Embodied hyperacuity from Bayesian perception: Shape and position discrimination with an iCub fingertip sensor

Nathan F. Lepora, Uriel Martinez-Hernandez, Hector Barron-Gonzalez, Mat Evans, Giorgio Metta, Tony J. Prescott

Abstract—Recent advances in modeling animal perception have motivated an approach of Bayesian perception applied to biomimetic robots. This study presents an initial application of Bayesian perception on an iCub fingertip sensor mounted on a dedicated positioning robot. We systematically probed the test system with five cylindrical stimuli offset by a range of positions relative to the fingertip. Testing the real-time speed and accuracy of shape and position discrimination, we achieved sub-millimeter accuracy with just a few taps. This result is apparently the first explicit demonstration of perceptual hyperacuity in robot touch, in that object positions are perceived more accurately than the taxel spacing. We also found substantial performance gains when the fingertip can reposition itself to avoid poor perceptual locations, which indicates that improved robot perception could mimic active perception in animals.

I. INTRODUCTION

A casual observer can tell that the perceptual abilities of the most advanced robots lag far behind those of animals, which is a major bottleneck for enabling robots to interact fully with their environments. Meanwhile, in neuroscience, two lines of investigation are converging on an understanding of animal perception as statistically optimal sensory processing. First, perception is considered as Bayesian inference from noisy and ambiguous sensations [1], [2]. Second, the decisions resulting from this inference are considered optimal in terms of minimizing the costs of making mistakes plus the costs of waiting to gather more sensory data for improved accuracy [3], [4]. These inferential and reactive aspects of perception have been embodied in robots with biomimetic whisker sensors, using Bayesian sequential analysis to accumulate evidence for competing perceptual hypotheses until reaching a decision threshold on the beliefs [5], [6].

This study presents an initial implementation of this biologically-inspired framework for perception in a robot using a biomimetic touch sensor based on the human fingertip [7] that was built for the iCub humanoid robot [8]. In doing so, we show that the approach applied previously to touch with whisker sensors applies also to the new problem domain of fingertip sensing, with potential applications in grasping and manipulation. To provide a standard set-up for systematic testing, the fingertip was mounted on a dedicated positioning robot (Fig. 1A) to test shape and position classification of range of cylindrical test objects (Fig. 1B). We emphasize that these are important percepts for the manipulation of objects; for example, shape discrimination is necessary to determine the grasp configuration for picking up an object, while position discrimination is necessary to appropriately position a manipulator onto an object.

There has recently been a surge in hardware development for robot touch sensing [9] accompanied by advances in tactile classification methods influenced by Bayesian methods from machine learning and probabilistic robotics [10], [11], [12], [13], [14], [15]. The approach to Bayesian perception in this study and related work [5], [6] differs by emphasizing the reactive aspects of decision making, rather than concentrating on the inference problem. In practise, this difference stems from the use of a threshold crossing rule on the beliefs for real-time decision making, rather than using a preset, fixed amount of data. In our opinion, utilizing the reactive aspects of perception is a necessary step to achieve robot perception comparable with that of animals.

Using Bayesian perception, we find that the fingertip sensor can achieve positional hyperacuity, in that the object position relative to the finger can be localized to within 25% of the physical taxel spacing (<1 mm mean error; 4 mm spacing). To the best of our knowledge, this result is the first explicit demonstration of hyperacuity in robot touch. We also find that the perceptual accuracy can be further improved by having the robot react to the sensory data during decision making. In particular, long reaction times indicate poor locations for perceiving the stimulus, which can be used to reactively reposition the robot for better perception.

