Artificial Tactile Perception for Surface Identification using a Triple Axis Accelerometer Probe

Patrick Dallaire, Daniel Emond, Philippe Giguere and Brahim Chaib-Draa

Abstract— **In recent years, autonomous robots have been increasingly deployed in unknown environments. In order to cope with the unknown, the capability to train autonomously the perception model of an environment is highly desirable. By developing proper sensing technology, this task can be significantly facilitated. In this paper, we explore the problem of artificial tactile perception, aimed at surface identification. To this end, we introduce a simple tactile probe based upon triple axis accelerometers. This tactile probe was tested on a large collection (28) of flat surfaces, using a controlled test bed. In a first set of experiments, we demonstrated the discrimination capabilities of the probe, by achieving a surface recognition rate of 96.7% with 1 second of data, using a Support Vector Machine classifier. We also demonstrate that similar results can be achieved without the need for ground truth or the actual number of surfaces using Dirichlet process mixture models, a Bayesian nonparametric approach. These two experiments indicate that tactile sensing is, thus, a potentially viable solution for autonomous surface identification.**

I. Introduction

Tactile sensing is a fundamental mechanism which allows animals to interact with their environment, perform object identification and enables them to better understand their surroundings. In order to replicate a tactile sensing mechanism on a robot, several obstacles must be overcome. In a sense, this explains the slow development of tactile sensing technology when compared to the vision sensing technology [1]. Nonetheless, over the years, many types of tactile probes were developed to mimic the sense of touch, either for texture recognition or form perception.

In this paper, we present an improved tactile probe that employs an accelerometer as a transducer. Tactile probes have the capacity to gather information related to the mechanical properties of surfaces. These could be the elasticity, the hardness, or simply the fine surface texture often present on many different objects. As this information is not available to vision system, tactile-based systems might offer a more robust way to object identification, since they often differ widely in material composition.

Although shape estimation is an important part of tactile perception, we focus strictly on surface recognition in this paper. Surface recognition performed through the use of a contact sensor has a tremendous potential. Indeed, tactile sensing is, by its very nature, immune to numerous problems that plague vision-based sensing, such as illumination changes, occlusion, or the high-dimensionality output (millions of pixels) of cameras. Consequently, tactile perception could be used in challenging environments, where vision systems are difficult to operate. In outdoor settings for example, large and varying illumination changes are present, complicating the use of vision.

The information captured by a tactile probe can be used to train an artificial perception system. If the data labels are known, supervised learning techniques can be used. If the labels are unknown, one can employ an unsupervised learning method to perform the same task.

The rest of this work is organized as follows. Section II discusses other designs of surface sensing mechanisms. Section III describes our tactile probe and the data gathering process. Section IV describes the supervised classification method employed to qualify the performance of the tactile probe. Section V presents an unsupervised learning method used to learn autonomously to differentiate different surfaces, without the need to specify the number of classes. Finally, section VI concludes this paper and presents future works.

II. Previous Work

Numerous tactile sensing technologies have been proposed throughout the years. They are in the likes of artificial whisker [2][3][4][5], array of whiskers [6][7], or artificial antennas [8]. These devices can generate surface profile, perform rudimentary object recognition or provide distance estimation. Even finger-like devices [9] were developed for the purpose of enabling manipulators with slip conditions detection, as well as providing some form of texture sensing.

Schultz *et al.* [10] explored the use of power spectral analysis on strain gauges signals from friction-induced vibrations picked up by whiskers, in order to differentiate between smooth and rough textures. Still, no formal classification results were produced. Hipp *et al.* [11] presented more extensive results regarding texture classification with actuated whiskers. The classification, using multidimensional Gaussian density estimators, achieved a success rate of 39% for eight different grades of sand paper. Fend *et al.* [12], using a microphone to record the vibrations induced in genuine rat whiskers, conducted experiments over 11 surfaces. The power spectra of individual sweeps were smoothed and combined together to generate an average power spectrum. By comparing the Euclidean distances between these power curves, Fend and his colleagues concluded that texture identification could be improved by using all whiskers at the same time and by increasing the number of sweeps.

