A Novel Bayesian Filtering Approach to Tactile Object Recognition

G. Vezzani[1], N. Jamali[2], U. Pattacini[2], G. Battistelli[3], L. Chisci[3], and L. Natale[2]

Abstract— **This paper addresses tactile object recognition, i.e. the identification of an object among a set of known objects, given tactile measurements. The solution of this problem can improve perception capabilities of autonomous robots and complement vision. Such a system is fundamental for the operation of autonomous robots that are often required to recognize objects while interacting with the environment.**

The proposed approach is innovative for three reasons. First, tactile recognition is cast into a tactile localization problem wherein multiple models are fit to the available measurements and objects are recognized by selecting the model that minimizes the localization error. Second, the measurements consist only of 3D contact point coordinates, which provide poor information for the recognition task. Lastly, we make use of a novel and effective filtering algorithm, named Memory Unscented Particle Filter (MUPF), which solves the 6 degree-of-freedom localization (and recognition) problem recursively by using only contact point measurements. The performance of the proposed approach has been assessed both in simulation and on a real robotic system equipped with tactile sensors (i.e., the iCub humanoid robot). The experiments show that our approach provides good recognition performance and is able to discriminate objects that are similar even in presence of noisy measurements.

I. INTRODUCTION

Findings in human physiology testify to how the sense of touch is irreplaceable for human beings [1], [2], especially during exploration in the dark or in the presence of visual occlusions. Like humans, autonomous robots can take advantage of haptic perception to make manipulation and recognition tasks more efficient, complementing vision – when unavailable or imprecise.

In the last few decades, rapid advances in tactile technology have made it possible to build tactile systems that are reliable enough to be deployed on real robots at a reasonable cost [3]–[5]. Among possible applications, researchers have investigated the problem of object recognition using tactile feedback. Various tactile features have been proposed to perform material [6]–[8] and local curvature classification [9]. The problem of object recognition, however, requires that features from various contact locations are integrated in a coherent representation. Object recognition using tactile feedback is, therefore, a challenging problem that requires filtering techniques for fusing noisy measurements.

Different methods have been proposed in the literature in order to solve tactile object recognition. They can be classified depending on the type of information they use and the object features they recover, namely, material and shape properties. Some researchers have focused on identifying material properties [6]–[8]. Decherchi *et al.* use multiple techniques to classify object materials with tactile data [7]. Liu *et al.* [8] apply a dynamic friction model to determine physical properties of surfaces while a robotic finger slides along the object with different speeds.

To recognize object shapes, a viable approach is to recover local geometry from each contact point, i.e., surface normal and curvature. By using a cylindrical tactile sensor, Fearing *et al.* propose a nonlinear, model-based inversion to recover contact surface curvatures [9]. Contact location point-clouds have also been used to reconstruct object shapes with computer graphic techniques [10]–[13]. Allen *et al.* fit points from tactile sensors readings to super-quadric surfaces to reconstruct unknown shapes [12]. A similar approach, proposed by Charlebois [14], uses tensor B-spline surfaces instead of super-quadratic surfaces. Through these methods, arbitrary object shapes can be identified by estimating surface curvatures.

Another solution to recognizing object shapes is to use machine learning techniques on the output of tactile sensor arrays. In this case, object features are extracted from the tactile data. A classifier is then trained to predict the shapes of novel objects [15], [16].

This paper proposes a different approach to the problem of tactile object recognition. Under the assumption that object models are known, the tactile recognition task is solved using a *nonlinear multimodal filtering approach* and is framed as a *tactile localization problem*. The robot explores an object using its tactile sensors, registering the 3D coordinates of the finger-object contact locations. The contact locations collected during the exploration are, then, compared with the object models. The object is recognized as the object whose model better fits the measurements, i.e., the object model with the lowest localization error.

This technique works with measurements consisting only of a set of 3D contact point coordinates, which can be collected by the robot using tactile feedback. Such data provide very basic, and noisy information, making the tactile recognition task more challenging.

The adopted algorithm, named *Memory Unscented Particle Filter* (MUPF) [17], is designed to efficiently solve

[1]G. Vezzani is with the Istituto Italiano di Tecnologia (IIT), iCub Facility, Via Morego 30, Genova, Italy. She is also with the University of Genova, Via All'Opera Pia, 13, 16145 Genova giulia.vezzani@iit.it.

[2]N. Jamali, U. Pattacini and L. Natale are with the Istituto Italiano di Tecnologia (IIT), iCub Facility, Via Morego 30, Genova, Italy nawid.jamali@iit.it, ugo.pattacini@iit.it, lorenzo.natale@iit.it.

[3]G. Battistelli and L. Chisci are with the Dipartimento di Ingegneria dell'Informazione, Universita' degli Studi di Firenze, Via S. Marta 3, Firenze, Italy giorgio.battistelli@unifi.it, luigi.chisci@unifi.it

the global 6 degree-of-freedom localization problem. The MUPF relies on the *Unscented Particle Filter* (UPF) [18]. We demonstrate that our approach is effective in solving the tactile recognition tasks in simulation as well as in a real experiment using the iCub humanoid robot and its tactile system [3].

The paper is organized as follows. Section II provides a brief introduction to nonlinear filtering techniques, followed by a mathematical (Bayesian) formulation of the tactile localization problem and the MUPF description. Section III presents the exploration strategy for acquiring measurements. Section IV demonstrates the effectiveness of the proposed solution by means of simulation and experimental tests on the iCub humanoid robot. In Section V we give our concluding remarks and future directions.