This work was supported by EU Framework project EFAA (ICT270490). N. Lepora, U. Martinez-Hernandez, H. Barron-Gonzalez and T. Prescott are with the Adaptive Touch Laboratory at Sheffield, Department of Psychology, University of Sheffield, Western Bank, Sheffield S10 2TN, UK. Email: {n.lepora, uriel.martinez, h.barron, t.j.prescott}@sheffield.ac.uk. Giorgio Metta is with the Italian Institute of Technology (IIT), Italy. Email: giorgio.metta@iit.it
II. METHODS

A. Robot experiments with an iCub fingertip sensor

The iCub is a humanoid robot about the size of a 3 year old infant [8]. To enable the robot to grasp and manipulate objects, touch sensors were initially integrated into its hands [7] and then later a tactile skin onto its forearms and body [16]. The iCub fingertips are 14.5 mm long and 13 mm wide with a rounded shape that resembles a human fingertip. They consist of an inner support wrapped with a flexible printed circuit board (PCB) containing twelve conductive patches for the touch sensor ‘taxels’. The PCB is covered first with a ∼2 mm layer of non-conductive soft silicone foam and then with a thin layer of conductive silicone rubber. The PCB and silicone layers together comprise a capacitive touch sensor that detects pressure via the capacitance change due to compressing the foam between the conductive layers.

The present experiments test the capabilities of the iCub fingertip sensor mounted on an xy-positioning robot. This robot can move the sensor over a horizontal plane in a highly controlled and repeatable manner onto various test stimuli to 50 µm accuracy (and was developed originally for testing tactile sensors based on rodent whiskers [17]). The fingertip was angled so the rod axis lay across the fingertip (vertically across the taxels in Fig. 2), and moved so that the rod initially contacted the fingertip at its base and finally contacted only the tip. In each case, an x-range of 16 mm was considered. This gave 320 taps per object at increments of 0.05 mm. Each tap of the fingertip against the object resulted in a 1 sec time series of pressure readings for all 12 taxels covering the fingertip (Figs 2B-E). This data was sampled at 50 Hz, giving 50 samples per taxel per tap. One training and one test set was collected for each of the five rods and these data used for off-line Monte Carlo validation of the Bayesian perception.

B. Bayesian sequential analysis for perception

Here we adopt a statistical method for tactile perception based on Bayesian sequential analysis, which is closely related to a leading model of perceptual decision making in neuroscience [3], [4], [5]. The approach has two aspects: an inference part based on Bayesian filtering and a decision-making part that reacts to the inference. In this section, we describe the method of Bayesian sequential analysis, and in the following section describe how it is applied to our data.

The inference part of this Bayesian model of tactile perception uses Bayes’ rule to update the posterior probabilities \( P(c_n | z_t) \), or ‘beliefs’, for \( N \) perceptual classes \( c_n \) after receiving new measurement data \( z_t \). This is achieved by using Bayes’ rule recursively as a Bayesian filter, with the prior given by the posterior on the previous time-step, which is combined with the present likelihood \( P(z_t | c_n) \) to give

\[
P(c_n | z_t) = \frac{P(z_t | c_n)P(c_n | z_{t-1})}{P(z_t | z_{t-1})}.
\]  

(1)

For the measurement model of the likelihoods, we assume the probability distributions \( P(z_t | c_n) \) are identically distributed and independent of time \( t \) and measurement history (hence...
we drop \( z_{1:t-1} \) from the posteriors); the construction from the training data is described below. The marginal probabilities are conditioned on the preceding measurement and are found from summing over all \( N \) classes

\[
P(z_t| z_{t-1}) = \sum_{n=1}^{N} P(z_t|c_n)P(c_n| z_{t-1})
\]

(2)

to give properly normalized posteriors \( \sum_{t=1}^{N} P(c_n| z_t) = 1 \).

Taking a sequence of measurements \( z_1, \ldots, z_t \) gives a sequence of posteriors \( P(c_n| z_1), \ldots, P(c_n| z_t) \) for each class, which are calculated by iterating over the relations (1,2) starting from uniform priors \( P(c_n) = 1/N \).

The decision-making part of Bayesian sequential analysis uses a threshold-crossing decision rule on the belief, which allows the perception to be reactive to the quality of the test data and can enforce optimal balance of speed versus accuracy in the decision making [3], [4], [5], [18]. At least one posterior must cross a preset probability threshold to trigger the decision, after which the maximal a posteriori (MAP) estimate is taken for the perceptual class:

\[
\text{if any } P(c_n| z_t) > \theta \text{ then } c = \arg \max_{c_n \in C} P(c_n| z_t).
\]

(3)

This decision rule implements free-response perception, in that the decision occurs at a time determined by the decision-maker. Previous robot work has demonstrated superior performance compared to fixed reaction time methods [5], [6].