P. Dallaire dallaire@damas.ift.ulaval.ca
D. Emond daniel.emond@ulaval.ca
P. Giguere philippe.giguere@ift.ulaval.ca
B. Chaib-Draa chaib@damas.ift.ulaval.ca
The authors are all with the Department of Computer Science and Software Engineering, Laval University, Quebec City, QC G1V 0A6, Canada.

For their part, Roy *et al.* [13] presented a tactile probe, which tapped against a surface instead of rubbing it. The classification was achieved by comparing windowed power spectra of the sound generated by the impact. The classification success rate, over 6 surfaces, was 95%. However, it is unclear whether this approach can detect small asperities, an important cue in texture identification.

Recently, de Boissieu *et al.* [14] produced a finger-like device sensitive to friction forces. The classification experiments were performed on 10 types of paper with a success rate around 70%. This is an impressive result considering the small differences in textures. Even so, the significant wear to the rubber tip formed during the experiments indicates a lack of physical robustness, significantly reducing its applicability to industrial applications.

III. Tactile Sensor and Acquisition System

A. Tactile Sensor Design

Our tactile probe goes beyond simply capturing surface textures, as was done previously [10][11][12]. Instead, the weight and inertia of the probe's stylus acts as a probing mechanism. The dynamic behavior of the probe will depend highly on the viscoelastic properties of the surface, as well as its texture. For example, surfaces such as rubber or textile cannot induce high-frequency components in the accelerations, in contrast with hard surfaces like steel.

The tactile sensor we present in this paper is an improved version of an earlier prototype [15]. The newer version consists in a 2.38 *mm* diameter standard steel stylus attached on a triple axis digital MEMS accelerometer (ADXL345 from Analog Devices). Both are attached on a mounting bracket, which has a 45 degree angle relative to the surface to be probed, as depicted in Fig. 1. The accelerometer is placed near the tip of the probe to ensure maximum vibration capture. The combined weight of the stylus, the accelerometer and the mounting bracket is approximately 16 grams. The mounting bracket was fixed to a turntable with a flexible attachment made of thick leather. This allows the probe to rotate freely in a plane perpendicular to the X axis, while granting small rotations in other directions. This permits the capture of accelerations in all three axes. Moreover, the flexible attachment reduces the mechanical coupling between the turntable and the mounting bracket.

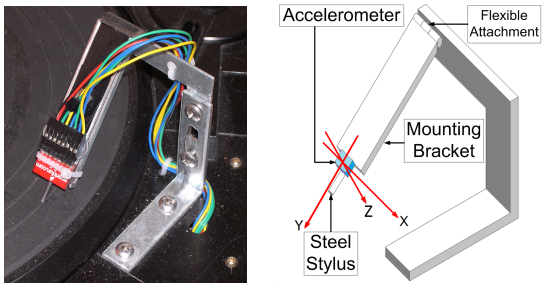

Fig. 1. Design of the tactile probe with a triple axis accelerometer.

The passive tactile probe gathers information about a surface through a rubbing action, while the surface is rotated by the turntable. To minimize tangential forces on the stylus' tip, the tactile probe is oriented in such a way that the X axis of the accelerometer is parallel to a line going from the tip to the rotation center of the turntable. A fine tip was selected, as to be able to track fine surface asperities. The speed of the surface was kept constant by the turntable. For each data collection, the Revolution per Minute (RPM) was maintained to 45.0 and measured by a digital tachometer (TC811B, Reliability Direct). The distance between the center of the turntable and the tip was 125.0 *mm*, giving a constant surface speed of 58.9 *cm/s* at the stylus' tip.

(a) Linoleum (b) Leatherette (c) Grass carpet

Fig. 2. Picture of 3 discs used in the surface identification experiments. The complete test set was comprised of 28 discs.

Data collection from the accelerometer device was performed via a Serial Peripheral Interface Bus (SPI) to Universal Serial Bus (USB) adapter (SUB-20-B from Dimax.) The accelerometer was configured to collect information at a data rate of 800*Hz*. The accelerometer data resolution was set to 3.9 milli-G, with a maximum range of $\pm$5G.