II. METHODOLOGY

We introduce hereinafter the problem of tactile object recognition. Let k denote the number of objects of interest, each object being represented by a mesh model consisting of triangular faces $\{f_i\}$. A set of measurements $\{\boldsymbol{y}_t\}_{t=1}^{L}$ is collected using the tactile sensors by detecting contacts on the surface of object k^* (one of the k objects). It is assumed that object is attached to a surface and, thus, does not move during the exploration. Each measurement provides the 3D coordinates of the contact point, i.e. $\{\boldsymbol{y}_t = (x_t, y_t, z_t)\}_{t=1}^{L}$. The goal is to infer on which object the measurements have been collected. In the described scenario, the solution is given by the object model that best fits the available measurements.

A. Recognition as Multi-object Localization

We address the tactile object recognition problem as a localization problem applied to multiple objects, where the solution is provided by the object whose localization error is the lowest among all the considered objects.

Then, once a localization algorithm is provided, object recognition is achieved by simply running such an algorithm for each of the given object models. For each possible object $l \in \{1, \ldots, k\}$, the algorithm finds the pose $\hat{\boldsymbol{x}}_l$ that makes the object model representing the l^{th} object best fit the set of measurements. Once the pose $\hat{\boldsymbol{x}}_l$ is calculated for each object models $l \in \{1, \ldots, k\}$, a suitable *performance index*, $\mathcal{I}_l$, is used in order to measure the fitness of each object model in the estimated pose l. To this end, the following *performance index* is introduced:

$$\mathcal{I}_l = \frac{1}{L} \sum_{i}^{L} d_{i,l}, \tag{1}$$

where L is the number of measurements and $d_{i,l}$ is the distance between the i^{th} measurement and the object l in the estimated pose $\hat{\boldsymbol{x}}_l$. In other words, given the set of measurements and the estimated pose, the proposed performance index is the average of the distances between each measurement and the object model in the estimated pose.

TABLE I
TACTILE RECOGNITION ALGORITHM

1: **Data:** k object models, a set of tactile measurements $\{\boldsymbol{y}_t\}_{t=1}^{L}$ on object k^*
2: **for** $l = 1, \ldots, k$ **do**
3: **Localization algorithm:**
data: object model l, set of measurements on object k^*
output: $\hat{\boldsymbol{x}}_l$
4: **end for**
5: Choose $\hat{k}$ as:
$$\hat{k} = \arg\min_{l} \mathcal{I}_l.$$
where $\mathcal{I}_l = \frac{1}{L}\sum_{i}^{L} d_{i,l}$ and $d_{i,l}$ is the distance between the i^{th} measurement and the object l in the estimated pose $\hat{\boldsymbol{x}}_l$
6: Recognition is successful if $\hat{k} = k^*$

Finally, after the k executions of the localization algorithm, the quantities $\mathcal{I}_l$, for $l = 1, \ldots, k$, are available and the solution $\hat{k}$ for the tactile recognition problem is given by:

$$\hat{k} = \arg\min_{l} \mathcal{I}_l. \tag{2}$$

Clearly, the recognition is successful when $\hat{k} = k^*$. The steps of our algorithm for the tactile recognition stated as a localization problem are outlined in Table I.

B. The Filtering Approach

In the present work, tactile localization is cast into the Bayesian framework and addressed as a nonlinear multi-modal filtering problem. In Bayesian filtering we seek to estimate the system state $\boldsymbol{x}_t \in \mathbb{R}^n$ at time t, assuming that its time evolution can be described by a *Markov transition density* $\varphi_{t+1|t}(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)$. Suppose also that the measurements $\{\boldsymbol{y}_t\}_{t=1}^{L} \in \mathbb{R}^p$ depend on the state $\boldsymbol{x}_t$ through a *measurement likelihood* function $\ell(\boldsymbol{y}_t|\boldsymbol{x}_t)$, which denotes the probability that the measurement will take value $\boldsymbol{y}_t$ given the state $\boldsymbol{x}_t$. From a Bayesian viewpoint, the goal of the filtering problem is to recursively compute the conditional PDFs $p_{t|t}(\boldsymbol{x}) = p(\boldsymbol{x}_t = \boldsymbol{x}|\boldsymbol{y}_t)$ and $p_{t+1|t}(\boldsymbol{x}) = p(\boldsymbol{x}_{t+1} = \boldsymbol{x}|\boldsymbol{y}_t)$, by taking advantage of the observations $\{\boldsymbol{y}_t\}_{t=1}^{L}$.

The solution of the filtering problem is given by the Bayesian recursion and consists of the following Bayes and Chapman-Kolmogorov equations, respectively:

$$p_{t|t}(\boldsymbol{x}) = \frac{\ell_t(\boldsymbol{y}_t|\boldsymbol{x}) p_{t|t-1}(\boldsymbol{x})}{\int \ell_t(\boldsymbol{y}_t|\boldsymbol{\xi}) p_{t|t-1}(\boldsymbol{\xi}) d\boldsymbol{\xi}} \tag{3}$$

$$p_{t+1|t}(\boldsymbol{x}) = \int \varphi_{t+1|t}(\boldsymbol{x}|\boldsymbol{\xi}) p_{t|t}(\boldsymbol{\xi}) d\boldsymbol{\xi}\,. \tag{4}$$

An analytical solution to (3) and (4) is available only in a few cases, among which the linear-Gaussian is the most notable, leading to the well-known Kalman filter recursion. In many practical applications such as our problem, the transition and likelihood models are usually affected by nonlinearities and/or non-Gaussian noise distributions, thus precluding analytical solutions and making approximation techniques the only viable approach.