In this study, we also consider a modification of standard sequential analysis in which null decisions are allowed if a preset ‘boredom’ deadline \( D \) is reached before crossing the probability threshold:

\[
\text{if } t = D \text{ then } c = \emptyset.
\]

(4)

In a previous study with robot whisker sensors, we interpreted a null decision as referring to forced versus unforced choices [6]. Here we make a different interpretation: if a null decision is reported, then it signals that the present location gives a poor perception of the object and the robot should move its sensor elsewhere (Fig. 3). This is important, because past null decisions should then be included in the reaction times (rather than considering only reaction times for decisions that were made, as done previously [6]).

C. Bayesian perception applied to robot touch

The above Bayesian sequential analysis approach for robot perception was then applied to the fingertip data.

The perceptual classes denoted the cylinder being contacted and the position of the fingertip relative to the cylinder. To provide sufficient training data for each position class, we separated the data into 16 groups of 20 taps, each spanning 1 mm (delineated by the tick marks on Fig. 2A). Using all taps in a position class aided the robustness of the classifier, by training across the range of positions. This gave 16 position classes and 5 shape classes, or 80 classes in total.

The training data was then used to construct a measurement model of the likelihoods based on a ‘bag of measurements’ histogram approach. The histogram of pressure readings \( s \) of the training data for percept class \( c_n \) defines a sample distribution

\[
P_k(s|c_n) = \frac{h_k(s)}{\sum_s h_k(s)},
\]

(5)

where \( h_k(s) \) is the number of measurement values \( s \) in the histogram occurring for taxel \( k \). Altogether, 1000 samples \( s \) were used for each taxel of each class, corresponding to 20 taps of 50 samples each. We extracted 12 histograms (one for each taxel) for each of the 80 training classes. Note that the pressure values must be binned to construct the histogram, which was implemented by uniformly partitioning each pressure range over 100 bins (see also [5], [19]).

The measurement model then gives the likelihoods for a test tap (denoted \( z_t \) in the Bayesian update rule) from the geometric mean over the sample distributions (5) evaluated over all samples \( s_j \) in the test tap

\[
P(z_t|c_n) = \left( \prod_{j=1}^{\text{tap}_j} \prod_{k=1}^{\text{tax}_k} P_k(s_j|c_n) \right)^{\text{ta}_j \text{tap}_j}
\]

(6)

where \( J = 50 \) and \( K = 12 \) are the time samples per tap and the number of taxels respectively. This model assumes a bag of measurements in which all samples are treated as independent and identically distributed for each taxel. The geometric mean prevents the product (6) from producing vanishingly small likelihoods by ensuring that the probabilities remain almost invariant for large sample numbers [5].

The test data was then used to construct sequences of taps to assess the speed and accuracy of the robot perception. A Monte Carlo procedure was employed to emulate real-time performance. Each test instance corresponded a tap sequence of known shape and position class constructed by drawing taps in a random order with replacement from the appropriate class of test data. Then the perception algorithm processed this test sequence until making a decision, after which the shape and position errors and reaction time (in taps) were reported. By performing many test instances (typically 2500
Fig. 4. Decision errors plotted against probability threshold. Panel A shows the mean absolute error for estimating rod diameter and panel B shows the mean absolute error for finger position. We distinguish fixed perception with no deadline (black lines) from when moves were allowed after a null decision upon reaching a deadline (colored lines).

Fig. 5. Decision errors plotted against reaction times. Plots are as in Fig. 4, apart from plotting the mean absolute errors against mean reaction time (by relating belief threshold to reaction time via Fig. 7B). We interpret these results as the mean real-time performance of accuracy versus reaction time.

per belief threshold) with randomly chosen test classes, the distributions of shape and position errors and reaction times could then be characterized. In general, the error distributions were near Gaussian centered on zero while the reaction time distributions were skewed as in Fig. 7. Suitable statistics to summarize the real-time performance are the mean absolute errors and the mean reaction times, which are displayed in the following results.

