2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Methods for Safe Human-Robot-Interaction Using Capacitive Tactile Proximity Sensors

Stefan Escaida Navarro, Maximiliano Marufo, Yitao Ding, Stephan Puls, Dirk Göger, Björn Hein and Heinz Wörn

***Abstract*— In this paper we base upon capacitive tactile proximity sensor modules developed in a previous work to demonstrate applications for safe human-robot-interaction. Arranged as a matrix, the modules can be used to model events in the near proximity of the robot surface, closing the near field perception gap in robotics. The central application investigated here is object tracking. Several results are shown: the tracking of two human hands as well as the handling of occlusions and the prediction of collision for object trajectories. These results are important for novel pretouch- and touch-based human-robot interaction strategies and for assessing and implementing safety capabilities with these sensor systems.**

I. Introduction

Research in robotics has long been focused on a system that on the one hand is able to interact autonomously with the environment and on the other hand guarantees a safe interaction with the human. It is a current research goal that the functionality and efficiency of the robotic system is hindered to the least possible degree by safety concerns. For their perception of the environment and implementation of safety features robots of any kind, humanoids, service robots and industrial robots are equipped with many sensors. In this paper we present methods for modular tactile proximity sensors that strengthen the perception in the near proximity of the robot and its endeffector or gripper. This type of sensing is one of the most neglected ones but is of utmost importance for safety because it closes a still existing perception gap.

Typical sensors employed for implementing a safety aware robotic system are 2D and 3D camera systems as well as tactile sensors of various designs and working principles. But, these sensors are inadequate for the detection and modeling of events in the near proximity of the robot: the perception of camera systems is encumbered by occlusions, extraneous light influence and shadows. Using cameras to increase safety is therefore problematic. Tactile perception on the other hand can only detect the presence of obstacles when the contact has already occurred. To react appropriately to unforeseen events it is necessary to restrict the robot's velocity at the cost of the time efficiency in which tasks can be executed.

To close the above mentioned perception gap a dual-mode tactile proximity sensor with a capacitive working principle was developed in a previous work [1]. Sensor identification for the tactile mode and further sensor designs where presented in [2]. Of special importance to the present paper is a modular unit that can be used to create an intelligent skin for a robot arm. Using an array of 3×16 of these modules, as shown in Fig. 1, the paper has contributions in two main aspects: we show how tracking in the proximity mode is enabling for 2-hand interaction and that a robust halt- and contact-prediction can be implemented. These features can be very useful for a robot system to avoid collisions by correcting its path according to the predicted situation, instead of an emergency halt. At a signal processing level we also present an improvement to the demodulation algorithms that leads to better framerates compared to the previous design. The remainder of the paper is structured as follows: in section II we review related work from the field, in section III we discuss the working principle of the sensor, the design of the modules used and the algorithms we used and developed. The results of our experiments for halt- and contact-prediction are shown in sec. IV and finally, conclusions are given in sec. V.

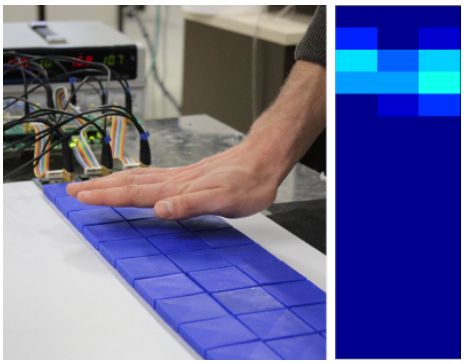


Fig. 1. A matrix arrangement of capacitive tactile proximity sensors modules can be used to increase safety and for combined pretouch- and touch-based interaction of the human with the robot. Here an arrangement of 3×16 is shown including the electronics for signal processing in the background.