Most of the existing approximation techniques can be divided in two families: Kalman-filtering-like approaches,

and sequential Monte Carlo methods. The algorithms belonging to the former family are characterized by a lower computational cost, but are not appropriate for multimodal distributions like the one arising in the tactile localization. On the other hand, sequential Monte Carlo methods, also known as particle filters [19], can deal with arbitrary nonlinearities and distributions and supply a complete representation of the posterior state distributions. The main drawback of the latter techniques is that, in order to make the approximation sufficiently accurate, the number N of particles must increase exponentially with the dimension n of the vector to be estimated (since it is required to sample in a subset of $\mathbb{R}^n$). Many variants of particle filtering have been proposed in order to mitigate such a downside and approximate the posterior reasonably well with a moderate number of particles. Among the most effective variants, there is the Unscented Particle Filter (UPF) that exploits the Unscented Kalman Filter (UKF) to improve performance [18].

Given the dimensionality of the problem (Section II-C shows that the quantity to be estimated belongs to $\mathbb{R}^6$), the strong nonlinearity and the non-Gaussian nature of the distributions, the UPF is the most suitable candidate for the solution of the tactile localization. In short, the UPF is a variant of the particle filter, which exploits a UKF for each particle. The UPF propagates a set of extended particles, with the i^{th} particle comprising a weight w_t^i, a mean $\bar{\boldsymbol{x}}_t^i$ and a covariance $P_{t|t}^i$. Then, the UKF is applied to each particle mean and covariance so as to move the particle towards the measurements, thus achieving a more dense sampling in the most relevant areas of the search space. After this step, the current observation is used in weight computation and the algorithm proceeds like a standard particle filter.

C. Localization Problem Formulation

Hereafter, we adapt the filtering notation provided in the previous section to our case study. Our goal can be formulated as a peculiar filtering problem. The object k^* to be localized and, then, recognized among the other objects, is rigidly attached to a surface, hence it does not undergo any movement during the exploration. Thus, the entity to be estimated consists of the object pose, which does not depend on time. For this reason, in our problem, the system state $\boldsymbol{x}$ is in $\mathbb{R}^6$ (i.e. $n = 6$) and is defined by:

$$\boldsymbol{x} = \begin{bmatrix} x, & y, & z, & \phi, & \theta, & \psi \end{bmatrix}^T, \tag{5}$$

where x, y, z are the coordinates of the center of the reference system attached to the object model and ϕ, θ, ψ are the three Euler angles representing orientation.

The observations $\{\boldsymbol{y}_t\}_{t=1}^L$ exploited to localize the object consist of the tactile measurements

$$\{\boldsymbol{y}_t = (x_t, y_t, z_t)\}_{t=1}^L, \tag{6}$$

which represent the Cartesian positions of the contact points in $\mathbb{R}^3$ (i.e. $p = 3$).

In order to correctly formulate the tactile localization problem in the filtering framework, we need to define other mathematical quantities, as follows:

- In cases the state to be estimated is stationary, as in our problem, the state transition equation can by expressed as:
$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t + \boldsymbol{\omega}_t, \tag{7}$$
where $\boldsymbol{\omega}_t$ is a small artificial noise [20]. This term is introduced in order to allow the filtering technique to change the estimate of $\boldsymbol{x}$ and then converge to the final solution. We model $\boldsymbol{\omega}_t$ as a Gaussian noise with zero mean and covariance Q.
- The likelihood function $\ell_t(\boldsymbol{y}_t|\boldsymbol{x}_t)$ is based on the so-called *proximity model* [21], in which the measurements are considered independent of each other and corrupted by Gaussian noise. For each observation, the likelihood function depends on the distance between the measurement and the object model, hence the name "proximity". The likelihood is defined as:
$$\ell_t(\boldsymbol{y}_t|\boldsymbol{x}) \propto \max_i \ell_{t,i}(\boldsymbol{y}_t|\boldsymbol{x}), \tag{8}$$
where $\ell_{t,i}(\boldsymbol{y}_t|\boldsymbol{x})$ is the likelihood of the measurement $\boldsymbol{y}_t$ produced by the i^{th} face of the object model, when the object is in the pose $\boldsymbol{x}$. The quantity $\ell_{t,i}(\boldsymbol{y}_t|\boldsymbol{x})$ is assumed to be Gaussian, with variance σ_p^2 and is described by:
$$\ell_{t,i}(\boldsymbol{y}_t|\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left(-\frac{1}{2}\,\frac{d_i(\boldsymbol{y}_t, \boldsymbol{x})^2}{\sigma_p^2}\right), \tag{9}$$
where $d_i(\boldsymbol{y}_t, \boldsymbol{x})$ is the shortest Euclidean distance of $\boldsymbol{y}_t$ from the face f_i when the object is in the pose $\boldsymbol{x}$. For instance, supposing that f_i is the representation of the i^{th} face in the object reference system, the distance $d_i(\boldsymbol{y}_t, \boldsymbol{x})$ is given by:
$$d_i(\boldsymbol{y}_t, \boldsymbol{x}) = \min_{\boldsymbol{p} \in f_i} \|\boldsymbol{y}_t^{\boldsymbol{x}} - \boldsymbol{p}\|,$$
where $\|\cdot\|$ is the Euclidean norm and $\boldsymbol{y}_t^{\boldsymbol{x}}$ denotes the transformation of the measurement $\boldsymbol{y}_t$ using the roto-translation matrix corresponding to the state $\boldsymbol{x}$.
- Further, in order to exploit UPF algorithm, we need to define a measurement function, namely, a mapping expressing the observation $\boldsymbol{y}_t$ as a function of the state $\boldsymbol{x}$ and a measurement noise $\boldsymbol{\nu}_t$:
$$\boldsymbol{y}_t = \boldsymbol{h}_t(\boldsymbol{x}) + \boldsymbol{\nu}_t\,. \tag{10}$$
The tactile sensors provide the 3D position of a point on the object surface which is touched during the exploration and, consequently, is the nearest to the sensors itself. For this reason, the measurement equation we propose for tactile sensors is:
$$\boldsymbol{h}_t(\boldsymbol{x}) = \arg\min_{\boldsymbol{p} \in \partial\mathcal{O}} \|\boldsymbol{p}_t^{\boldsymbol{x}} - \boldsymbol{p}\|, \tag{11}$$
where $\partial\mathcal{O} = \bigcup_i f_i$ and $\boldsymbol{p}_t^x$ is the position of the tactile sensor.