B. Data Set Gathering

The surfaces in the test set were made into 12-inches discs, as shown Fig. 2, so they can be placed on the turntable. In total, 28 different surfaces were used to make these discs. This large number of surfaces was selected to thoroughly test the discrimination capabilities of the probe. The probe support had adjustable height to maintain a probe angle of 45 degree, regardless of a discs' thickness. Combined with the controlled speed of the turntable, this ensured that the data gathering process was kept similar between discs.

Two data sets were collected at the same rotation speed. The first data set, $\mathbf{S}_{initial}$, consisted of approximately 20 minutes of recording for each disc. A second data set, $\mathbf{S}_{valid}$, was collected later to evaluate the impact on the mechanical manipulation between disc changes on the probe's measurements. It contained the equivalent of 1 minute of recording, again for each of the 28 surfaces.

C. Selected Signal Features

The raw data from all three accelerometers' measurements was divided into 3-by-W sub-matrices $\mathbf{a}_t$, with W being the width of the time window. The following seven features were extracted from each row of $\mathbf{a}_t$:

- f_1: variance,
- f_2: skewness,
- f_3: kurtosis,

- f_4: fifth moment,
- f_5: sum of the variation over time:

$$\sum_{t=1}^{W-1}(|a(t)-a(t+1)|)$$

- f_6: number of times 20 uniformly separated thresholds are crossed, and
- f_7: sum of higher half of amplitude spectrum.

Note that these features are a subset of the features employed in [15], originally adapted from [16]. The mean values of $\mathbf{a}_t$ were not used, avoiding the issue of accelerometer drift or probe angle impacting the classification. After extraction from a data set, each feature was normalized with a factor g_i so as to have a standard deviation of one. In total, this yielded a 21 dimensional data point $\mathbf{x}_i$ per time-window $\mathbf{a}_t$, when all three axes, i.e. the rows $\mathbf{a}_{1t}$, $\mathbf{a}_{2t}$, and $\mathbf{a}_{3t}$ were used:

$$\mathbf{x}_i = [g_1 f_1(\mathbf{a}_{1t}), g_2 f(\mathbf{a}_{1t}), \ldots, g_{21} f_7(\mathbf{a}_{3t})]. \quad (1)$$

From observation, we noted that some of these features f_i had non-Gaussian distributions. This was not an issue for supervised learning with an appropriate classifier, such as a Support Vector Machine (see Section IV). However, a subsequent feature linearization step, described in Section V-C, was required for clustering.

IV. Supervised Classification

A. Classifier: Support Vector Machine

We quantified the surface identification performance of our tactile probe by performing supervised learning on data sets $\mathbf{S}_{initial}$ and $\mathbf{S}_{valid}$. Classification was performed using Support Vector Machines (SVM) from the LIBSVM library [17]. This is in contrast with our previous results [15], where classification was performed with an artificial neural network. We selected a Radial Basis Function (RBF) kernel:

$$K_{kernel}(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma||\mathbf{x}_i - \mathbf{x}_j||^2), \gamma > 0.$$

The best values for γ and the regularization term C of the SVM were found using stratified n-fold cross-validation.

B. Classification Results

The data set $\mathbf{S}_{initial}$ was split in half, with the odd samples placed in $\mathbf{S}_{train}$ and the even samples placed in $\mathbf{S}_{test}$. This ensured that samples in both sets were equally spread over time: some features f_i extracted from soft surfaces varied over time, as these surfaces experienced some wear. For time-window size $W = 800$ (corresponding to 1 second of data at 800 *Hz*) and using 50% of the data set, all 28 surfaces were identified correctly (100.0%) in $\mathbf{S}_{test}$. The SVM trained on $\mathbf{S}_{train}$ was then used to classify the validation test set $\mathbf{S}_{valid}$ collected two weeks later. The overall classification rate was lower (96.7%), with the corresponding confusion matrix shown in Fig. 3. These results point towards some possible drift in the probe's response over long periods of time. We are further investigating this matter.

