2013 IEEE International Conference on Robotics and Automation (ICRA)
Karlsruhe, Germany, May 6-10, 2013

Dual-Mode Compliant Optical Tactile Sensor

Espen Knoop and Jonathan Rossiter

***Abstract*— Tactile force sensing and compliance are key elements of safe and natural-feeling human-robot interaction.**

We present an optical tactile sensor in the form of a compliant elastomer 'fingertip' tracked by a high-speed low-resolution image sensor with on-board signal processing. We propose a dual-mode bio-mimetic control loop, where in *reflex mode* the sensor sends fast reflexive action commands directly to actuators, bypassing the central controller to minimise reaction times. For higher-level interpretation, a slower *explore mode* enables more sophisticated processing of the sensory input by the central controller.

We demonstrate sensing of normal force in both modes of operation, showing that in *reflex mode* we are able to rapidly detect the presence of forces and compute an approximate magnitude estimate while in *explore mode* we are able to perform more accurate force measurements.

I. INTRODUCTION

Efficient human-robot interaction requires tactile force sensing. Soft and compliant robotic manipulators are safer, and also allow for more natural interaction through increased anthropomorphism.

Interaction with traditional, rigid robots carries a risk of the robot applying excessive force, and would also be perceived as mechanical and artificial. Interaction with soft and compliant robots would be perceived as far more natural and is therefore much preferable. Chorley et al. [1] argue that robot manipulators should mimic the human finger to gain the same desirable properties of compliance, conformability and mechanical strength. Any human-interacting robot will require extensive force sensing capabilities.

The way humans interact with objects through touch has multiple modes. Reflexes allow for rapid muscular reactions to stimuli, typically in combination with fast movement. We are also able to manipulate delicate objects — requiring precise estimation of gripping force — and detect very fine surface features.

Rather than trying to replicate the sensing range and resolution of the human finger, we have here focused on developing a simple optical tactile sensor capable of sensing forces up to 10 N, with a compliance and form factor mimicking the human finger.

We have developed a low-cost optical tactile force sensor, or 'robotic finger'. The fingertip is made of soft, compliant and transparent elastomer with an opaque rubber skin, and a dot pattern is applied to the inside skin surface. The application of a force to the fingertip causes it to deform and the dots to move. A camera chip is mounted behind the fingertip,

Department of Engineering Mathematics, Faculty of Engineering, University of Bristol, Bristol, BS8 1TR, espen.knoop@bris.ac.uk, jonathan.rossiter@bris.ac.uk

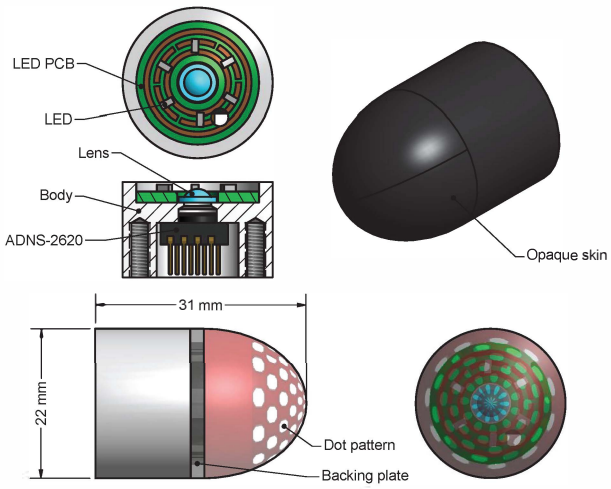

Fig. 1. Computer model of the complete sensor assembly. Anticlockwise from top left: Top view of body; side section view of body; side view of body and transparent elastomer; top view of body and transparent elastomer; complete assembly with opaque skin.

along with an array of LEDs for illumination. The camera tracks the motion of the dots, estimating displacement of the fingertip and hence inferring force. A model of the sensor is presented in Fig. 1.

The camera chip is a high-speed, low-resolution chip designed for tracking applications. It generates 18×18 pixel grayscale images, and features on-board Digital Signal Processing (DSP) computing image statistics at 1500 Hz.