II. Related Work

Many physical effects have been utilized to build tactile sensors such as capacitive [1], resistive [3] or optical approaches [4]. Proximity sensing is commonly implemented via optical [5] or capacitive measurement [6]. However, most of current approaches are only able to sense either

All authors are affiliated with the Institute for Process Control and Robotics, Karlsruhe Institute of Technology, Germany. Emails: {stefan.navarro, stephan.puls, bjoern.hein, heinz.woern}@kit.edu, {uyddx, yitao.ding}@student.kit.edu, dirk@goeger.net

tactile influence *or* object proximity but not both in a single sensor. Optical proximity sensors have the advantages of good precision and low circuit complexity [5]. As shown in [7], when integrating optical proximity sensors into a robotic gripper *spatial resolution* can be achieved, as well as for 360° sensing around a mobile platform [5]. Similar to cameras, optical sensing principles are prone to changing illumination, visibility conditions, mirroring effects and translucence. Thus, their application in safety critical environments is problematic. Capacitive proximity sensors show strong non-linear signal behavior, due to dependency on object size, its material and its grounding state. Therefore, the identification of object-sensor-distance is challenging. On the other hand, using multiple sensors can help to improve distance estimations. Also, capacitive sensors are insensible to optical conditions and can determine additional object properties, e.g. fluid level in a translucent object. Due to this robustness, the scope of application is broader than with optical counterparts.

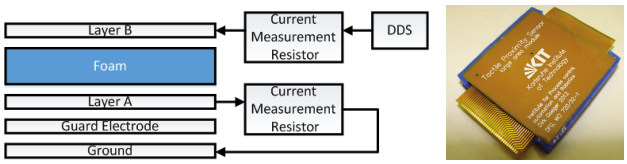
Application wise, proximity sensors can support and allow for robust gripping tasks. In the works [7] and [6], IR and capacitive sensing principles are respectively exploited to take object-gripper-distance into account. Thus, robust gripping can be achieved without prior exact object location information. Tactile information can be used to evaluate the stability of a grip, as shown in [8], [9]. Moreover, slip detection of a grip can be accomplished using tactile sensing, e.g. [10]. In [11] robust tracking of objects within the proximity of a gripper-like end-effector that has four IR proximity sensors is demonstrated. Multi-object tracking seems feasible in principle, but is restricted by the number of sensors.

Another field of application is the interaction of robots with their environment. Especially the avoidance of collisions during robot motion is important. In [12], distance sensing laser range finders are used to avoid dynamic obstacles with a mobile platform. Similarily the proximity sensor network of [5] is used for the same task. In [13], concepts for safeguarding human-robot cooperation in commercial robotics applications are presented, which are based on capacitive proximity sensors and haptic input devices. This bandwith of applications for proximity and tactile sensing clearly demonstrates the usefullness of integrated multi-modal sensors.

III. System Description

A. Sensor Platform

The basic working principle for tactile proximity sensor illustrated in fig. 2(a) was first presented in [1]. In tactile mode the capacitance between layer A and layer B is measured. The guard electrode is switched to ground to keep layer A capacitive coupled to ground. A force on the sensor decreases the distance between both layers and therefore changes the capacitance. In proximity mode layer B and the environment are capacitive coupled. The guard electrode is switched to the same potential as layer B to shield it from the ground layer. Objects in proximity will increase



(a) Schematic structure of the tactile proximity sensor.

(b) Big area module.

Fig. 2. Overview of the tactile proximity sensor module used in this work.

the coupling. In order to measure capacitance a direct digital synthesizer (DDS) generates a signal of approx. $100kHz$ which is routed to the capacitor plates. A simple measuring resistor in each circuit measures the current flow through the capacitors. The sensor modules shown in 2(b) used in this work were described in detail in our previous work [2]. Each module has a footprint of $4 \times 4cm$ and can detect objects in the proximity of about $10cm$. It is designed to be used for reveting a robot arm, up to 16 modules can be connected together to form a stripe that is connected to one channel on the signal processing board. Therefore the signal processing board with 10 channels is able to drive up to 10×16 sensor modules. Signal processing in the digital domain (filtering, demodulation) is further implemented on an FPGA-board. Enough modules were built to implement an array 3×16 sensors for this work.