C. Relevance of using 3 axis

Our previous tactile probe design [15] used a single-axis accelerometer. Using a 3-axis accelerometer allowed for a greater amount of information to be collected per unit of time, as well as collecting information pertaining to the probe's lateral sliding motion. This translated in improved performances, particularly for small time-window sizes W. To confirm this, we performed the same experiments as in Section IV-B, but using only one of the 3 outputs of the accelerometer at a time (X, Y or Z), as well as a shorter time-window size $W = 200$. A coarse grid search with 50 % fold cross-validation was performed to find the C and γ values producing the highest testing classification rate, for each axis These parameters were used with the results reported in Table I. These results indicate that the X axis is the least informative for surface identification, while the Z axis is the most. This is not surprising, given that accelerations in the Z direction correspond to a rotation of the probe around the flexible attachment. Using all three axes surpassed any single-axis results, confirming the improvement over the original single axis accelerometer design.

TABLE I
Classification success rates for single and 3-axes modes, with a time window size $W = 200$.

Axis	$\mathbf{S}_{train}$	$\mathbf{S}_{test}$	$\mathbf{S}_{valid}$
X	88.0%	85.1%	66.0%
Y	92.9%	90.4%	73.5%
Z	96.6%	95.0%	81.4%
XYZ	99.6%	99.5%	92.9%

D. Supervised Classification Results as a Function of W

Using a smaller window size W implies that less information is available to perform classification. Assuming that each feature f_i behaves as an estimator, the standard deviation of the distributions associated with each feature f_i should be proportional to $\frac{1}{\sqrt{W}}$. The estimated Bayes error for normal distributions can be found using the erf function:

$$G_{rate} = \frac{1 + erf(\frac{A}{\sqrt{W}})}{2}. \quad (2)$$

The curves plotted in Fig. 4 correspond to the success rates, as a function of time window size W. A proper value of A in Eq. 2 was selected to match the top curve. The relative closeness of this match indicates that, overall, the distributions in feature space become wider as W is decreased. Moreover, this test demonstrates that surfaces can be quickly identified: more than 50 % of the samples in the validation test set are correctly identified with as little as 15 *milliseconds* of data.

V. Unsupervised Classification

One of our long-term objectives is to achieve autonomous learning of low-level perception models of the environment by mobile agents, by relying on unsupervised learning techniques. A number of unsupervised learning techniques have been proposed, such as k-means or Normalized-Cut [18].

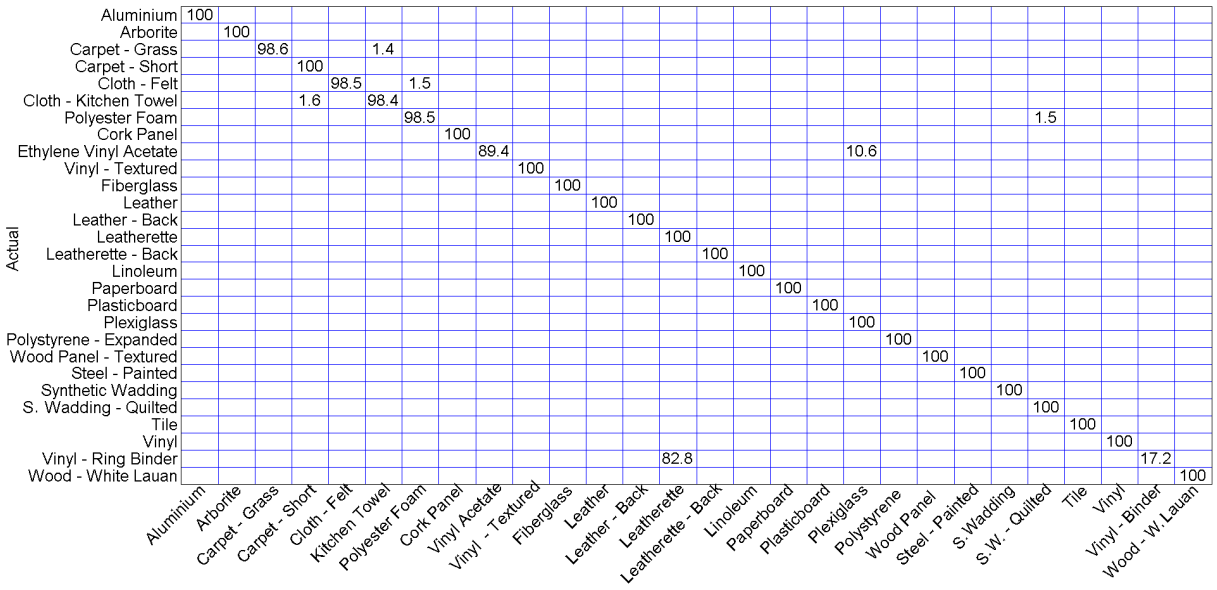

Fig. 3. Confusion matrix for the validation test set $\mathbf{S}_{valid}$ using time-window size of $W = 800$ samples.