Optical tactile sensing has been studied previously by Kamiyama et al. [2, 3, 4], with the *GelForce*. The sensing surface is an elastomer pad, in which coloured dots are suspended. A camera tracks the dots, and force is estimated from the inferred motion. Obinata et al. [5] have developed an optical elastomer 'fingertip' sensor where a camera tracks markers suspended in the elastomer. They estimate normal force from the size of the contact area on the sensor.

Chorley et al. [6, 7] have developed a similar tactile sensing system, where the sensor is a gel-filled rubber 'finger' with nodule markers protruding from the inside of the sensor skin, and have in particular studied tactile detection of edges.

The *GelSight* system [8] also uses a camera to track the displacement of an elastomer; microscopic surface geometry is captured by a high-resolution camera from which 3D surface models are generated.

Alternative tactile sensing approaches include an optical tactile sensor using a waveguide probed by compliant sensing elements [9], as well as semiconductor pressure sensing methods embedded in elastomer [10].

978-1-4673-5643-5/13/$31.00 ©2013 IEEE

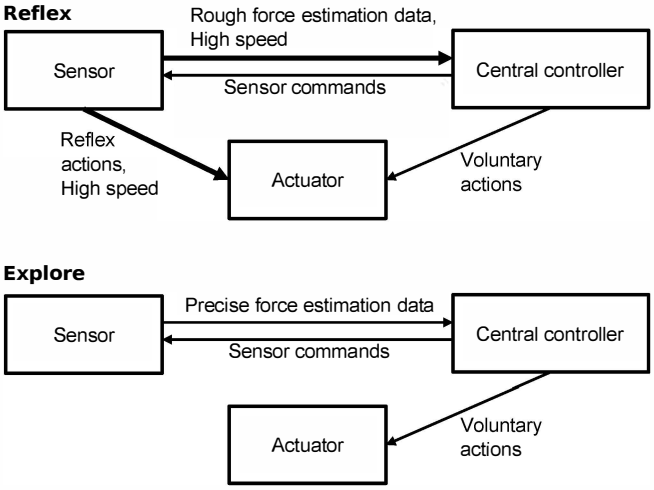

Fig. 2. Schematic representation of the two sensory modes. In *reflex mode*, the sensor can send commands to the actuator that override those sent by the central controller to provide fast reaction times.

In contrast to the work described above, we exploit the capabilities of the sensor to mimic the multi-modal interaction mechanisms in humans. This is implemented through the use of a dual-mode sensing scheme, providing a *reflex mode* for fast response times and an *explore mode* for more precise force measurement. The use of a low-resolution high-speed camera gives quick response times at the expense of precision, however we believe that the precision achieved can be sufficient for many robotics applications.

Possible applications include robotic hands or grippers for human interaction, allowing for soft and compliant interaction while at the same time yielding fast response times required for a robotic manipulator moving at moderately high speeds.

II. SENSING STRATEGIES

Biological organisms commonly have multiple sensory pathways. When performing voluntary actions such as exploring an environment, sensory information is processed by the brain at a relatively low speed, due to the high cognitive load, and motor commands are then sent from the cerebral cortex. On the other hand, reflexes provide more rapid response to stimuli by means of a control loop passing through the spinal cord and bypassing higher-level cognitive parts of the brain [11, pp 91, 94].

We mimic this behaviour in the dual-mode sensor developed here, see Fig. 2. In *explore mode*, the sensor outputs a continuous stream of data requiring processing by a central controller. This is analogous to higher-level cognitive processing of sensory input in humans. In *reflex mode*, basic processing of the sensory input is carried out by the sensor itself. This allows for fast notification of the central controller, and for the sensor to send fast commands directly to actuators given some condition on the sensory input — analogous to reflexes bypassing the cognitive parts of the brain. Features of the two modes are summarised in Table I.

TABLE I

COMPARISON OF THE SENSOR'S TWO MODES OF OPERATION.