1) Signal Demodulation: The amplitude of the measured current is proportional to the capacitance value. Therefore a demodulation algorithm needs to be implemented in the FPGA in order to obtain the amplitude value. The A/D-converter uses a sample rate of approx. $400kHz$, exactly four times higher than the signal frequency generated by the DDS. The current layout of the electronics does not allow the triggering of the A/D-converter, only specify its frequency, meaning the phase shift between sampling and signal generation is unknown. Therefore the old demodulation implementation (pictured by the blue part in Fig. 3) used a phase control algorithm to tune the phase to zero at the A/D-converter by manipulating the phase of the signal from the DDS. Up to 16 modules are multiplexed to a single A/D-converter and therefore the controller needed time to tune itself after each switching operation at the multiplexer. With this implementation a framerate approx. 10 fps was achieved. Furthemore, because of the signal processing path including filtering and communication via TCP/IP to Matlab a delay of about 0.5 seconds is introduced to the signal.

The relation of A/D conversion frequency to the signal frequency by exactly factor of 4 enables a new approach for demodulation (shown as the green part in Fig.: 3) using a simplified DFT in order to optain the amplitude value.

$$DFT : X(k) = \sum_{n=0}^{N-1} x(n) e^{-i2\pi kn/N} \qquad (1)$$

The DFT provides the amplitude independently from the phase of the signal. In this case we only need $X(k = 1)$ with $N = 4$ (the relating factor between sampling and signal

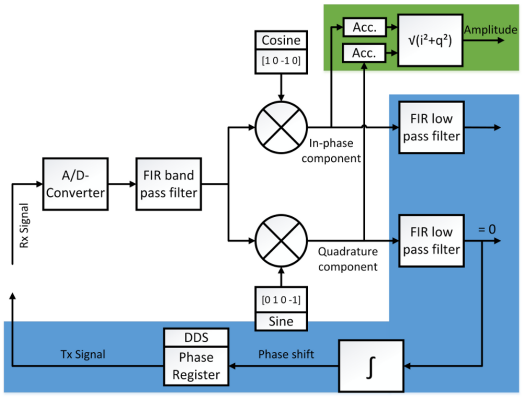


Fig. 3. Schematic structure of the signal demodulation path. The blue area is part of the first implementation, the green area is part of the new implementation.

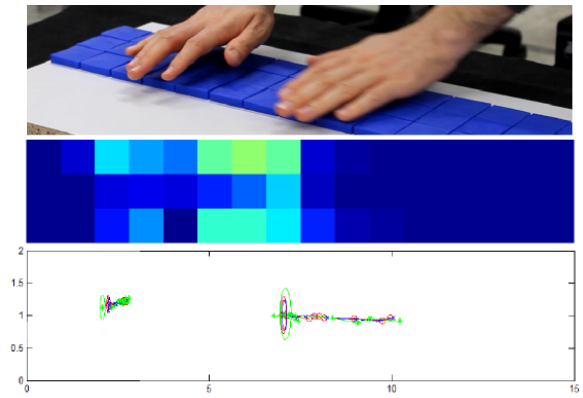


Fig. 4. Tracking enables to distinguish two hands even if the separation in sensor data is not evident.

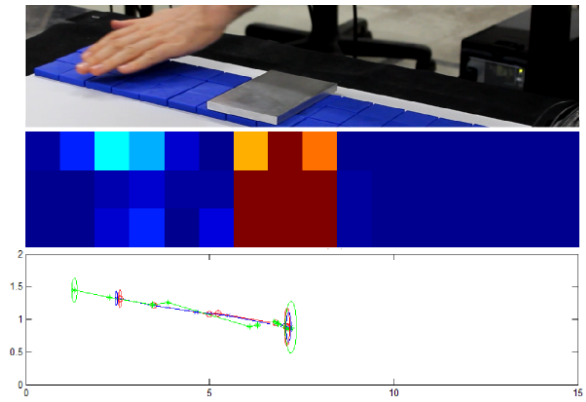


Fig. 5. Tracking enables to handle occlusions in the proximity sensing.