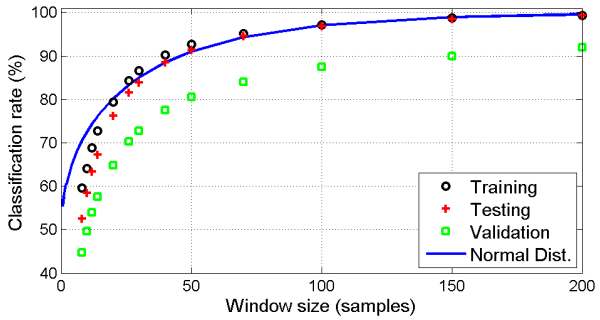

Fig. 4. Variation of the classification success rate, as a function of the window size W. The curve indicated as Normal Dist. correspond to Eq. 2.

A common drawback is the need to provide the number of clusters present in a data set. The method described in the following section, the Dirichlet Process Mixture Model [19], does not have such constraint. Instead, this method estimates the number of clusters, based on a prior distribution.

A. Clustering: Dirichlet Process Mixture Models

The fundamental aspect of the Dirichlet process mixture models (DPMMs) is the Dirichlet process (DP). The DP is a stochastic process, used in Bayesian nonparametric models of data, and is a distribution over distributions. It is parameterized by a base distribution G_0, which can be seen as a prior guess, and a concentration parameter α, determining the variance of the distribution. Using this distribution on the parameters of mixture models leads to the following specification of DPMMs:

$$\begin{aligned} G \mid \alpha, G_0 &\sim \mathrm{DP}(\alpha, G_0) \\ \theta_i \mid G &\sim G \\ \mathbf{x}_i \mid \theta_i &\sim F(\theta_i) \end{aligned} \tag{3}$$

where F is the data distribution parameterized by θ. The distribution G, corresponding to the unknown mixture distribution, is discrete with probability one and infinite dimensional, providing DPMMs with countably infinite number of mixture components.

The discreteness and clustering properties of DPs are essential for clustering with DPMMs. These properties become evident when the distribution G is integrated out to obtain a prior distribution on component parameters θ. In fact, as demonstrated in [20], the prior probability for an observation $\mathbf{x}_i$ to join an existing class or to create a new one is:

$$\begin{aligned} p(c_i = j \mid c_{1:i-1}) &= \frac{n_{-i,j}}{i - 1 + \alpha} \\ p(c_i \neq j \text{ for all } j < i) &= \frac{\alpha}{i - 1 + \alpha} \end{aligned} \tag{4}$$

where c_i is the class of $\mathbf{x}_i$ and $n_{-i,j}$ denotes the number of data belonging to class j, when excluding the i^{th} observation. Eq. 4 ensures that large clusters tend to get bigger, since the probability to join a class is proportional to its size. On the other hand, the concentration parameter α, which controls the new class probability, influences the total number of clusters.

To complete the previous prior, we have to specify the data distribution F. Then, given a set of observations, along with DP parameters G_0 and α, the DPMM yields a posterior distribution over the component parameters of a mixture. Exact inference for this posterior is analytically intractable. However, it is possible to directly sample this posterior with Markov Chain Monte Carlo methods.

B. Dirichlet process mixture of Gaussians

The traditional approach to Gaussians mixture model is to fix the number of components *a priori*. However, from a Bayesian nonparametric perspective, it is not necessary to fix this number, since it can be automatically inferred from the data. Thus, the number of clusters in a data set can be estimated automatically. The DPMM, as shown in [21], can be used to produce an infinite Gaussian mixture, thus avoiding the finite representation. This model has been

further extended in [22], with vague priors over the model hyperparameters. This section introduces a model based on these previous approaches.