The Bayesian framework provides a recursive method for solving the tactile localization problem. In fact, the measurements are recursively used and, at each time step, the pose $\boldsymbol{x}$ of the stationary object is refined, converging to the final estimate by incorporating more recent measurements.

D. The Memory Unscented Particle Filter

The employed algorithm for solving the tactile localization problem is a modification of the standard Unscented Particle Filter. The tactile measurements are relatively uninformative if used individually, since they are three-dimensional vectors in a 6D space. This fact implies that the standard UPF algorithm is not well suited to this problem, since it exploits only the current measurement $\boldsymbol{y}_t$ at each time step. Such a behavior is somewhat critical, because the algorithm might end up limiting the search within wrong sub-regions, thus ruling out potential representative solutions.

In order to overcome this drawback, we make use of a variant of the UPF, referred to as Memory UPF (MUPF) [17]. In the MUPF algorithm, a limited number of past measurements are used during each iteration. The main two changes applied to the standard UPF can be summarized as follows.

First, the importance weights $\{w_t^i\}_{i=1}^N$ are updated by resorting also to past observations. In case the number of tactile acquisitions remains limited (as often occurs in practice), a growing memory strategy can be adopted by computing the importance weights with all the measurements collected up to time t. On the other hand, in case a larger number of measurements is required, for example, due to the complexity of the object model, the computational burden can be reduced by following a moving window strategy, where only a given number m of the most recent acquisitions are used at each time instant.

Second, since in the first iterations only a few measurements are available, that is, we do not have sufficient information yet, all particles are retained so as to account for likely solutions, in accordance with the multimodal nature of the problem. This amounts to skipping the standard resampling step for a certain number of initial time instants (in the experimental results reported in the following sections, for the first two time instants). The degeneration of the weights in the first iterations is avoided by setting the weights of all particles equal to $1/N$, N being the number of particles.

After all the L measurements have been processed, i.e. when $t = L$, the algorithm outputs, as final estimate of the object configuration, the corrected particle $\bar{\boldsymbol{x}}_L^i$ corresponding to the highest value of the estimated posterior distribution $\hat{p}_{L|L}(\cdot)$. The adoption of a maximum *a posteriori* probability (MAP) criterion is motivated by the strongly multimodal nature of the density, due to the fact that, in the presence of symmetries in the object, there might exist multiple values of $\boldsymbol{x}$ compatible with the measurements. In fact, in a multimodal case, taking the expected value as estimate is not meaningful. Recalling that each corrected particle can be considered corresponding to a Gaussian distribution with mean $\bar{\boldsymbol{x}}_t^i$ and covariance $P_{t|t}^i$, the estimated posterior $\hat{p}_{L|L}(\cdot)$ can be obtained as:

$$\hat{p}_{L|L}(\boldsymbol{x}) = \sum_{i=1}^{N} w_L^i \, \mathcal{N}(\boldsymbol{x}; \bar{\boldsymbol{x}}_L^i, P_{L|L}^i). \tag{12}$$

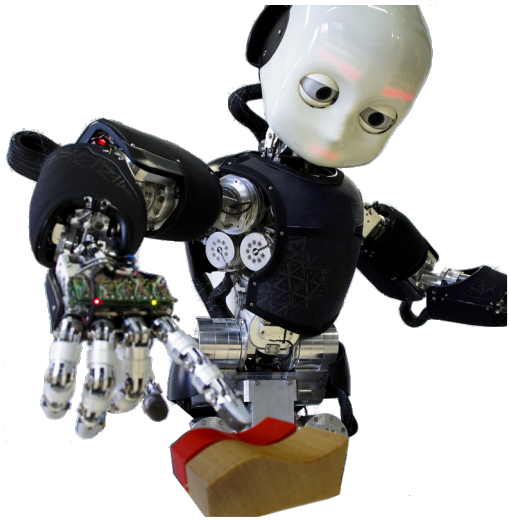

Fig. 1. Experimental setup for data collection: the iCub robot is touching the object with its index fingertip.

Hence, the particle with the MAP probability [22] can be readily obtained as:

$$\hat{\boldsymbol{x}} = \arg \max_{j \in \{1,\dots,k\}} \hat{p}_{L|L}(\bar{\boldsymbol{x}}_L^j). \tag{13}$$

III. DATA ACQUISITION

As shown in Fig. 1, the experimental setup consists of the iCub robot [23] and six objects of interest (i.e. $k = 6$) for acquiring tactile data in our experiments. The objects, as shown in Fig. 2, are made of wooden geometric shapes. The objects are deliberately selected to have overlapping shapes with strong similarities in order to test our method in a challenging setting. For example, objects (a) and (b) have similar geometric configurations: one has a smooth arched surface and the other a saw-tooth surface, respectively. With the same principle we also selected objects (c) and (d), that have same general shape, the only difference being in the smoothness of the surfaces. Objects (e) and (f) can only be discriminated by the bottom edge: one has a straight edge, while the other has a curved edge.