Reflex Mode	**Explore Mode**
High speed	Low speed
Low level of detail	High level of detail
Detect presence of force	Accurate force estimation
Rough magnitude estimation	
Computed by sensor	Computed centrally

A robot could have a large number of sensors without increasing the computational load on the central controller. The robot could 'focus its attention' on one sensor and still be able to respond rapidly to stimulus appearing at any sensor through reflex actions. A decentralised control structure will also be more robust, allowing for the robot to respond to stimuli in the event of a failure of the central control system.

After a reflex action has been taken in response to stimulus of a sensor, that sensor could be switched to *explore mode* and the robot could move more slowly to investigate, e.g., the shape and size of the object it has collided with. Alternatively, for a robotic hand locating and picking up an object, *reflex mode* could be used when quickly exploring the environment to locate the object and *explore mode* could then be used to pick up the object with an appropriate force.

III. SENSOR DEVELOPMENT

The fingertip sensor (Fig. 1) and the control algorithms have been optimised for sensing of normal force, but the algorithms could be readily extended to sense shear force with the same hardware.

A. Hardware Design

The sensor is based around a moulded transparent fingertip fabricated from Terpene Phenolic Resin. For tracking by the image sensor, a semi-random pattern of ~ 1 mm diameter white dots spaced ~ 2 mm apart was painted on the surface of the fingertip before applying an opaque black rubber coating. The opaque 'skin' means that external lighting conditions do not affect sensor performance.

The fingertip exhibits a compliance similar to that of a human finger.

An *Avago ADNS-2620* low-cost tracking image sensor, featuring an 18×18 pixel camera and on-board DSP was used. The DSP computes image statistics: Movement in x- and y-directions; maximum, minimum and average pixel value; and image quality(proprietary measure of the number of features visible by the sensor — algorithm not known to us). The framerate of the sensor is 1500 Hz, and image statistics are available at this frequency. Full images can be extracted from the sensor one pixel at a time, requiring 324 frames for each image. We use the statistics for *reflex mode* operation and the images for *explore mode* operation, yielding a speed of 1500 Hz in *reflex mode* and 4.6 Hz in *explore mode*.

The tracking image sensor is interfaced to a *Microchip dsPIC30F6012A* microcontroller communicating with a PC over a USB connection.

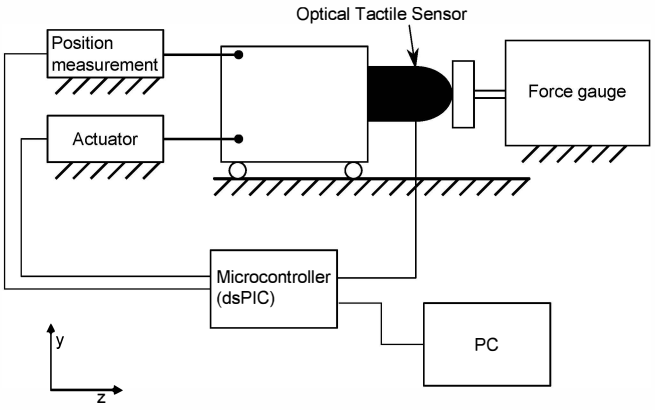

Fig. 3. Experimental setup, featuring linear actuation of the sensor in the z-direction with measurement of force and displacement. The x-direction is into the page.

A commercialised tactile sensor system could feature a single sensor chip including a camera and a microcontroller for DSP and communication. The microcontroller would estimate the sensory input based on image statistics to allowing for very fast reactions (*reflex mode*), and would also send full images to a central controller (*explore mode*).

B. Experimental Setup

The experimental setup for the tactile sensor is presented in Fig. 3. This features automatic position control in the z-direction with a displacement error less than 50 μm. A force gauge was mounted on the test rig, as shown, with a resolution of 0.01 N.

The *Microchip* microcontroller was used to communicate with the image sensor as well as interface with the actuator and the digital position sensor.

C. Force-Displacement Characteristics

We require a model of the relationship between normal force F and normal displacement of the fingertip z. To a first approximation, we assume no time-dependant stress relaxation.