To define a Dirichlet process mixture of Gaussians (DP-MoG) from the general DPMM, we specify the distribution F in the feature space as Gaussian:

$$\mathbf{x}_i \mid \theta_i \sim \mathcal{N}(\boldsymbol{\mu}_{c_i}, S_{c_i}^{-1}) \tag{5}$$

where $\boldsymbol{\mu}$ is the mean and S is the inverse covariance matrix. These parameters should then be assigned a prior distribution G_0. Following the conditionally conjugate model of [22], this G_0 prior is:

$$\boldsymbol{\mu}_j \mid \boldsymbol{\xi}, R \sim \mathcal{N}(\boldsymbol{\xi}, R^{-1}) \tag{6}$$

$$S_j \mid \beta, \Sigma \sim \mathcal{W}(\beta, (\beta\Sigma)^{-1}) \tag{7}$$

where $\mathcal{W}$ denotes the Wishart distribution with β degrees of freedom and scale matrix $(\beta\Sigma)^{-1}$. This prior determines the expected covariance of a cluster and its expected position in the feature space. Moreover, to complete the model, we also put an inverse-gamma prior on α:

$$\alpha^{-1} \sim \mathcal{G}(1/2, 1/2). \tag{8}$$

Since the base distribution G_0 might not be exact, vague hyperpriors are applied on hyperparameters $\boldsymbol{\xi}$, R, β and Σ to obtain a robust model learning G_0. This part assumes properly scaled data and is omitted for brevity [22].

C. Experiments with Unlabeled Data Sets

We tried to achieve autonomous learning of surfaces, with as little information as possible provided by a human. This task is significantly harder than the supervised learning experiments presented in Section IV. For the unsupervised learning experiments, the information provided was limited to which features to extract: no labels were given, and the true number of surfaces was unspecified. We removed features f_2, f_3 and f_4; their respective distribution had high kurtosis, violating the assumption that features are normally distributed. Transformations were applied on the remaining features because i) they are strictly positive and ii) their variance appeared proportional to the distance from the origin, giving the data sets a conical shape with apex on the origin. As a result, we used features $\sqrt{f_1}$, $\log(f_5)$, $\log(f_6)$ and $\log(f_7)$, after which we performed a Principal Component Analysis for dimensionality reduction. For the experiments, we generated data sets using time-windows size W of 200, 400, 800 and 1600.

We used the DPMoG model as a prior on Gaussian mixtures. For the learning part, we employed a Gibbs sampling approach based on auxiliary parameters to sample posterior mixtures [20]. Markov Chain Monte Carlo methods are often computationally expensive for large data sets. Hence, we reduced the size of all training sets to 50 samples $\mathbf{x}_i$ per surface.

We first evaluated the capacity to recover the number of distinct surfaces in our tactile data sets for various windows size W, by sampling the posterior distribution over the number of classes. Fig. 5 shows the marginal posterior distributions over K, with the tails truncated for visibility. If we only consider the largest window size $W = 1600$, the distribution gives highest probability to the true number of surfaces: 28. Using a smaller window size W yields worst estimates, but they still give good probability to the true value. This goes in line with the previous results of Section IV-D that showed that classification is increasingly harder with a smaller W.

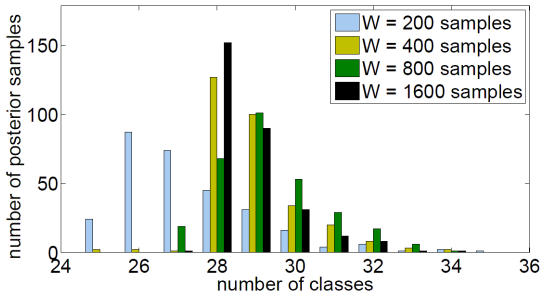

Fig. 5. Estimating the number of surfaces in an unlabeled data set. The most probable number of surfaces are 28 (W=1600), 29 (W=800), 28 (for W=400) and 26 (W=200), all close or equal to the true value of 28.