frequency) to be calculated which simplifies the calculation to:

$$e^{-i2\pi 0/4} = 1 \quad (2)$$

$$e^{-i2\pi 1/4} = i \quad (3)$$

$$e^{-i2\pi 2/4} = -1 \quad (4)$$

$$e^{-i2\pi 3/4} = -i \quad (5)$$

$$DFT:$$

$$X(1) = x(0) - x(3) + i(x(1) - x(4)) \quad (6)$$

The amplitude A is the absolute value of X(1):

$$A = abs(X(1)) = \sqrt{[x(0) - x(3)]^2 + [x(1) - x(4)]^2} \quad (7)$$

This can be easily implemented in the existing FPGA by using two additional multipliers. The new method avoids the use of a PI-controller for the phase and allows shorter waiting times after switching between each module. Depending on the configurations of the band- and low-pass filters a framerate of approx. 25 fps can be achieved.

B. Image Processing

As the data of the proximity-/tactile-sensor matrix corresponds to a two-dimensional planar image, we can analyze these images using moments up to the 2^{nd} order [14]. The two-dimensional $(p+q)^{th}$ order moment $m_{p,q}$ of an image is defined as the following double sum over all image pixels (x, y) and their values $f(x, y)$:

$$m_{p,q} = \sum_x \sum_y x^p y^q f(x, y) \qquad p, q \geq 0 \ . \quad (8)$$

The features used in this work are the area and the centroid of the foreground portion in the image.

C. Tracking

It is worth noting again that the proximity signal measured depends (at least) on size *and* distance of the object, making different sized objects indistiguishable to the sensor array if only a single snapshot is available. Also, spatial resolution is very low compared to vision sensors, making it essential to gain information from *time series* about the objets. A tracking scheme allows for smoothing of the noisy position and trajectories of objects being perceived. Higher-level features can be extracted from the trajectories to allow interaction with the robot, for instance through gestures. A robust estimation of the object's dynamics is necessary for the tasks of halt or contact prediction. When more than one object is present, the tracking algorithm allows to find the correct correspondence of object instances in case of occlusions, as illustrated by fig. 5. Also, when objects get closely together the correspondence of measurement on the sensor to the objects cannot be readily determined. Through tracking, the correspondence can be traced back to the time when the separation was evident as shown by fig. 4. For tracking in the proximity mode we implemented a Kalman-Filter as described in [15].

With the Kalman-Filter the current dynamics state of the object is estimated based on state history, a noisy dynamics model for the motion of the object combined with a noisy measurement taken at each sample time. In each step the predicted state of the object, according to the dynamics model and the previous state, is combined with the measurement such that the combined uncertainty inherited from prediction and measurement is minimized. In our implementation we

use the dynamics model for rigid body motion given by:

$$pos(t) = p + vt + \frac{1}{2}at^2 \tag{9}$$

At each time step the next position and velocity are predicted based on the previous values for position p, velocity v and acceleration a. The acceleration of the object is propagated as constant but is assumed to be noisy, leading to the uncertainty in the prediction. The modeled uncertainty in the acceleration should vary from scenario to scenario. The motion of human hands has high variance in its acceleration whereas a robot controlled movement is commonly subject to constant accelerations and target speeds. In this fact we see a chance for performing pretouch object categorization by analyzing its movement profile.

A measurement for the current object position is the weighted mean, according to eq. 8, of the sensor values of the sensors that are considered to be activated because of the object's presence. For a single object or multiple objects clearly separated this means all non-zero values of the corresponding region are included. For multiple objects that activate overlapping regions of sensors it has to be decided which measurements correspond to which object in order to measure the position. For the case of the human hands an object size of 3×4 sensor cells is assumed. A mask of the object size is applied along the Y axis (longer side of the sensor) of the foreground region, flush with the lower and upper boundary of the foreground respectively. Then a weighted centroid c_k is calculated from the mask as a position estimate. An example for two hand tracking is shown in Fig. 4. For this case of the rather narrow sensor arrangement this simple solution works more robustly than trying to find the region limits by means of the watershed-algorithm [16] or using gradient images. Also with our scheme, when objects (hands) completely overlap, the same measurements are assigned to both instances, which is a desirable behaviour. Nonetheless, for higher resolutions and a greater number of sensor modules more elaborate schemes and more features, such as the orientation of the shape, should be considered.