An empirical approach was taken where a function was fitted to a set of data points generated experimentally. The sensor was moved to 50 positions between 0 mm and 5 mm, and the force was recorded. Between each data point, the sensor was moved back to zero displacement to minimise the effect of stress relaxation. This was repeated 4 times.

We found that the exponential function

$$F = 1.1293\left(e^{0.4677z} - 1\right), \tag{1}$$

which minimised the Root Mean Square (RMS) error, described the data well; a RMS error of 0.14 N was achieved. The data and fitted exponential are shown in Fig. 4.

D. Reflex Mode Algorithm

In *reflex mode*, we wish to mimic reflexes and involuntary actions in living organisms. The sensor should therefore respond to stimuli as quickly as possible and output an approximate estimate of the magnitude of the applied force. The force estimate is computed by the microcontroller interfacing

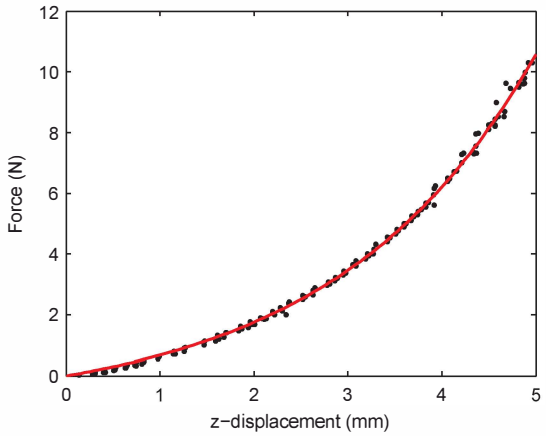

Fig. 4. Force against displacement for elastomer fingertip, with least-squares curve fitted.

TABLE II
IMAGE STATISTICS COMPUTED BY MOUSE SENSOR

Name	Description	Range
DX	Movement in x-direction	-128 to 127
DY	Movement in y-direction	-128 to 127
MAX	Maximum pixel value	0 to 63
MIN	Minimum pixel value	0 to 63
SQ	Image quality/number of features	0 to 255
AVG	Average pixel value	0 to 63

directly with the optical sensor, requiring a computationally simple force estimation algorithm for minimal response time.

Reflex mode is based on the high-speed image statistics computed by the optical sensor, listed in Table II. We wish to determine from the statistics both the presence and magnitude of an applied force.

The DX and DY statistics correspond to the shear force applied to the sensor and are not correlated to normal displacement. We therefore use the remaining 4 image statistics to infer normal force.

Detecting initial contact and the presence of a force can be considered a binary classification problem: we wish to classify data points as either 'force present' or 'no force present'. Let the logical variable FP denote 'force present'. A computationally simple classifier is required, so a decision tree was used. For an in-depth explanation of decision trees, see Breiman et al. [12].

To train the classifier, 1000 data points were generated by recording all the statistics with the sensor undergoing slow (2 mm/s) reciprocating motion into a flat plate. The data set was used to generate a decision tree. Some pruning of the tree was done, to further speed up point classification. The generated tree is given by

$$FP = (\texttt{MAX} > 40 \vee (\texttt{MAX} > 38 \wedge \texttt{SQ} < 82)). \tag{2}$$

If a data point is classed as FP, we wish to then estimate the magnitude of the force. Again, this should be computationally efficient. To a first approximation, we assume force is a linear function of the statistics, and we use least-squares regression on the generated data set to find the best linear approximation.

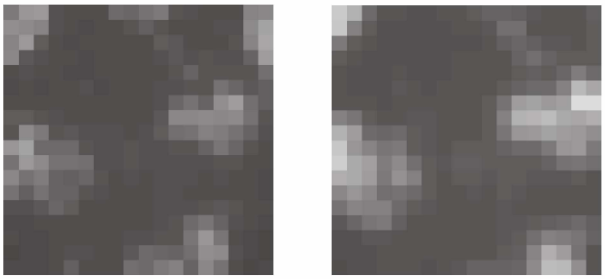

Fig. 5. Images captured by optical sensor, in its natural state (left) and with a force of 5 N applied (right). With the force applied, the sensor skin is displaced towards the optical sensor. Images are grayscale, with pixel values from 0 (black) to 63 (white).