To measure the clustering accuracy, we defined a metric based on pairwise correct classification. It used the hidden labels to compute a dissimilarity measure with the true clusters. An error is made when 2 data points are assigned to the same cluster and should not be, or when they are in different clusters and should be in the same one. The resulting clustering accuracy is then 1 minus the error ratio. A score of 1 is only achieved for a perfect clustering. Fig. 6 shows the clustering accuracy obtained as the Markov chain evolved. The initial region corresponds to a burn-in period, and therefore should not be considered until stationarity is (probably) reached.

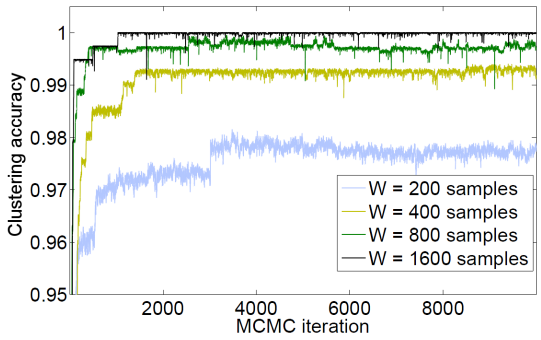

Fig. 6. Clustering accuracy on unlabeled data sets as a function of the MCMC iteration, for different time window sizes W. For these experiments, the true number of classes was unspecified.

We used the MAP estimate learned from the unlabeled data set to predict samples from new data sets $\mathbf{S}_i$. The data sets $\mathbf{S}_i$ were comprised of samples $\mathbf{x}_i$ immediately following the unlabeled training set, with 2800 samples $\mathbf{x}_i$ in each of them. To determine the class, we computed the likelihood of each Gaussian component, one for each class. The label assigned to a sample $\mathbf{x}_i$ corresponded to the class with the

highest likelihood. Classification results with this method are presented in Table II. As can be seen from these results, the Gaussian components found using the DPMoG method can be employed successfully to classify unseen labels.

TABLE II

MAP PREDICTION RESULTS ON FIVE DATA SETS $\mathbf{S}_i$ OF UNSEEN SAMPLES, WITH THE TRAINED DPMOG MODEL.

Test set	$\mathbf{S}_1$	$\mathbf{S}_2$	$\mathbf{S}_3$	$\mathbf{S}_4$	$\mathbf{S}_5$
Success rate	99.96%	99.96%	99.39%	99.50%	98.61%

D. Comparison with other Unsupervised Learning Methods

We compared the DPMoG results with two other unsupervised learning algorithms: k-means and Expectation-Maximization for Gaussian mixtures. Contrary to the DPMoG we employed, these two algorithms required the true number of classes to be provided. Each method was ran 10,000 times, and only the best model was kept for comparison purposes. Table III reports the clustering accuracies for all three methods. The averaged clustering accuracies from Fig. 6, excluding the burn-in period, compares favorably with these two methods. This is despite the fact that DPMoG had to estimate the number of clusters in the data set.

TABLE III

COMPARISON OF CLUSTERING ACCURACIES BETWEEN DIFFERENT UNSUPERVISED LEARNING METHODS.

Time-window	$W = 200$	$W = 400$	$W = 800$	$W = 1600$
k-means	0.9783	0.9877	0.9947	0.9961
EM	**0.9789**	0.9807	0.9915	0.9931
DPMoG	0.9776	**0.9924**	**0.9974**	**0.9999**

VI. CONCLUSION AND FUTURE WORKS

In this paper, we presented an improved version of a tactile probe, employing a triple axis accelerometer fixed near the tip of the stylus. Surface classification was based on seven features extracted from the accelerometer's measurements, in the time and frequency domains. The good discrimination capabilities of our tactile probe was established by using a Support Vector Machine classifier, on a data set comprising 28 different surfaces. We evaluated the impact of using 3 axes compared to 1, as well as the impact of the window size W on the classification. Finally, we demonstrated that information captured by this tactile probe can be easily used in the context of unsupervised learning, even when the number of surfaces in the data set is unknown. This was accomplished through the use of a Bayesian nonparametric modeling named Dirichlet process mixture. Overall, these experiments demonstrated that tactile sensing has the potential for surface identification by autonomous agents.