The correspondence of centroids c_k to the objects is calculated using the relative distance $rel_{i,k}$ of c_k to the position pos_i^{pred} of object O_i predicted by the Kalman-Filter:

$$rel_{i,k} = \frac{||pos_i^{pred} - c_k||}{\sum_j ||pos_i^{pred} - c_j||} \tag{10}$$

The centroid $c_l = argmin_{i,k}(rel_{i,k})$ with overall smallest distance $rel_{j,l}$ is assigned object O_j. This is done successively until all centroids are assigned. Multiple objects that from the sensing point of view are joined from the beginning cannot be recognized as such until separation is evident.

D. Proximity Mode Calibration

As stated previously, the proximity-signal is non-linearly related to the object's distance, material, size and electrical-state (floating, grounded). Therefore, we calibrated the proximity mode output of the sensor for the two types of objects used in this paper: a grounded aluminum cylinder and a human hand. The raw sensor outputs are offsetted for each sensor stripe so the first step is to get zero mean values after sensor initialization. Since the sensor signal is approximately exponentially related to the distance, in the next step the logarithm of the values is taken. These corrected values are then calibrated using a least-squares polynomial regression as seen in Fig. 6, good distance estimation through the polynomial expression is achieved. The calibrated sensor values are used for the experiments in sec IV.

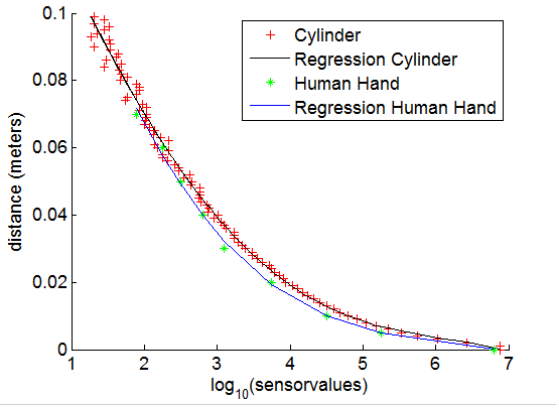


Fig. 6. Sample data and least-squares polynomial fits for calibrating sensor values to distance for a human hand and the aluminum cylinder.

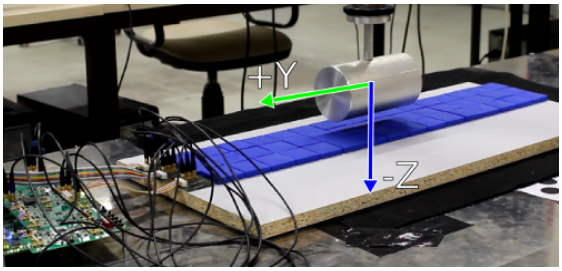


Fig. 7. Setup for the experiments of halt- and contact-prediction with an aluminum cylinder (length 15cm, diameter 8.5 cm and weight of 2kg).

For the cylinder samples at an interval of 1mm were taken and it is noticeable that with increasing distance the measurement becomes noisier. Therefore a next step in the tracking would be to consider this noise explicitly in the tracking for fine-tuning. A detection range of about $10cm$ is achievable for the hand and the cylinder.