Note that this was fitted to the data points with force present — points with no force present will be rejected by the classifier. We also improve our force estimates by setting negative forces to zero.

E. Explore Mode Algorithm

For more refined force estimates, the images generated by the optical sensor are passed from the microcontroller to the central controller for analysis. We use image processing methods to infer motion of the sensor skin from the sequence of images, from which force is then estimated. Sample images from the optical sensor are shown in Fig. 5.

Let $I(\mathbf{x}^{im}, t)$ be the pixel intensity as a function of space and time, where $\mathbf{x}^{im} = \begin{bmatrix} x^{im} & y^{im} \end{bmatrix}$ is position in the image plane and t is time. We write image plane variables with superscript *im*. Assuming a pinhole camera and small z-displacements, we can use the optical flow equation

$$\nabla I \bullet \mathbf{x}^{im} + \frac{\partial I}{\partial t} = 0 \tag{3}$$

to estimate motion from the sequence of images. See Szeliski [13] for details on camera modelling.

The sensor skin can be modelled as a plane parallel to the image plane that can translate in the x-, y- and z-directions. This assumes that any force is applied uniformly over the tip of the sensor, and that there is no moment acting on the skin about the z-axis. If the motion is constrained in this way, it can be shown that the motion field of the image can be expressed as

$$\mathbf{u} = \begin{bmatrix} \mathbf{x}^{im} & 1 \end{bmatrix} \begin{bmatrix} v_d^{im} & 0 \\ 0 & v_d^{im} \\ v_{x0}^{im} & v_{y0}^{im} \end{bmatrix} \tag{4}$$

where v_d^{im} is the dilation rate which is dependent on the z-velocity and v_{x0}^{im} and v_{y0}^{im} are the base velocities in the x- and y-directions.

Given a sequence of images, the most likely values of v_d^{im}, v_{x0}^{im} and v_{y0}^{im} are found for each frame using least-squares. Integration of the velocity components then yields the displacements of the sensor skin, from which the force is estimated. As we are estimating the normal force acting on the sensor, we are interested in the displacement of the sensor in the z-direction.

This method allows for computation of the displacement of any image, and will give good estimates of the displacements of the sensor skin provided there are some features in the image. This makes it very robust with regards to changes in the dot pattern e.g. through large x- and y-displacements or mechanical damage to the sensor skin.

However, numerical integration of the velocity estimates makes this method prone to integration drift. An alternative method is required with no integration drift.

Now, the assumption that pixel intensities are transferred can be extended inductively to $I_0(\mathbf{x}^{im}) = I(\mathbf{x}^{im} + \Delta\mathbf{x}^{im}, t)\ \forall t$ where $I_0(\mathbf{x}^{im}) = I(\mathbf{x}^{im}, t_0)$ is a reference image. Making the same assumptions of the motion of the sensor skin, we can express $\Delta\mathbf{x}^{im}$ as

$$\Delta\mathbf{x}^{im} = \begin{bmatrix} \mathbf{x}^{im} & 1 \end{bmatrix} \begin{bmatrix} t_d^{im} & 0 \\ 0 & t_d^{im} \\ t_x^{im} & t_y^{im} \end{bmatrix} \tag{5}$$

where t_d^{im} is now the dilation which is dependent on the total z-displacement and t_x^{im} and t_y^{im} are functions of the total x- and y-displacements. Let us define the vector $\mathbf{T}^{im} = \begin{bmatrix} t_x^{im} & t_y^{im} & t_d^{im} \end{bmatrix}$.

The correlation between images, $\mathrm{corr}(I_1(\mathbf{x}^{im}), I_2(\mathbf{x}^{im}))$, is a normalised measure of the distance between them taking into account overall changes in brightness levels that can arise as the sensor skin moves closer to the light source, and is thus preferable over conventional distance metrics such as Euclidean.