The robot touches the object at various locations with the tip of its index finger. The fingertip is 14.5 [mm] long, 13 [mm] wide. Each finger is equipped with tactile sensors [4]. A contact location is registered when the tactile sensors are activated. In our experiments the object is anchored to the surface of a table in front of the robot, hence, it does not move during the exploration. The choice of the exploratory area depends on the the size of the object. We sample an area of 40×50 [mm^2] (Fig. 3), using a grid search with a cell size of 2.5×2.5 [mm^2].

At the beginning of the exploration, the robot's index finger is placed at an arbitrary position close to the object. Then, the robot is commanded to sample a location of interest. We will refer to the location of interest as a waypoint. Since we do not have a priori knowledge of the shape of the object, the height of the waypoint is set to an arbitrary value larger than the height of the object. As reported in the flow chart

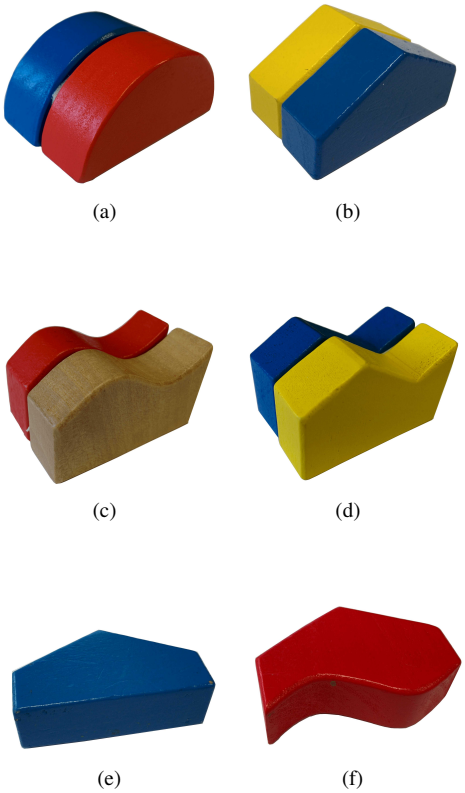

Fig. 2. Objects used for experimental evaluation of the method.

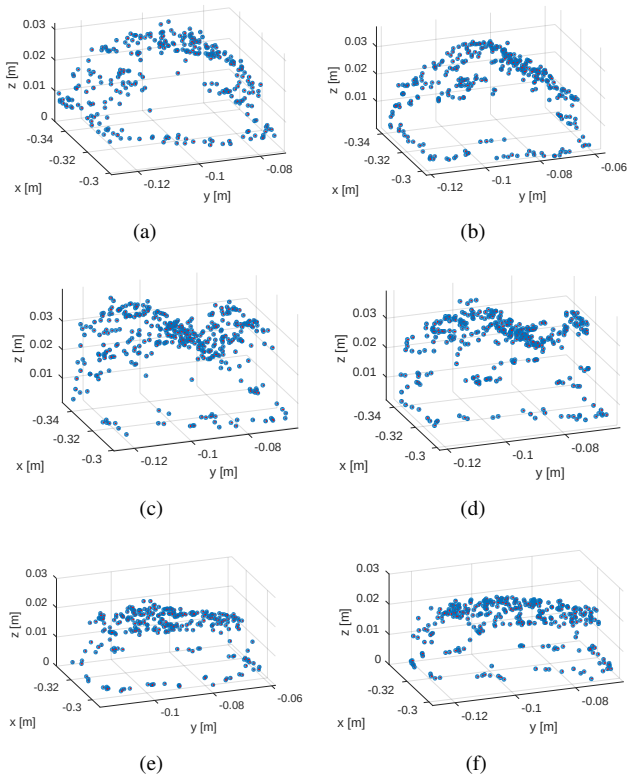

Fig. 3. For each object, the tactile data collected by using the exploration strategy of Section III are shown. The letters identifying the different plots ((a) - (f)) correspond to the objects according to the notation of Fig. 2.

of Fig. 4, the robot moves the finger toward the waypoint. After that, the robot extends its finger downward to detect a contact. If no contact is detected when the finger is fully extended, the robot sets the waypoint to the current location of the finger and retracts it. This process is repeated until the finger makes a contact with a surface – either the object or the table. When a contact is detected, the location of the contact is registered and the next waypoint is set to the next point in the grid. This process is repeated until the area is entirely covered. The tactile data collected for each object with this exploration strategy are shown in Fig. 3.

IV. RESULTS

The algorithm evaluation is performed first with synthetic measurements (Section IV-A) and then with real measurements (Section IV-B), collected through the exploration strategy described in Section III. In both scenarios, the aim is to recognize the true object labeled as k^*, among the set of six objects shown in Fig. 2.

The C++ implementation of the MUPF algorithm used to carry out our experiments is publicly available on GitHub[1].

A. Simulation Results

The synthetic measurements consist of six sets of 3D points (around 170 triplets for each set), each sampled on the surface of one specific model. We refer to the 3D points sampled on object (a) as *set of measurements (a)*. The same notation is used for the other objects. The synthetically-generated data are noiseless.