IV. Experiments And Results

In addition to the qualitative results shown for the tracking of hands in the previous section, in this section two experiments regarding the halting prediction and contact prediction are shown. These experiments represent two possible scenarios: 1) an object is moving towards the sensor array mounted on a robot manipulator and 2) the arm itself is moving and the sensors detect some feature of the environment is approaching. The general setup used for the experiments is shown in Fig. 7. To avoid damaging the sensors by dragging them after contact has been established, both types of

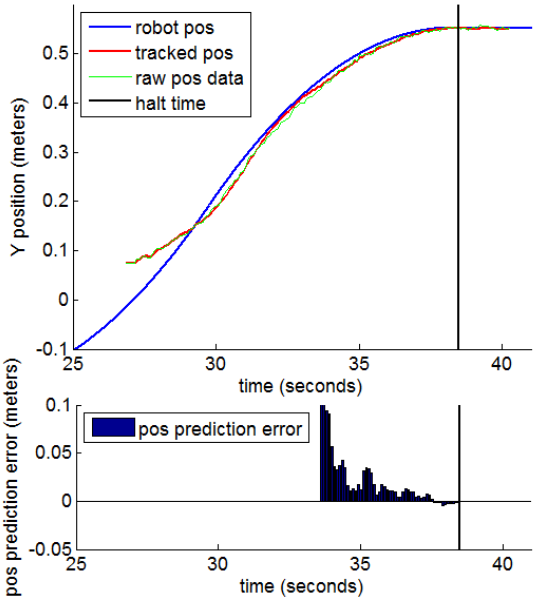


Fig. 8. Real and tracked movement (above) and halt position prediction error (below) in Y-direction of the aluminum cylinder.

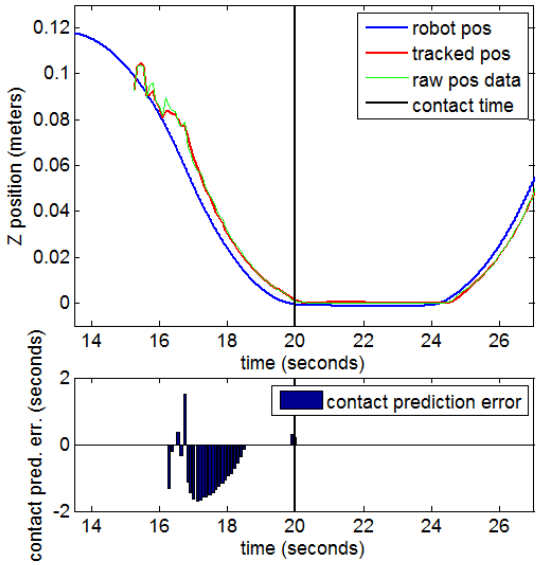


Fig. 9. Real and tracked movement (above) and contact time prediction error (below).

movement, parallel to the sensor (in Y) and perpendicular to the sensor (in Z), are evaluated independently. The foremost application that can be derived from these experiments is a smart collision avoidance strategy for robot manipulators equipped with tactile proximity sensors. To our knowledge no contact prediction using proximity sensing, evaluated against the same sensor in tactile mode has been presented by other authors so far. This type of evaluation is essential for assessing the extent of the safety awareness that is possible.

There is a time-delay due to signal processing as explained in III-A.1. At this point it has not been exactly determined and for this reason the robot control data of the object's motion is not considered for the quantitative evaluation.

1) Halt Position Prediction: For this experiment we consider the scenario of tracking an object moved by a robot arm parallel to the sensor surface. The robot is programmed to perform a trajectory of start and stopping constrained to a constant acceleration for starting and braking. Therefore, the acceleration variance (noise) in the dynamics model of the tracking algorithm is set very low. The robot trajectory is represented by the blue line in Fig. 8. The movement parameters of the braking part of the trajectory of the object are estimated using the tracking algorithm indicated by the red line in Fig. 8. Once a negative acceleration is detected, a prediction about the halting position of the movement is calculated according to the equations of motion (Eq. 9).

The actual final position measured as the mean of 20 sensor values after the halt of the motion is then used to estimate the precision of the prediction. At each sample time the difference of the actual and the predicted final position in Y is shown in Fig. 8. From the estimated trajectory in Fig. 8 it is recognizable that the measurement of the object position is quite precise. Shortly after the braking phase begins, the error in the halt position estimation drops below $5cm$, 4 seconds before halt it is below $2.5cm$. This result is satisfactory since the error is quite below the length of a sensor module ($4cm$).