In order to estimate the displacements of an image relative to a reference image we thus pose the maximisation problem

$$\arg\max_{\mathbf{T}^{im}} \mathrm{corr}(I(\mathbf{x}^{im} + \Delta\mathbf{x}^{im}(\mathbf{T}^{im}), t), I_0(\mathbf{x}^{im})) \tag{6}$$

Note that for this analysis we have assumed $I(\mathbf{x}^{im}, t)$ to be a continuous function such that the elements of $\mathbf{T}^{im}$ take on real numbers. The images are sampled at discrete pixels, so linear interpolation was used to make $I(\mathbf{x}^{im}, t)$ continuous. However, this means that a closed-form solution of the problem no longer exists. Instead, the maximisation problem is solved using the simplex method.

For sufficiently small $\mathbf{T}^{im}$, i.e. sufficient overlap between the current image and the reference image, this method is expected to find a strong peak in the correlation, however it breaks down for large $\mathbf{T}^{im}$.

The value of

$$\max_{\mathbf{T}^{im}} \mathrm{corr}(I(\mathbf{x}^{im} + \Delta\mathbf{x}^{im}(\mathbf{T}^{im}), t), I_0(\mathbf{x}^{im})) \tag{7}$$

serves as an indication of the performance of this method: a correlation close to 1 indicates images with good overlap and thus good results whereas a correlation close to 0 indicates failure of this approach.

This gives us two methods for motion estimation: (1) Estimating the velocity between frames and integrating to find displacement, which is robust but prone to integration drift, and; (2) Estimating the displacement relative to a reference image, which gives absolute position but breaks down for large displacements.

A 2-mode motion estimation scheme is thus proposed: For each image, the maximisation problem in (6) is solved and the correlation at the optimum is evaluated. If the correlation is above a threshold, we use $\mathbf{T}^{im}$ for our current position estimate. Otherwise, the motion field of the current image relative to the previous image, $\mathbf{u}^{im}$, is computed and the integral of this is added to the current position estimate. The scheme thus uses the absolute motion estimate for small displacements, switches to the differential method for larger displacements, and crucially switches back to the absolute method once the sensor skin comes close to its original position thus eliminating any integration drift.

For the ideal pinhole camera model and small displacements, and for the motion considered here, it can be shown that the relationship between t_d^{im} and z is given by

$$z = Z(1/t_d^{im} - 1) \tag{8}$$

where Z is the distance from the camera origin to the sensor skin at rest. We determine the value of Z experimentally using least-squares regression.

IV. EXPERIMENTAL RESULTS

We carried out experiments to evaluate the performance of the sensor and algorithms developed in detecting and measuring normal force.

A. Reflex Mode Results

We here evaluate the performance of the *reflex mode* algorithm on the training dataset generated in Sec. III-D. The large number of data points combined with the low dimensionality of the data means that overfitting of the data set is not likely. Future work will include evaluation of the algorithm on different data sets and for different loading conditions.

Fig. 6 shows the result of applying the decision tree and the force estimation to the training dataset. The error rate of the decision tree classification is 2.1 % ($\neg FP$ classified as FP: 0.7 %, FP classified as $\neg FP$: 1.4 %). Importantly, FP is classed as $\neg FP$ only for displacement smaller than 0.21 mm corresponding to forces smaller than 0.09 N. Any force larger than this will be detected.

Note that the experiment was done under quasi-static conditions, so results do not show response time. Allowing for some computational overhead we estimate that a response time of < 2 ms would be possible. For reference, Kamiyama et al. [4] achieve a response time of 45 ms.

To a first approximation, the linear force estimation is seen to perform well. Estimates are high for low and intermediate forces, and at higher forces there is a greater spread in the force estimates. A RMS error of 0.65 N is achieved overall, with a RMS error of 0.30 N for forces smaller than 1 N. A more sophisticated force estimation algorithm would be expected to improved performance here, but at a higher computational cost.