III. RESULTS

A. Initial observations

Fingertip pressure data was collected for five steel rods (diameters 4 mm, 6 mm, 8 mm, 10 mm and 12 mm) by passing the fingertip slowly over each rod tapping at a constant rate of once per second (see methods; configuration shown in Fig. 1). For the example data shown in Fig. 2A (4 mm diameter rod), at the start of the movement only the bottom taxels at the fingertip base are in contact (Fig. 2B), then the lower-center taxels contact and the bottom taxels detach (Fig. 2C), then the upper-center taxels contact and the lower-center taxels detach (Fig. 2D) and finally only the top-most taxels at the tip of the finger contact (Fig. 2E).

Notice that the pattern of taxel pressures depends on both the curvature of the surface being contacted and the position of the fingertip over the object, permitting the simultaneous perception of object shape and location. Another key observation is that each taxel has a receptive field size of about 5 mm (100 taps on Fig. 2a), with small pressure readings at the edges of this field that increase to peak in the center. The centers of the taxels are spaced about every 4 mm, so the receptive fields for different taxels overlap to permit tactile coverage across the entire fingertip. Position classes were considered here at 1 mm intervals along the 16 mm range (tickmarks on Fig. 2A), at about 25% of the taxel spacing.

B. Bayesian tactile perception from a fixed position class

The first validation study considered tactile perception in which the test data was drawn from a single test shape and position class. Real-time performance was assessed by repeatedly drawing random sequences of test data to give accurate statistics on the mean absolute shape and position errors and mean reaction times (Methods, Sec. II-C), with the test shapes and positions also randomly varied. We used a Bayesian method of free-response perception in which the decision was reported after a variable number of taps set...
by when at least one of the posterior probabilities reached a preset belief threshold. The posteriors were updated using Bayes’ rule recurrently as a Bayesian filter with the priors given by the posteriors on the previous tap.

Average decision errors for perceived shape and position were examined over belief thresholds ranging from 0.05 to 0.99999 (Fig. 4; black curves). Statistically robust estimates of the mean errors were found by averaging the absolute classification errors over many test instances. These mean errors decreased steadily with belief threshold, reaching a minimum of about 2.2 mm for shape and 0.6 mm for position for thresholds above 0.9 (Fig. 4; black curves).

In accordance with the perception being freely responsive, the number of taps to reach a decision had a reaction time distribution (Fig. 7A). Increasing the belief threshold increased the mean reaction time (Fig. 7B; black curve) and decreased the decision errors. Treating this threshold as an implicit parameter gave a direct plot of decision error against mean reaction time (Fig. 5; black curves), in which for each threshold the mean decision accuracy from Fig. 4 was plotted against the mean reaction time from Fig. 7B. The resulting error plots decreased with longer reaction times, reaching their minimum values after about 5–10 taps.

The mean decision errors and reaction times were non-uniform across finger position and rod diameter, showing significant variance with test class (Fig. 8). In particular, neighboring position classes varied substantially in their classification properties, indicating a lack of overall structure in the variation of these errors. That being said, there was maybe a slight tendency for finger placements between 2–7 mm to yield better position accuracy than elsewhere.

In summary, for the above Bayesian perception, the decisions of rod diameter reached an average error of about one shape class after a few taps (e.g. the 8 mm rod was commonly mistaken for a 6 mm or 10 mm rod but not for the 4 mm or 12 mm rod). Meanwhile, the decisions on position were less than the 1 mm width of a perceptual class, indicating that about half found the correct position and the other half were commonly about 1 mm out. These position errors were ≤25% of the 4 mm spacing between taxels. In consequence, the position discrimination performs at sub-taxel resolution, and thus constitutes an embodiment of hyperacuity in a biomimetic robot with brain-based perception.

C. Bayesian perception with random move after deadline

The second validation study considered tactile perception against a ‘boredom’ deadline, after which a null decision is reported and the fingertip sensor moved randomly to retry decision making at another location. The same method of Bayesian perception was used as for the fixed-position study from Sec. III-B, but with a deadline number of taps for the null decision (Fig. 3). This method can still be considered Bayesian sequential analysis, even though the standard optimality theorems [18] would need re-examining with respect to the costs of null outcomes. This study interprets this cost as from the accumulated delay in reaction time for the previous null results before finally making a decision.