2) Contact Time Prediction: In the event that contact with the robot is unavoidable, one of the main features that are necessary to assess the safety of the situation with proximity sensing is to be able to predict the movement parameters such as the velocity at the time of impact or the time of impact itself. How to detect *where* the robot will halt was evaluated in the previous experiment IV-.1. For detecting *when* impact will occur it is necessary to track the motion of the object perpendicular to the sensor, i.e. in Z direction. In our experiment the object movement starts outside the sensing range of the proximity sensors and moves along the Z axis in negative direction, shortly after entering the sensing range of the sensor the object motion brakes until finally contacting the sensor itself and triggering the tactile modality of the cells. The object's motion is represented in Fig. 9 where blue is the actual trajectory given by the robot control, green represents the raw measurements and red is the filtered measurement. The halting point for the object's motion was set slightly lower than the contact point with the sensor's surface, so that the tactile modality is activated and the motion has a crossing point for $Z = 0$.

It can be seen that the measurements and the tracking follow the trajectory of the object very closely. This also shows that a very precise calibration of sensor signal to distance is possible (see III-D). Fig. 9 shows the error in predicting the actual time of impact at each sample time with respect to the impact time measured with the tactile mode. The total braking time is about 3.25 seconds and the estimate is best at 1.5 seconds before impact. After 18.5

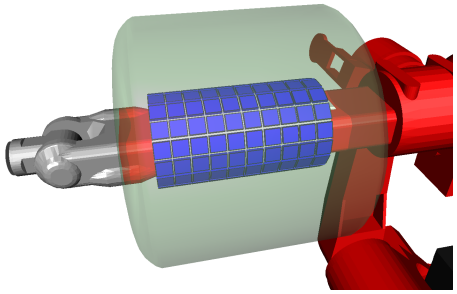


Fig. 10. A vision for the integration of the sensors on the exterior of a robotic manipulator.

seconds no more prediction solutions are available because the predicted motion is has no crossing point at $Z = 0$. Prediction of contact time satisfactory, but this experiment shows that even with precise measurements the object's distance the exact estimation of the movement's parameters is challenging. A basic safety reaction is an emergency stop as soon as the object enters the sensing range of the sensor. Using the contact time and position prediction presented here it is possible to implement more complex strategies that can actively avoid contact.

V. Conclusions and Future Work

In this paper we have shown applications for a matrix arrangement of modules of capacitive tactile proximity sensors. The sensing principle presents some challenges in relating the proximity signal to distance values, but nevertheless the scope of application is broader than optical counterparts. These sensors are able to close an existing near field perception gap in robotics. The challenge adressed in this paper is to model events in the near surroundings of the robot, namely tracking of objects and performing analysis of their motion for halt and contact prediction.

Firstly, we showed improvements on the previous signal processing design which results in higher framerates for the sensors which are beneficial for the prediction and tracking tasks. We have shown that the human hands can be tracked effectively, a combined touch- and proximity-based interaction should be easily be implemented with state of the art methods for gesture recognition. For object motion controlled by a robot we have shown that brake and contact prediction is possible with encouraging results. These results are a foundation for proximity and touch based safe human-robot-interaction. Prediction of contact has several applications for safety and can be used to implement collision avoiding strategies for preventing damage because of collision caused by robot motion. When the contact is intended, this can also be useful for proximity/touch based control of the robot.

For a future work we want investigate how to relate object's motion detected by the sensors to the type of object. Human hands motion is very dynamic whereas robot controlled motion are often constrained to constant acceleration and target speed. Then, indirectly regression models can be adjusted to better estimate the distance and size. Further steps are also the integration of the sensors into a gripper and onto a robot platform and its motion control, as seen in Fig. 10, as well as developing more applications, such as pretouch object exploration. In general the developed capabilities are aimed to be included in a *Skill* library so they can be used within complex execution and collaboration plans.