In Table II, the MUPF parameter set used for running the simulated tests is shown. Matrix Q and σ_p are respectively the covariances of the process noise $\boldsymbol{\omega}_t$ and measurement noise $\boldsymbol{\nu}_t$; P_0 is the covariance matrix representing the initial uncertainty and N is the number of particles. The covariance Q is chosen such that it takes into account the stationarity of the object, similarly, the value of the covariance σ_p models the measurement noise. An arbitrarily large value is instead chosen for P_0 matrix. The selected number of particles N is a trade-off between algorithm execution time and reliability. In order to determine a good value for m, which is the number of most recent measurements used at each time instant, we run the MUPF algorithm for each object. Fig. 5 displays how the localization errors vary with different values of m in the range from 1 to L. The figure is for the data collected in the real experiments. The results of the simulated data, which were similar, have been omitted for clarity. Since the localization errors do not decrease significantly for $m > L/2$, $m = L/2$ has been chosen.

Fig. 6 shows the performance achieved with the simulated measurements. The results are grouped in 6 *experiments*, according to the exploited set of measurements, from (a) to (f) (see x axis of the plot). For each experiment, we show the average localization errors on 10 trials for all 6 object

[1]DOI:10.5281/zenodo.45493.

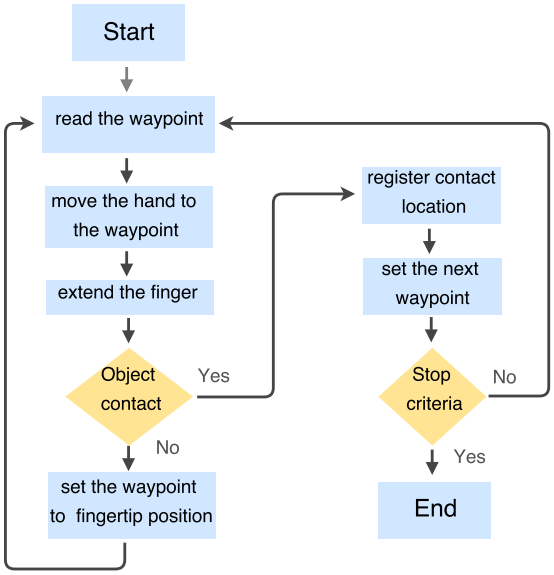

Fig. 4. A flow chart showing the object-surface sampling.

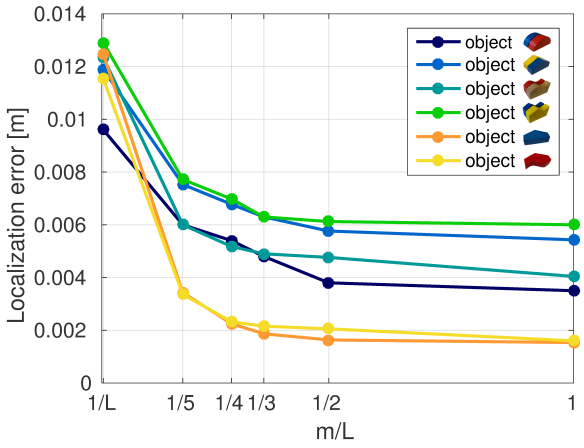

Fig. 5. The localization errors obtained with real measurements with different values of m, from 1 to L. For the sake of clarity, the results from the simulated data have not been plotted as it exhibits a similar trend.

TABLE II

PARAMETERS SET FOR THE MUPF IN SIMULATION

Q	diag([10^{-4}, 10^{-4}, 10^{-4}, 10^{-2} 10^{-2}, 10^{-2}]) [m], [rad]
σ_p	10^{-4} [m]
P_0	diag([0.04, 0.04, 0.04, π^2, $(\pi/2)^2$, π^2]) [m], [rad]
N	700
m	$L/2$

models. If we consider the first experiment on the left, where the set of measurements (a) is used, the object $\hat{k}$ selected as solution by virtue of (2) is the one featured by the lowest localization error. Fig. 6 shows how the object is correctly recognized (i.e. $\hat{k} = k^*$), since the bar corresponding to object (a) (dark blue) provides the smallest localization error for that experiment. With the same procedure, we can infer that all objects are successfully recognized.

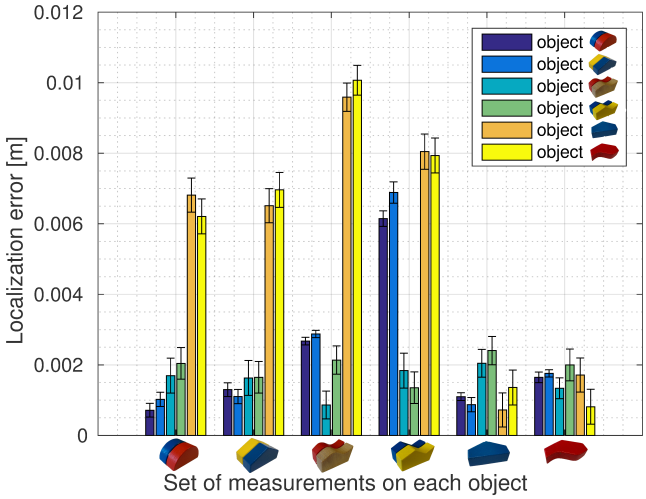

Fig. 6. MUPF performance with simulated measurements. The results are grouped according to the exploited set of measurements: from the left to the right, the results obtained for the measurements sampled on object (a) to object (f) are shown. For each experiment, the average localization errors on 10 trials obtained for all the object models are shown.