We used random moves to re-position the fingertip because of the lack of structure in the error dependency on rod shape and location (Fig. 8). Discarding the data before the random move also simplified the validation. In general, the number of null decisions (move requests) increased as the belief threshold became closer to one and also as the deadline became shorter (Fig. 6). This behavior was expected, because more decisions should time out as the thresholds become harder to reach within a maximum allowed decision time.

The principal effect of allowing repositioning of the fingertip is improved decision accuracy (Fig. 4; colored lines) and reaction times (Fig. 6B; colored lines). The most dramatic improvements were for the shortest considered deadline of 5 taps (Fig. 4; blue line), which reduced the average shape errors to about 1.4 mm and the average position errors to about 0.1 mm at the highest belief thresholds (compared with 2.1 mm and 0.6 mm originally). This was accompanied with an improvement in reaction time, provided the belief threshold was not too large. Both effects combined together to give a pronounced improvement in the speed-accuracy plots of decision making (Fig. 5; colored lines).

These improvements in reaction speed and decision accuracy are a consequence of rejecting positions where the perception is poor or ambiguous. Although there is a cost to the reaction time in losing the data from the previous position, this can potentially be outweighed by the benefits of a location with faster and more accurate decisions. Therefore, Bayesian perception can give a natural way to assess the quality of a location for perceptual decisions, which can be utilized to re-position the sensor for improved perception.
IV. CONCLUSIONS

We have implemented a brain-based framework of Bayesian perception in a robot using a biomimetic sensor based on the human fingertip. For this study, the touch sensor was mounted on a dedicated positioning robot to test its capabilities in a controlled setting (Fig. 1A). For test stimuli, we used five cylindrical steel rods with diameters ranging from 4 mm to 12 mm (Fig. 1B) offset by a 16 mm range of positions relative to the fingertip. The robot gathered sensory data by tapping its fingertip against the curved surface. These experiments were suitable for characterizing both the shape and position-sensing capabilities of the iCub fingertip, and inform about the real-time capabilities of Bayesian perception with this tactile sensor.

Using a simple ‘fixed-position’ instantiation of Bayesian perception, the fingertip sensor could reach mean accuracies of ∼2 mm for the rod diameter and below 1 mm for relative fingertip-object position with just a few taps within a single position class. Given the taxel spacing was around 4 mm, the fingertip could achieve positional accuracy with mean errors less than 25% of the spatial resolution between taxels. In the biological literature, achieving a perceptual accuracy that is better than the sensor resolution is known as hyperacuity.

Progressing to a more complex reactive implementation of Bayesian perception, we found substantial improvements in shape and position discrimination when the robot could respond to the sensory data by randomly re-positioning the fingertip after a deadline. Hence, this movement strategy permitted an even greater degree of hyperacuity in the positional perception. The reaction time in Bayesian perception assessed the quality of a location for perceptual discrimination, which could be utilized to re-position the sensor for improved performance. Several other lines of enquiry remain open for improving robot perception. For example, the reactive perception considered here used random moves rather than purposively trying to improve the decision accuracy. In our opinion, significant gains could arise from allowing these movements to be guided with active perception [20], [21], such as by deciding the best move to disambiguate competing hypotheses during the perceptual process.

To the best of our knowledge, we have presented the first explicit demonstration of hyperacuity in robot touch. Hyperacuity is a generic aspect of animal perception, and while best known in vision also occurs for tactile perception. For example, Braille reading involves perceiving spatial patterns of a finer resolution than the spacing between touch receptors in the fingertips [22]. The hyperacuity in the present robot experiment was a consequence of the method of Bayesian perception being capable of utilizing the graded population response across taxels to interpolate between taxel positions. It arose from both the implementation of Bayesian perception and the morphology of the artificial fingertip sensor, coupled together by the statistics of the stimuli, and hence we refer to it as embodied hyperacuity. We expect that tactile hyperacuity will be a general phenomenon when Bayesian perception is applied to positioning tasks beyond those considered here, and could be a key consideration in the design of tactile sensors (as it is in the composition of their biological counterparts).

REFERENCES