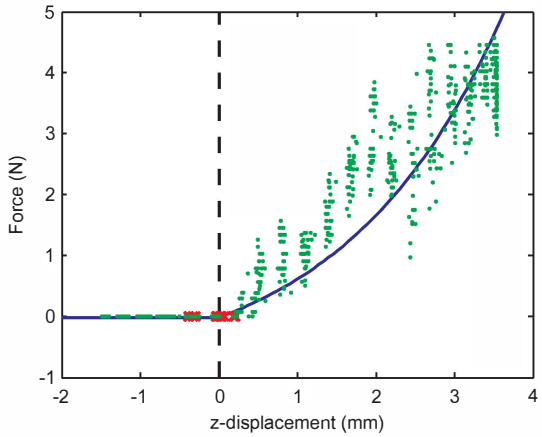

Fig. 6. Results of *reflex mode* binary classification and force estimation algorithms. Correctly classified points shown as green dots, misclassifications shown as red crosses. Blue line shows true force magnitude. Black dashed line indicates point of contact.

B. Explore Mode Results

A sequence of 1000 images was generated at 4.6 Hz with the sensor undergoing slow (2 mm/s) reciprocating motion into a flat plate. The 2-mode motion estimation scheme was then applied to compute $\mathbf{T}^{im}$. Again, we are studying normal motion so the feature of interest is t_d^{im}. The correlation threshold for switching to the relative mode was set to 0.95. Fig. 7 shows the computed value of t_d^{im}, the image dilation, for each data point together with the z-displacement of the sensor skin. Points are colour-coded by the motion estimation mode used. It can be seen that for small displacements the absolute motion estimate is primarily used, and for large displacements the relative motion estimate is used. It is also seen that there is no integration drift in the force estimates; at zero displacement t_d^{im} returns to 1.

Using the method of least-squares, the optimal value of Z in (8) was found to be $Z = 16.03$, achieving a RMS error of 0.39 mm. The resulting relationship between z and t_d^{im} has been plotted in Fig. 7, and there is seen to be good agreement between the calculated z-t_d^{im} relationship and the data.

C. Explore Mode Force Estimation Test

Combining the force-displacement relationship from (1) with the z-t_d^{im} relationship from (8), we can estimate force from a sequence of images using the *explore mode* algorithm.

A dataset (50 data points) was generated for displacement going from 0 mm to 5 mm and back to 0 mm, with force and displacement measured at quasi-static conditions. Fig. 8 shows the z-displacement and force estimated from the images together with their true values. It is seen that both the displacement and force estimates follow the true values closely, with a maximum displacement error of 0.80 mm and a maximum force error of 2.14 N. As the data in Fig. 7 is more noisy than the data in Fig. 4, we expect the tracking algorithm to be the main source of error rather than the system exhibiting hysteresis. Again, experiments are done under quasi-static conditions. In *explore mode*, the response time will be limited by the framerate (4.6 Hz).

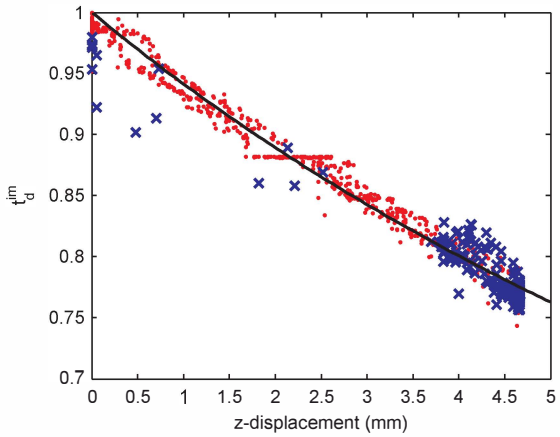

Fig. 7. Computed values of image dilation t_d^{im} against z-displacement of sensor. Red dots are computed using the absolute motion estimate, blue crosses are computed with the differential method. Also shown is the calculated z-t_d^{im} relationship, given by (8).

