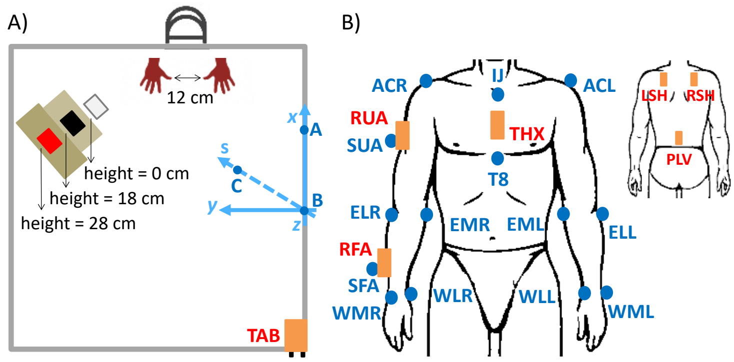

Fig. 1. A) Top view of table with hands silhouettes, boxes, IMU and markers global reference systems. B) IMUs and markers placement on participants' upper body.

2.4 Data analysis

Data analysis was conducted with Matlab® routines. The right wrist hit on the table was used for the temporal synchronization of data from both acquisition systems [19]. Then, markers coordinates were resampled at 50 Hz and data were expressed with respect to the global reference frame.

Anatomical reference systems of right forearm (Fig. 2.A) and right upper arm (Fig. 2.B) were defined from markers: x-axis was longitudinal to the segment; z-axis was perpendicular to the plane defined by x-axis and support s-axis; y-axis was obtained as the cross product between z-axis and x-axis. Trunk anatomical reference system was composed of a vertical x-axis, a horizontal y-axis and a z-axis obtained as their cross product (Fig. 2.C). IMUs data were referred to these anatomical axes by considering a constant transformation between IMUs and markers reference frames identified during the initial calibration in neutral position. The consistency of sensor fusion operation was verified by estimating elbow and shoulder angles from both IMUs and markers systems [20, 21] and calculating Root Mean Square Error (RMSE) values.

Finally, signals of x-acceleration of RUA-IMU and y-coordinate of WMR marker have been analysed to identify the presence of features characterizing the pick and place gesture performed by subjects and to discriminate which box was managed. In both signals, a double peak identifies "pick and place" gesture of the box. Signals trends then showed 15 couples of peaks related to the 15 pick gestures. For both signals of all subjects, the amplitude of each couple of consecutive peaks was averaged.

Then, mean (m) and standard deviation (σ) of the obtained 15 mean values were used to define a band (m ± σ/2). Each gesture was identified as the "pick and place" of a box which was in low, medium or high position, depending if the peaks fell below, inside or above the band, respectively. The accuracy of gesture recognition was estimated for all participants for both signals.

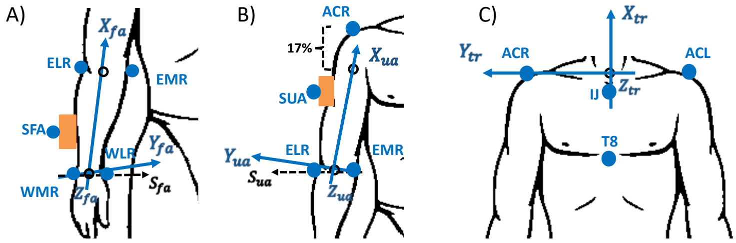

Fig. 2. Anatomical reference systems of A) forearm; B) upper arm; C) trunk.

3 Results and discussion

Fig. 3 shows joints angular trends estimated for one subject with both systems. The left panel contains elbow flexion-extension (FE) and pronation-supination (PS); the right panel shows shoulder FE, abduction-adduction (AA) and intra-extra rotation (IE). Angular trends are very similar, testifying the accuracy of the performed sensor fusion operation.