B. Experimental Results

Before showing the performance achieved using the real measurements, we provide a synthetic experiment to point out, from a quantitative viewpoint, that the task at hand is indeed challenging. The results of the experiment are shown in Fig. 7. The test consists of calculating the localization error of three different object models: (a), (b) and (c), using the set of *real* measurements (b). More precisely, the three profiles depicted in Fig. 7 represent how the localization error varies as the object models slide along the y axis of the frame attached to the object basis. Therefore, Fig. 7 reports the localization error versus the y displacement: a displacement equal to 0 represents the correct pose for the object (b), with respect to the set of measurements (b). By observing the trend of the localization errors, we can see how the localization error for object (b) is minimum for a displacement equal to 0, that is in fact the correct pose. However, object (a) and (c) provide an even lower localization error in correspondence of small displacements along y. This fact highlights how the similarity of objects and the noisy nature of the measurements could lead to wrong recognitions.

We discuss hereinafter the performance achieved with real data. The MUPF parameters used for the experimental tests are provided in Table III. The parameters have been chosen by taking into account considerations similar to those explained in Section IV-A. In particular, covariances Q and σ_p are tuned differently in order to take into account the measurement noise of the real data. The value of m is determined as described in the previous section, see Fig. 5.

Fig. 8 shows the results of the real experiments, which can be interpreted similarly to the data of Fig. 6. Two main differences can be noticed by comparing Fig. 6 and Fig. 8. First, the measurement noise causes higher aver-

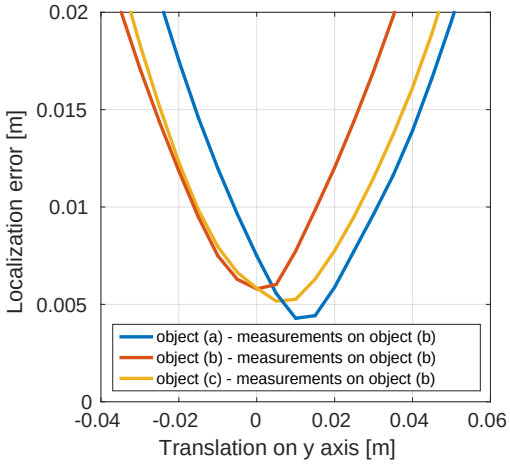

Fig. 7. Synthetic test showing the challenging nature of tactile recognition problem. We compute the localization errors with respect to the set of real measurements (b) and three object models: (a), (b), and (c). Each model is sliding along the y axis of the ground frame. Object (b) results in the lowest error at zero displacement, whereas, notably, object (a) and (c) give lower values for small nonzero displacements.

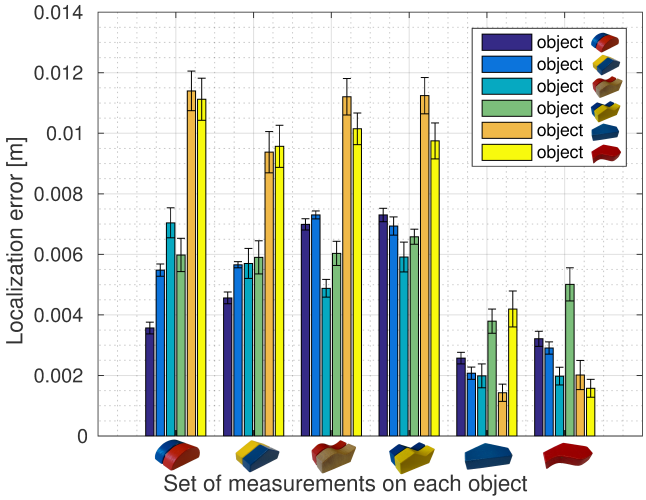

Fig. 8. MUPF performance with real measurements. The results are grouped according to the exploited set of measurements: from the left to the right, the results obtained with the measurements collected by touching object (a) to object (f) are shown. For each experiment, the average localization errors on 10 trials obtained for all the object models are shown.

TABLE III
PARAMETERS SET FOR THE MUPF IN REAL EXPERIMENTS

Q	diag([8 10^{-6}, 8 10^{-6}, 8 10^{-6}, 8 10^{-4} 8 10^{-4}, 8 10^{-4}]) [m,rad]
σ_p	4 10^{-4} [m]
P_0	diag([0.04, 0.04, 0.04, π^2, $(\pi/2)^2$, π^2]) [m], [rad]
N	1200
m	$L/2$

age localization errors. Therefore we manage to correctly recognize only 4 objects out of 6 in the real scenario, compared with the 100% overall classification score achieved in simulation. In particular, when the MUPF is executed using set of measurements (b), the solution $\hat{k}$ is given by object (a) and, analogously, when measurements belong to object (d), $\hat{k}$ comes out to be object (c). However, we could reasonably consider these two misclassifications acceptable, considering the high level of similarity between the pairs of objects and the noise in the measurements. In addition, the limited resolution of the tactile sensor and the size of the fingertip (approximately 6 $\times$ 6 [mm^2]) allow only a coarse discrimination of the shape of the object and hide finer details. It is expected that the performance of the recognition would increase using a smaller fingertip or sensors with higher resolution. Given these limitations, however, the carried out experiments demonstrate that the proposed algorithm achieves good performance.

V. CONCLUSIONS

In this paper, we propose a novel approach to the problem of tactile object recognition. We addressed the problem as a tactile localization on multiple objects and used a nonlinear filtering algorithm, named Memory Unscented Particle Filter, capable of recognizing objects by exploiting only contact point measurements. The effectiveness of our approach is demonstrated both in simulation and with a real robot.

