ISMCR 2018 - 21th International Symposium on Measurement and Control in Robotics
26-28 September 2018 - International CBRNE Institute, MONS, Belgium
ROBOTICS FOR THE CHANGING WORLD

IMU based gesture recognition for mobile robot control using Online Lazy Neighborhood Graph search

P Kulkarni, B Illing, B Gaspers, B Brüggemann, D Schulz

Fraunhofer FKIE, Fraunhoferstraße 20, 53343 Wachtberg

E-mail: padmaja.kulkarni@fkie.fraunhofer.de

Abstract. In this paper, we present and evaluate a framework for gesture recognition using four wearable IMUs to indirectly control a mobile robot. Six gestures involving different hand and arm motions are defined. A novel algorithm based on Online Lazy Neighborhood Graph (OLNG) search is used to recognize the gestures. We use this algorithm to classify the gestures online and trigger predefined behaviors. Experiments show that the framework is able to correctly detect and classify six different gestures in real-time with an average success rate of 81.6%, while keeping the false positive rate low by design and using only 126 training samples.

1. Introduction

For robots to be able to work in unstructured environments, areas dangerous to humans, or disaster sites, human intelligence is still vital. In such cases, teleoperation of robots could be one of the solutions. With recent advancements in robotics, the complexity of using robots has also increased. Despite this fact, currently used technology limits the majority of man-machine interfaces to text or GUI based interfaces and joysticks. Such types of control can become cumbersome in case of, for example, robots with a heavy control box or high degrees of freedom. Often, working in disaster areas could be stressful for an operator. Hence, alternate and intuitive control paradigms need to be developed. Gesture-based control seems particularly useful as it can be very intuitive [1].

Vision-based gesture control is well researched but the set-up time and dependency on controlled environmental conditions, like lighting, make it less suitable for teleoperation in disaster areas. On the other hand, Hoffmann et al. [2] developed an Inertial Measurement Unit (IMU) based control for a robot manipulator, which does not need any infrastructure. They transformed human arm motions into corresponding robotic manipulator motions using five IMUs attached to the sleeve of a wearable-jacket. They showed that teleoperation performed in this way is very efficient and intuitive [3]. However, to trigger some predefined manipulator motion or to trigger robot base motions this direct control method cannot be used.

This paper presents an extension of the work done by Hoffmann et al. [2] in the area of wearable IMUs. We present a framework based on OLNG search, which is able to identify dynamic gestures in real-time and can be used to trigger predefined robot motions. The main contribution of this work is an implementation and evaluation of an OLNG search based algorithm for gesture recognition and robotics application. Prior to this, OLNG search algorithm was primarily used in the area of computer graphics [4].

2. Related work

Most approaches in the field of gesture recognition are based on vision, IMU, and Electromyography (EMG) signals [5] [6].

For IMU based approaches many use glove mounted sensors. Mummadi et al. [7] use an IMU based glove for real-time sign language recognition. They use various machine learning algorithms such as Support Vector Machines, Naive Bayes, Multi Layer Perceptron, and Random Forest to classify the gestures. Wu et al. [8] use a data glove with perception and Hidden Markov Models (HMMs) to classify hand gestures. Georgi et al. [6] couple IMU based motion with EMG muscle activity to recognize hand and finger gestures. They use HMMs for the gesture recognition and obtain 74.3% in accuracy with person independent settings. These methods only classify static gestures with hand and fingers [6] [7] and need a huge database (about 1000 samples of each kind) [7]. In the domain of commercial products, the Myo armband by Thalmic Labs uses EMG signals along with IMUs to detect up to five different gestures and motions of the arm, but these gestures are pre-set.

For vision based gesture recognition the Microsoft Kinect is widely utilized. OpenNI or Kinect SDK are used for motion tracking. For gesture identification, algorithms like Dynamic Time Warping, Artificial Neural Networks, or HMMs are implemented, for an overview see [9]. Amin et al. [10] developed a vision based technique to identify hand gestures using Principal Component Analysis and Gabor representation. However, vision based approaches suffer from limitations like occlusion and are vulnerable to bad performance from ambient lighting and background changes.

3. Approach

We assume that the IMU readings are available in the form of vectors at a discrete time interval $(\ldots, \vec{\alpha}_{t-2}, \vec{\alpha}_{t-1}, \vec{\alpha}_t, \ldots)$, where $\vec{\alpha}$ is a vector of Euler angles and t is time. This vector is referred to as input vector in this paper. An underlying training database consists of Euler angles obtained from the four IMUs. A 12-dimensional vector forms a data point in it. For building the database, sequences of such vectors labelled with the corresponding gesture names are saved while the gestures are being performed. Every vector in the database has a unique index *i,* which is later used for the identification of a particular vector.