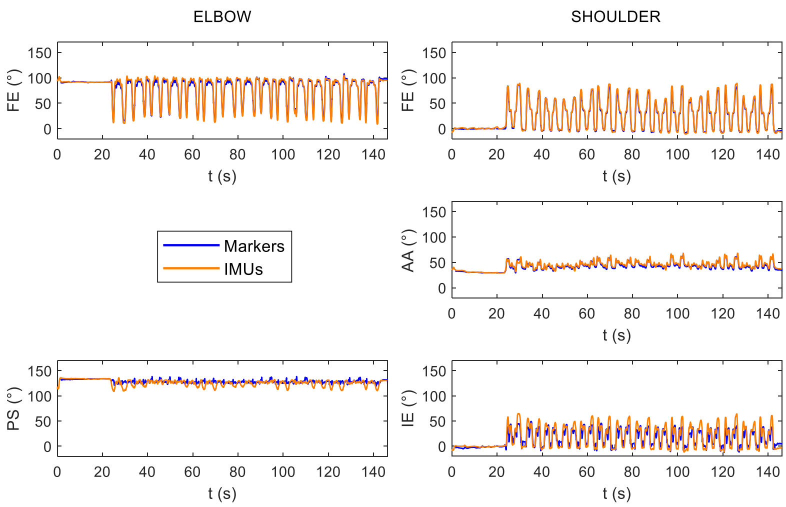

Fig. 3. Joints angular trends estimated with both markers and IMUs.

Table 1 contains RMSE values quantifying the correspondence of angular trends calculated using markers and IMUs data. Except in some cases, (elbow FE of subject 6, elbow AA and PS of subject 8, shoulder IE of subject 1), RMSE values lower than 9.8° reflect this correspondence in line with literature [22]. Errors can be due to skin artefacts and IMUs magnetic disturbances. Fig. 4 shows the distinction between pick and place of a low, a medium and a high box (grey, black and red dots, respectively), considering both x-acceleration of RUA-IMU and y-coordinate of WMR marker. The inertial feature had an accuracy of 98.7%, by mismatching medium and high gestures in only 2 cases out of 150.

Table 1. RMSE values of elbow FE and PS and shoulder FE, AA and IE angles (°).

Sub	Elbow FE	Elbow PS	Shoulder FE	Shoulder AA	Shoulder IE
1	3.5	5.3	3.9	6.3	11.9
2	3.0	2.9	8.2	3.8	4.6
3	3.9	5.8	3.0	2.8	9.8
4	4.0	4.0	2.4	2.3	7.3
5	3.7	6.3	3.3	4.1	4.4
6	13.3	5.5	2.7	1.6	5.9
7	3.8	5.9	2.2	1.9	9.0
8	4.2	12.8	6.4	8.4	7.5
9	3.4	3.8	3.1	1.9	3.8
10	4.7	6.4	2.6	2.3	8.1

This could be justified by the smaller height difference between black and red boxes. Instead, as expected, the accuracy of the spatial feature was of 100%. Overall, both signals were suitable for distinguishing among the three gestures for all subjects.

In conclusion, the present database appears to be congruent, complementary and suitable for features identification. Hence, future plans are first to identify other motion features able to distinguish among gestures. Then, these features could be used for development and training of a prediction algorithm of human upper limbs kinematics in an industrial context.

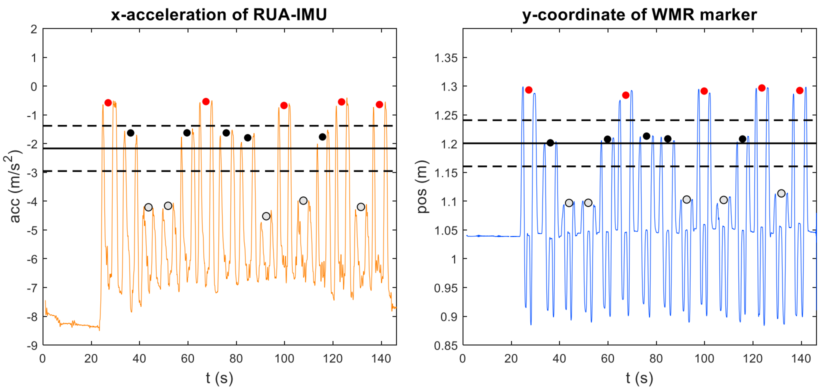

Fig. 4 The algorithm adopted to distinguish among low, medium and high gestures. Two signals: x-acceleration of RUA-IMU (left); y-coordinate of WMR marker (right).