Acknowledgment

The authors would like to thank Mirko Kunze for the provided interpolator block in Simulink that implements the robot movement profiles used in our experiments. This work was funded by the FP7 EU-Project *SkillPro*.

References

[1] D. Göger, M. Blankertz, and H. Wörn, "A Tactile Proximity Sensor," in *IEEE Senors 2010*, 2010, pp. 589 –594.

[2] D. Göger, H. Alagi, and H. Wörn, "Tactile proximity sensors for robotic applications," in *International Conference on Industrial Technology (ICIT)*, 2013.

[3] K. Weiss, "Ein ortsauflösendes taktiles Sensorsystem für Mehrfinger-Greifer," Ph.D. dissertation, Universität Karlsruhe (TH), Institut für Prozessrechentechnik Automation und Robotik, 2006.

[4] S. De Rossi, N. Vitiello, T. Lenzi, R. Ronsse, B. Koopman, A. Persichetti, F. Giovacchini, F. Vecchi, A. Ijspeert, H. van der Kooij, and M. Carrozza, "Soft artificial tactile sensors for the measurement of human-robot interaction in the rehabilitation of the lower limb," in *Conf Proc IEEE Eng Med Biol Soc.*, 2010.

[5] K. Terada, Y. Suzuki, H. Hasegawa, S. Sone, A. Ming, M. Ishikawa, and M. Shimojo, "Development of omni-directional and fast-responsive net-structure proximity sensor," in *International Conference on Intelligent Robots and Systems*, 2011.

[6] B. Mayton, L. LeGrand, and J. Smith, "An electric field pretouch system for grasping and co-manipulation," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, may 2010, pp. 831 –838.

[7] H. Hasegawa, Y. Mizoguchi, K. Tadakuma, A. Ming, M. Ishikawa, and M. Shimojo, "Development of intelligent robot hand using proximity, contact and slip sensing," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, may 2010, pp. 777 –784.

[8] H. Dang, J. Weisz, and P. Allen, "Blind grasping: Stable robotic grasping using tactile feedback and hand kinematics," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, may 2011, pp. 5917 –5922.

[9] J. Schill, J. Laaksonen, M. Przybylski, V. Kyrki, T. Asfour, and R. Dillmann, "Learning continuous grasp stability for a humanoid robot hand based on tactile sensing," in *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS EMBS International Conference on*, june 2012, pp. 1901 –1906.

[10] J. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. Kuchenbecker, "Human-inspired robotic grasp control with tactile sensing," *Robotics, IEEE Transactions on*, vol. 27, no. 6, pp. 1067 –1079, dec. 2011.

[11] G. Petryk and M. Buehler, "Robust estimation of pre-contact object trajectories," in *In: Robot Control*, pp. 793–799.

[12] M. Göller, F. Steinhardt, T. Kerscher, J. Zöllner, and R. Dillmann, "Proactive avoidance of moving obstacles for a service robot utilizing a behavior-based control," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, oct. 2010, pp. 5984 – 5989.

[13] R. Koeppe, D. Engelhardt, A. Hagenauer, P. Heiligensetzer, B. Kneifel, A. Knipfer, and . K.Stoddard, "Robot-robot and human-robot cooperation in commercial robotics applications," in *Robotics Research*, ser. Springer Tracts in Advanced Robotics, P. Dario and R. Chatila, Eds. Springer Berlin Heidelberg, 2005, vol. 15, pp. 202–216.

[14] M.-K. Hu, "Visual Pattern Recognition by Moment Invariants," *IEEE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 1962.

[15] G. Welch and G. Bishop, "An Introduction to the Kalman Filter," UNC-CH Computer Science Technical Report 95-041, Tech. Rep., 1995.

[16] F. Meyer, "Topographic distance and watershed lines," *Signal Process.*, vol. 38, no. 1, pp. 113–125, July 1994.