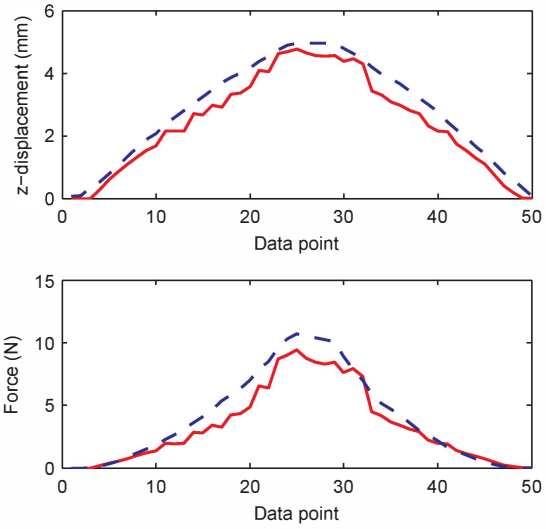

Fig. 8. Comparison of estimated (red solid lines) force and displacement to measured values (dashed blue lines) for experiment ramping displacement up and down, using *explore mode* algorithm.

The exponential force-displacement characteristics of the sensor tip means that for smaller forces the relative force error decreases. This is useful for many applications, allowing for precise and delicate manipulation as well as higher-force manipulation within a small range of sensor deformations. It can be seen in Fig. 8 that for forces smaller than ~ 2 N the force error is very small.

Kamiyama et al. [4] achieve a force magnitude resolution of 0.3 N, but at the expense of speed and higher component cost.

V. CONCLUSIONS

Optical tactile sensing has numerous advantages over traditional force sensing methods, including inherent compliance as well as resilience to mechanical damage as the sensing fingertip does not house any delicate electronics. The compliance of the sensor allows for soft and natural human-robot interaction.

Sensors can be made at low cost due to low mechanical complexity and low on-board computational requirements.

We use a dual-mode sensing scheme that mimics how biological organisms achieve fast reaction times or precise sensing depending on the situation. This dual-mode scheme is applicable to other robotic systems where many sensors and fast reaction times are required.

Further study will include the extension of sensor algorithms to detect shear force in both modes of operation, as well as further development of tracking algorithms to reduce error. Shear analysis will require study of the coupling between normal and shear force and displacement. Preliminary shear tests show promising results.

REFERENCES

[1] C. Chorley *et al.*, "A biologically inspired fingertip design for compliance and strength," *Proc. TAROS*, 2008.

[2] K. Kamiyama *et al.*, "Vision-based sensor for real-time measuring of surface traction fields," *IEEE Comput. Graph. Appl.*, 2005.

[3] K. Kamiyama *et al.*, "Evaluation of a vision-based tactile sensor," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2004.

[4] K. Kamiyama *et al.*, "Vision-based sensor for real-time measuring of surface traction fields," *IEEE Comput. Graph. Appl.*, 2005.

[5] G. Obinata *et al.*, *Mobile Robots: Perception and Navigation.* Pro Literatur Verlag, 2007, ch. Vision Based Tactile Sensor Using Transparent Elastic Fingertip for Dexterous Handling.

[6] C. Chorley *et al.*, "Development of a tactile sensor based on biologically inspired edge encoding," in *Proc. IEEE Int. Conf. Adv. Robot.*, 2009.

[7] C. Chorley *et al.*, "Tactile edge detection," in *Proc. IEEE Sensors*, 2010.

[8] M. K. Johnson *et al.*, "Microgeometry capture using an elastomeric sensor," *Proc. ACM SIGGRAPH*, 2011.

[9] M. Ohka *et al.*, "Sensing characteristics of an optical three-axis tactile sensor mounted on a multi-fingered robotic hand," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2005.

[10] T. Mukai, "Soft areal tactile sensors with embedded semiconductor pressure sensors in a structured elastic body," in *Proc. IEEE Sensors*, 2004.

[11] J. Bullock *et al.*, *Physiology*, ser. The National medical series for independent study. Lippincott Williams & Wilkins, 2001.

[12] L. Breiman *et al.*, *Classification and Regression Trees.* New York: Chapman & Hall, 1984.

[13] R. Szeliski, *Computer Vision: Algorithms and Applications*, ser. Texts in Computer Science. Springer, 2010.