References

1. Bauer, A., Wollherr, D., Buss, M.: Human-robot collaboration: A survey. Int. J.

Humanoid Robot. 5, 47–66 (2008).
2. Ajoudani, A., Zanchettin, A.M., et al.: Progress and prospects of the human–robot collaboration. Auton. Robots. 42, 957–975 (2018).
3. Mauro, S., Pastorelli, S., Scimmi, L.S.: Collision avoidance algorithm for collaborative robotics. Int. J. Autom. Technol. 11, 481–489 (2017)
4. Lasota, P.A., Fong, T., Shah, J.A.: A Survey of Methods for Safe Human-Robot Interaction. Found. Trends Robot. 5, 261–349 (2017).
5. Melchiorre, M., Scimmi, L.S., Mauro, S., Pastorelli, S.: Influence of human limb motion speed in a collaborative hand-over task. ICINCO 2018 - Proc. 15th Int. Conf. Informatics Control. Autom. Robot. 2, 349–356 (2018).
6. Perez-D'Arpino, C., Shah, J.A.: Fast target prediction of human reaching motion for cooperative human-robot manipulation tasks using time series classification. Proc. - IEEE Int. Conf. Robot. Autom. 2015–June, 6175–6182 (2015).
7. Pereira, A., Althoff, M.: Overapproximative arm occupancy prediction for human-robot co-existence built from archetypal movements. IEEE Int. Conf. Intell. Robot. Syst. 2016–Novem, 1394–1401 (2016).
8. Wang, Y., Ye, X., Yang, Y., Zhang, W.: Collision-free trajectory planning in human-robot interaction through hand movement prediction from vision. IEEE-RAS Int. Conf. Humanoid Robot. 305–310 (2017).
9. Mainprice, J., Berenson, D.: Human-robot collaborative manipulation planning using early prediction of human motion. IEEE Int. Conf. Intell. Robot. Syst. 299–306 (2013).
10. Ghosh, P., Song, J., Aksan, E., Hilliges, O.: Learning human motion models for long-Term predictions. Proc. - 2017 Int. Conf. 3D Vision, 3DV 2017. 458–466 (2018).
11. Butepage, J., Kjellstrom, H., Kragic, D.: Anticipating Many Futures: Online Human Motion Prediction and Generation for Human-Robot Interaction. Proc. - IEEE Int. Conf. Robot. Autom. 4563–4570 (2018).
12. Ionescu, C., Papava, D., Olaru, V., Sminchisescu, C.: Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments. IEEE Trans. Pattern Anal. Mach. Intell. 36, 1325–1339 (2014).
13. Xia, S., Wang, C., Chai, J., Hodgins, J.: Realtime style transfer for unlabeled heterogeneous human motion. ACM Trans. Graph. 34, 1–10 (2015).
14. Müller, M., Röder, T., Clausen, M., Eberhardt, B., Krüger, B., Weber, A.: Documentation mocap database hdm05. Tech. Rep. (2007)
15. Holden, D., Saito, J., Komura, T.: A deep learning framework for character motion synthesis and editing. ACM Trans. Graph. 35, 1–11 (2016).
16. Ofli, F., Chaudhry, R., et al.: Berkeley MHAD: A comprehensive Multimodal Human Action Database. Proc. IEEE Work. Appl. Comput. Vis. 53–60 (2013).
17. De La Torre, F., Hodgins, J., Bargteil, A.W., Martin, X., Macey, J.C., Collado, A., Beltran, P.: Guide to the Carnegie Mellon University Multimodal Activity (CMU-MMAC) Database. Robotics. 19 (2008)
18. De Leonardis, G., Rosati, S., Balestra, G., Agostini, V., Panero, E., Gastaldi, L., Knaflitz, M.: Human Activity Recognition by Wearable Sensors : Comparison of different classifiers for real-time applications. MeMeA 2018 - 2018 IEEE Int. Symp. Med. Meas. Appl. Proc. 1–6 (2018).
19. Digo, E., Pierro, G., Pastorelli, S., Gastaldi, L.: Tilt-twist Method using Inertial Sensors to assess Spinal Posture during Gait. In: International Conference on Robotics in Alpe-Adria Danube Region. pp. 384–392. Spinger, Cham (2019)
20. Rab, G., Petuskey, K., Bagley, A.: A method for determination of upper extremity kinematics. Gait Posture. 15, 113–119 (2002).
21. Wu, G., Van Der Helm, F.C.T., Veeger, H.E.J., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A.R., McQuade, K., Wang, X., Werner, F.W., Buchholz, B.:

ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion - Part II: Shoulder, elbow, wrist and hand. J. Biomech. 38, 981–992 (2005).
22. Peppoloni, L., Filippeschi, A., Ruffaldi, E., Avizzano, C.A.: A novel 7 degrees of freedom model for upper limb kinematic reconstruction based on wearable sensors. SISY 2013 - IEEE 11th Int. Symp. Intell. Syst. Informatics, Proc. 105–110 (2013).