For sequence matching, a window of m input vectors is defined. The distance between the current input vector and each vector in the training database is calculated and from this set of distances the indices of the k nearest vectors (neighbors) and their spatial distances from the input vector are obtained and stored using Fast Library for Approximate Nearest Neighbors (FLANN) in real-time. A matrix is created for the k nearest neighbors in the training database and for a window of m input vectors. Considering these $k \times m$ vectors as nodes, a graph is built. For each input vector, its k corresponding graph nodes are then sorted in ascending order of the database index i. For the sequence matching, nodes are chosen in such a way that their database indices are in strictly ascending order. A path is a sequence connected from the first vector in the window to the last vector. Figure 1 shows possible sequences for three input vectors. If no valid neighbor index for a particular input vector exists, then the neighbors of the next input vector are considered for the path. In that case, an extra path cost is added for skipping one input vector. In this way all possible paths are listed and the best path amongst them is chosen based on its minimum cost. OLNG search offers extremely fast sequence matching and is suitable for real-time applications due to its linear complexity [4].

For gesture recognition the following procedure is applied. For the whole database, k nearest neighbors to the current input vector are calculated using FLANN. OLNG is then used to find a matching and valid sequence. The best path found is saved along with its cost and sequence matching length for comparison. An arm and hand movement is considered to be a gesture if the sequence matching length is more than a certain threshold value and the cost of the path is less than another threshold. Both thresholds were chosen based on initial informal experiments and were tweaked and validated during our evaluation. Each recognized gesture triggers a specific motion of the robot. During the robot motion no gestures are detected. Gesture recognition restarts when the robot signals the completion of the motion.

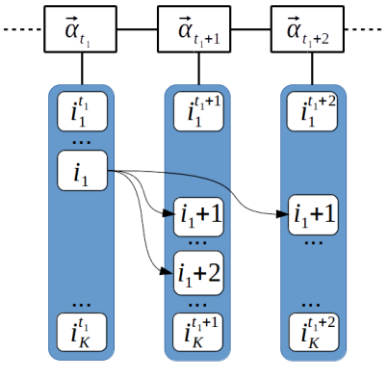

Figure 1. Path finding using OLNG search with k nearest neighbors [4].

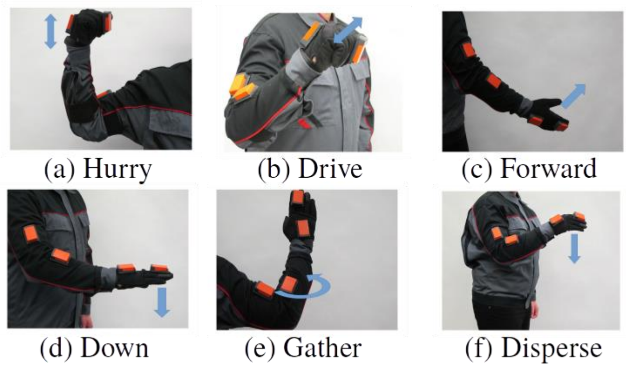

Figure 2. (a)-(f) show six different gestures defined for our experiment.

4. Experimental setup

We validated our OLNG search based algorithm by testing it for online gesture recognition. Xsens MTw sensors were used for the motion capturing and building a database. The sensor is a 9-axis IMU consisting of a 3-axis gyroscope, acceleration sensor, and magnetometer and it includes an extended Kalman filter. Internally the sensors operate at 75 Hz, but such a high rate does not offer much additional information during arm movements. Hence, we used an update frequency of 15 Hz. Four of such sensors were mounted on a jacket at humerus, radius, hand, and finger.

The training database used consisted of 21 motions of each gesture from three different users for a total of 126 samples. The test database consisted of 10 different motions of each gesture plus 10 random arm movements. To apply robot motions, a telemax manipulator from Telerob was used. The programming was done in C++ using MoveIt! motion planning framework with a Linux operating system and Robot Operating System as a middleware. For OLNG search we considered $k = 40$ nearest neighbors, a window of $m = 10$ vectors, and a path length of 9. Based on these fixed parameters the path cost was varied from 0 to 4 in steps of 0.05 squared radians.

We defined six gestures to test our algorithm. These are based on internationally used hand signals from a military context. Figure 2 shows the gestures used and the placement of the four IMUs on the jacket.