The promising results presented in this paper encourage us to keep working on tactile object recognition. A possible future application is to test our approach on a larger set of objects, also considering different properties. For example, the model could be extended by including local features, such as surface classification (e.g. local curvature, edge, corners) or material properties (e.g. stiffness, texture). A further extension of the work we presented consists of taking advantage of a more complex exploration strategy for data collection, by using multiple fingers at the same time. In fact, the exploitation of the knowledge of which finger has caused each tactile measurement could be very powerful and considerably improve the performance of our approach.

ACKNOWLEDGMENT

This research has received funding from the European Union Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 610967 (TACMAN).

REFERENCES

[1] P. Jenmalm, S. Dahlstedt, and R. S. Johansson, "Visual and tactile information about object-curvature control fingertip forces and grasp kinematics in human dexterous manipulation," *Journal of Neurophysiology*, vol. 84, no. 6, pp. 2984 – 2997, 2000.

[2] A. Talati, F. J. Valero-Cuevas, and J. Hirsch, "Visual and tactile guidance of dexterous manipulation tasks: an FMRI study 1, 2," *Perceptual and motor skills*, vol. 101, no. 1, pp. 317 – 334, 2005.

[3] N. Jamali, M. Maggiali, F. Giovannini, G. Metta, and L. Natale, "A new design of a fingertip for the iCub hand," in *28th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 2705 – 2710, Hamburg, Germany.

[4] A. Schmitz, P. Maiolino, M. Maggiali, L. Natale, G. Cannata, and G. Metta, "Methods and technologies for the implementation of large-scale robot tactile sensors," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 389 – 400, 2011.

[5] N. Wettels, V. J. Santos, R. S. Johansson, and G. E. Loeb, "Biomimetic tactile sensor array," *Advanced Robotics*, vol. 22, no. 8, pp. 829 – 849, 2008.

[6] N. Jamali and C. Sammut, "Majority voting: Material classification by tactile sensing using surface texture," *IEEE Transaction on Robotics*, vol. 27, no. 3, pp. 508 – 521, 2011.

[7] S. Decherchi, P. Gastaldo, R. S. Dahiya, M. Valle, and R. Zunino, "Tactile-data classification of contact materials using computational intelligence," in *25th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012, pp. 635 – 639.

[8] H. Liu, X. Song, J. Bimbo, L. Seneviratne, and K. Althoefer, "Surface material recognition through haptic exploration using an intelligent contact sensing finger," in *25th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012, pp. 52 – 57.

[9] R. S. Fearing and T. O. Binford, "Using a cylindrical tactile sensor for determining curvature," in *IEEE Transaction Robotics and Automation*, vol. 7, no. 6, 1991, pp. 806 – 817.

[10] R. A. Russell and S. Parkinson, "Sensing surface shape by touch," in *10th IEEE International Conference on Robotics and Automation (ICRA)*, 1993, pp. 423 – 428.

[11] M. Charlebois, K. Gupta, and S. Payandeh, "Shape description of curved surfaces from contact sensing using surface normals," in *International Journal of Robotics Research*, vol. 18, no. 8, 1999, pp. 779 – 787.

[12] P. K. Allen and K. S. Roberts, "Haptic object recognition using a multifingered dextrous hand," in *6th IEEE International Conference on Robotics and Automation (ICRA)*, 1989, pp. 342 – 347.

[13] P. K. Allen and P. Michelman, "Acquisition and interpretation of 3-d sensor data from touch," in *IEEE Transaction on Robotics and Automation*, vol. 6, no. 4, 1990, pp. 397 – 404.

[14] M. Charlebois, K. Gupta, and S. Payandeh, "Shape description of general, curved surfaces using tactile sensing and surface normal information," in *14th IEEE International Conference on Robotics and Automation (ICRA)*, 1997, pp. 2819 – 2824.

[15] H. Liu, X. Song, T. Nanayakkara, L. D. Seneviratne, and K. Althoefer, "A computationally fast algorithm for local contact shape and pose classification using a tactile array sensor," in *29th IEEE International Conference on Robotics and Automation (ICRA)*, 2012, pp. 1410 – 1415.

[16] A. R. Jimnez, A. S. Soembagijo, D. Reynaerts, H. Van Brussel, R. Ceres, and J. L. Pons, "Featureless classification of tactile contacts in a gripper using neural networks," in *Sensors and Actuators A, Physical*, vol. 62, no. 1, 1997, pp. 488 – 1491.

[17] G. Vezzani, U. Pattacini, G. Battistelli, L. Chisci, and L. Natale, "Memory unscented particle filter for 6-DOF tactile localization," *preprint available at arXiv:1607.02757*, 2016.

[18] R. Van Der Merwe, A. Doucet, N. De Freitas, and E. Wan, "The unscented particle filter," in *Advances in Neural Information Processing Systems 13*, 2001, pp. 584 – 590.

[19] A. Doucet, N. De Freitas, and N. Gordon, "An introduction to sequential Monte Carlo methods," in *Sequential Monte Carlo methods in practice*, Springer, Ed., 2001, pp. 3 – 14.

[20] D. Simon, *Optimal State Estimation*. Hoboken, New Jersey: Wiley, 2006.

[21] A. Petrovskaya and O. Khatib, "Global localization of objects via touch," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 569 – 585, 2011.

[22] S. Saha, Y. Boers, H. Driessen, P. K. Mandal, and A. Bagchi, "Particle based MAP state estimation: A comparison," in *12th International Conference on Information Fusion (FUSION)*, 2009, pp. 278 – 283, Seattle, USA.

[23] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The iCub humanoid robot: an open platform for research in embodied cognition." *Proc. 8th Work. Perform. Metrics Intell. Syst.*, pp. 50–56, 2008.