As part of future works, we are currently investigating the impact of the tactile probe's material on its sensitivity to surfaces. Our current probe design calls for a stiff material, but this has not been properly justified. Also, the experiments were conducted at a single velocity (58.9 *cm/s*). In our case, this might not be a major issue, since we assume that we can always control the velocity of the tactile probe, However, it would be important to quantify this impact. Finally, we are looking into how this rich tactile information can be exploited in the context of object recognition.

REFERENCES

[1] Mark R. Cutkosky, Robert D. Howe, and William R. Provancher. Force and tactile sensors. In Bruno Siciliano and Oussama Khatib, editors, *Springer Handbook of Robotics*, pages 455–476. Springer Berlin Heidelberg, 2008.

[2] D. Jung and A. Zelinsky. Whisker based mobile robot navigation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 2, pages 497–504 vol.2, Nov 1996.

[3] R.A. Russell. Using tactile whiskers to measure surface contours. In *IEEE Transactions on Robotics and Automation*, 1992.

[4] G.R. Scholz and C.D. Rahn. Profile sensing with an actuated whisker. *IEEE Trans. on Robotics and Automation*, 20(1):124–127, Feb. 2004.

[5] T.N. Clements and C.D. Rahn. Three-dimensional contact imaging with an actuated whisker. *IEEE Trans. on Robotics*, 22(4):844–848, 2006.

[6] R. A. Russell and J. A. Wijaya. Recognising and manipulating objects using data from a whisker sensor array. *Robotica*, 23(5):653–664, 2005.

[7] M. J. Pearson, A. G. Pipe, C. Melhuish, B. Mitchinson, and T. J. Prescott. Whiskerbot: A robotic active touch system modeled on the rat whisker sensory system. *Adaptive Behavior - Animals, Animats, Software Agents, Robots, Adaptive Systems*, 15(3):223–240, 2007.

[8] Noah J. Cowan, Emily J. Ma, Mark Cutkosky, and Robert J. Full. A biologically inspired passive antenna for steering control of a running robot. In *In International Symposium on Robotics Research*, pages 541–550. Springer, 2003.

[9] R.D. Howe and M.R. Cutkosky. Sensing skin acceleration for slip and texture perception. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 145–150, May 1989.

[10] A.E. Schultz, J.H. Solomon, M.A. Peshkin, and M.J. Hartmann. Multifunctional whisker arrays for distance detection, terrain mapping, and object feature extraction. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2588–2593, April 2005.

[11] J. Hipp, E. Arabzadeh, J. Conradt, E. Zorzin, C. Kayser, M. Diamond, and P. Knig. Texture signals in whisker vibrations. *Journal of Neurophysiology*, 95(3):1792–1799, 02 2006.

[12] M. Fend, S. Bovet, H. Yokoi, and R. Pfeifer. An active artificial whisker array for texture discrimination. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 2, pages 1044–1049 vol.2, Oct. 2003.

[13] Nicholas Roy, Gregory Dudek, and Paul Freedman. Surface sensing and classification for efficient mobile robot navigation. In *IEEE Transactions on Robotics and Automation*, 1996.

[14] F. de Boissieu, C. Godin, B. Guilhamat, D. David, C. Serviere, and D. Baudois. Tactile texture recognition with a 3-axial force MEMS integrated artificial finger. In *Proceedings of Robotics: Science and Systems*, Seattle, USA, June 2009.

[15] P. Giguere and G. Dudek. Surface identification using simple contact dynamics for mobile robots. In *IEEE International Conference on Robotics and Automation (ICRA)*, Kobe, Japan, May 2009.

[16] C. Weiss, H. Frohlich, and A. Zell. Vibration-based terrain classification using support vector machines. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4429–4434, Oct. 2006.

[17] Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[18] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2000.

[19] Y.W. Teh. Dirichlet Process. *Submitted to Encyclopedia of Machine Learning*, 2007.

[20] R.M. Neal. Markov chain sampling methods for Dirichlet process mixture models. *Journal of computational and graphical statistics*, 9(2):249–265, 2000.

[21] C.E. Rasmussen. The infinite Gaussian mixture model. *Advances in neural information processing systems*, 12:554–560, 2000.

[22] D.I.D. Görür. *Nonparametric Bayesian Discrete Latent Variable Models for Unsupervised Learning*. PhD thesis, Max Planck Inst. for Biological Cybernetics, 2007.