5. Results

Receiver Operating Characteristic (ROC) curves for all six gestures with varying path costs are shown in figure 3. Based on the ROC curves the threshold path cost was chosen to be 1.80 with 81.61% true-positive rate and false-positive rate of 15.12%. By including random arm movements in the test database, all thresholds were also validated to only recognize a gesture when one was performed. The confusion matrix corresponding to the chosen cost is shown in figure 4. 'Drive' and 'Hurry' were recognized the best at the chosen cost with 95.65% and 90.10% prediction rate respectively. It can be observed that the gesture 'Disperse' is more difficult to recognize as it partly shares similarity with the whole gesture 'Down'. The same goes for the gestures 'Gather' and 'Hurry', which leads to the non-symmetrical confusion matrix

6. Conclusion and future work

A novel algorithm based on OLNG search was implemented and tested for the application of gesture recognition in real-time. A software framework to trigger predefined robot motions based on a detected gesture was implemented. Experiments show that we could obtain a gesture recognition rate of 81.61% while keeping the false positive rate low. In the future, we would like to expand our training database on the fly by adding correctly recognized gestures to it. We would also like to extend our algorithm to match parts of start, middle, and end of a gesture to counter similarity between different gestures.

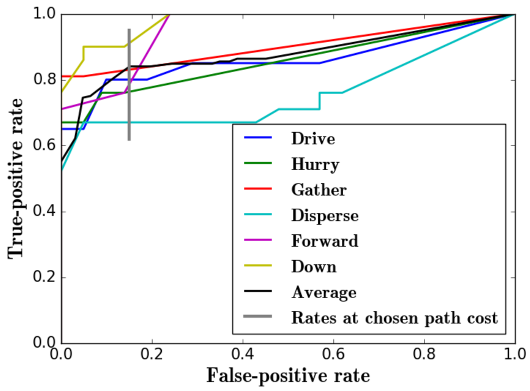

Figure 3. ROC curve for all gestures with variable threshold path cost.

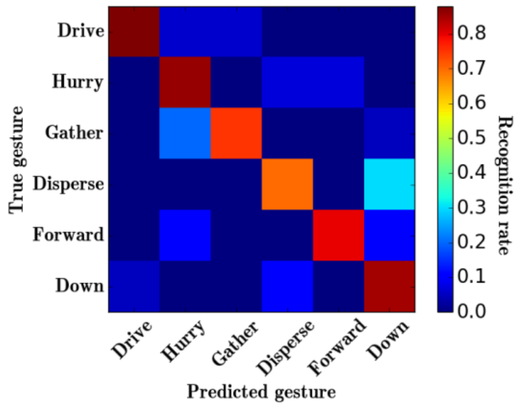

Figure 4. Confusion matrix for gestures with the chosen path cost.

References:

[1] Coronado E, Villalobos J, Bruno B and Mastrogiovanni F 2017 Gesture-based robot control: Design challenges and evaluation with humans *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 2761–2767

[2] Hoffmann J, Brüggemann B and Krüger B 2010 Automatic calibration of a motion capture system based on inertial sensors for tele-manipulation *Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO)* pp 121–128

[3] Brüggemann B, Gaspers B, Ciossek A, Pellenz J and Kroll N 2013 Comparison of different control methods for mobile manipulation using standardized tests *IEEE Int. Symp. on Safety, Security, and Rescue Robotics (SSRR)* pp 1–2

[4] Tautges J 2012 *Reconstruction of human motions based on low-dimensional control signals* Dissertation University of Bonn

[5] Kavarthapu D C and Mitra K 2017 Hand gesture sequence recognition using Inertial Motion Units (IMUs) *Asian Conf. on Pattern Recognition (ACPR)*

[6] Georgi M, Amma C and Schultz T 2015 Recognizing hand and finger gestures with IMU based motion and EMG based muscle activity sensing *Int. Conf. on Bio-inspired Systems and Signal Processing (BIOSIGNALS)* pp 99–108

[7] Mummadi C K, Leo F P P, Verma K D, Kasireddy S, Scholl P, Kempfle J and Laerhoven K V 2018 Real-time and embedded detection of hand gestures with an IMU-based glove *Informatics* **5** pp 28–45

[8] Jiangqin W, Wen G, Yibo S, Wei La and Bo P 1998 A simple sign language recognition system based on data glove *IEEE Int. Conf. on Signal Processing (ICSP)* pp 1257–1260

[9] Nandwana B, Tazi S, Trivedi S, Kumar D and Vipparthi S K 2017 A survey paper on hand gesture recognition *Int. Conf. on Communication Systems and Network Technologies (CSNT)* pp 147–152

[10] Amin M A and Yan H 2007 Sign language finger alphabet recognition from Gabor-PCA representation of hand gestures *Int. Conf. on Machine Learning and Cybernetics* (ICMLC) pp 2218–2223