Tactile Features: Recognising Touch Sensations with a Novel and Inexpensive Tactile Sensor

Tadeo Corradi, Peter Hall, and Pejman Iravani

University of Bath, Bath, UK
t.m.corradi@bath.ac.uk

Abstract. A simple and cost effective new tactile sensor is presented, based on a camera capturing images of the shading of a deformable rubber membrane. In Computer Vision, the issue of information encoding and classification is well studied. In this paper we explore different ways of encoding tactile images, including: Hu moments, Zernike Moments, Principal Component Analysis (PCA), Zernike PCA, and vectorized scaling. These encodings are tested by performing tactile shape recognition using a number of supervised approaches (Nearest Neighbor, Artificial Neural Networks, Support Vector Machines, Naive Bayes). In conclusion: the most effective way of representing tactile information is achieved by combining Zernike Moments and PCA, and the most accurate classifier is Nearest Neighbor, with which the system achieves a high degree (96.4%) of accuracy at recognising seven basic shapes.

Keywords: Haptic recognition, tactile features, tactile sensors, supervised learning.

1 Introduction

The aim of this paper is to find an accurate low-dimensional representation of a tactile image perceived by a novel tactile sensor developed by us, these representations are from now on referred to as 'encodings'. Tactile sensors and tactile information encoding have been focus of much research lately [5]. Whilst numerous standards exist in Computer Vision, there is no consensus on the best approach to encoding tactile sensing information [5], and the only tactile database known to us [24] is limited to a single sensor type. Unlike visual information, haptic information can be distributed over a potentially unknown geometry [5] (for example a single robotic hand can be fitted with many different combinations of tactile sensors), so the equivalent problem to 'camera calibration' is a significantly more difficult task. Whilst the majority of efforts have gone to low resolution sensor pads [2], [16], [19], [20], [26], a new biologically inspired sensor design, called the TacTip [3] aims to provide higher resolution whilst remaining inexpensive. This paper presents a similar, simplified, low cost tactile sensor and evaluate its accuracy recognising 7 basic tactile shapes (Corner, Cylinder, Edge, Flat-to-Edge, Flat, Nothing, Point), comparing a selection of encodings and a range of supervised classifiers.

M. Mistry et al. (Eds.): TAROS 2014, LNAI 8717, pp. 163–172, 2014.

2 Related Work

Tactile sensors can be designed using a variety of techniques, perhaps the most popular being resistive sensors [28]; but also including magnetic, piezo-electric, capacitive and others [5]. A large amount of effort has been put into texture recognition [7], [11], [14], [25], since texture is usually difficult to capture from vision alone. The most direct approach to tactile feature classification is to use the tactile images with no encoding and use a simple distance metric [23]. Recently, there have been several projects involving recognition by grasping using Pattern Recognition techniques to find the best dimensionality reduction function for tactile information. Early approaches focused on tailored designs [1] or classical Artificial Neural Networks (ANNs) [27]. More recently, PCA, moment analysis and binary (contact/no contact) have been compared in a system that integrates tactile and kinesthetic information for object recognition [8], finding that the use of central moments outperforms other encodings. A variation on Self-Organizing Maps (SOMs) [13] has been developed and applied to fusing propioceptive and tactile input for object recognition [12]. PCA and SOMs have been used to extract tactile features which were then used for object recognition [16]. Novel recursive gaussian kernels have been used to encode the various stages of contact during grasping leading to a robust online classifier [26].

2.1 The TacTip

Most previous studies are based on pressure sensor arrays. An innovative biologically inspired sensor was proposed recently [3] which uses a flexible hemispherical membrane with internal papillae which move as the membrane deforms whenever it touches an object. A digital camera records and transmits the image of the displaced papillae (see right side of Fig. 1). This sensor, called the TacTip, was shown to achieve a high degree of accuracy in sensing edges [4] to a point where a small object is clearly identifiable by a human from its tactile image and has been theoretically shown to have potential in tele-surgery [21]. More recently it has also been successfully used to identify textures [29]. The new sensor presented by this paper is an adaptation of the TacTip. No papillae nor internal gel is needed (significantly simplifying the sensors manufacture process and cost) and the shading pattern of light is used as input, instead of the papillae locations. This paper shows that the new sensor is effective at recognising tactile shapes.

3 Sensor Specification

3.1 Design

The new sensor consists of an opaque silicone rubber hemispherical membrane of radius 40mm and thickness 1mm, mounted at the end of a rigid opaque cylindrical ABS tube. At the base of the tube, there is a PC web-cam equipped with 8 white LEDs. The LEDs illuminate the rubber, the shading pattern of the image changes as the rubber makes contact with various surfaces (see Fig. 1).

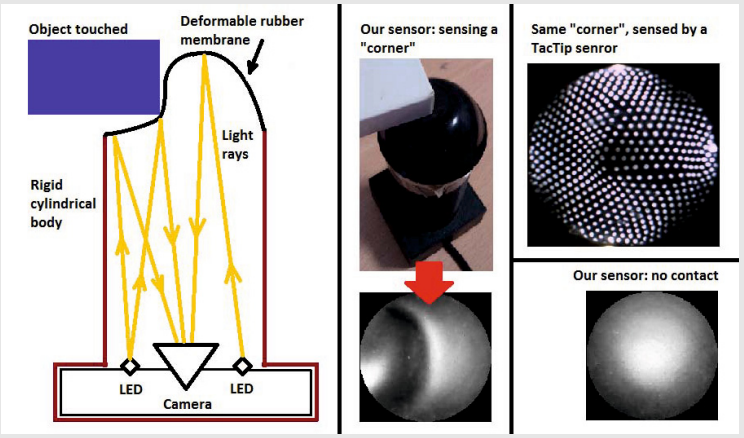

Fig. 1. The new tactile sensor design (left). The main body is 3D printed in ABS. The tip is a 1mm thick silicone rubber hemisphere. At the base (not visible) there is a USB eSecure©web-cam with 8 LEDs illuminating the inside of the silicone hemisphere (bottom right). As the tip makes contact with an object, it deforms resulting in a specific shading pattern (middle). As a comparison, the same tactile shape as perceived by a TacTip is shown (top right).

4 Methods

4.1 Preprocessing: Discrete Derivative

The shading pattern is related to the angle between the membrane's normal and the light rays going to the camera. Therefore drastic changes in luminosity are to be expected whenever the discrete spatial derivative of the normal of the surface is highest, that is where the rubber is most sharply bent (see Fig. 2). This concept motivates the analysis of the images' discrete derivative's magnitude matrix $D(I)$, defined, for any square image matrix $I \in \mathbb{R}^{w \times w}$, as:

$$D(I)_{i,j} := +\sqrt{(I_{i-1,j} - I_{i+1,j})^2 + (I_{i,j-1} + I_{i,j+1})^2}, \quad \forall i,j \in [1, w-1] \qquad (1)$$

In the experiments described below, encodings will be applied to the raw image received by the camera, and to the magnitude of its discrete derivative, $D(I)$.

4.2 Rotationally Invariant Encodings

Due to the circular geometry of the sensor image, a rotation invariant encoding was required. Five alternatives were explored: Hu moments [10], Zernike Moments [30], Principal Component Analysis (with regularized rotation), Zernike-PCA (PCA applied to the Zernike moments), and image scaling.

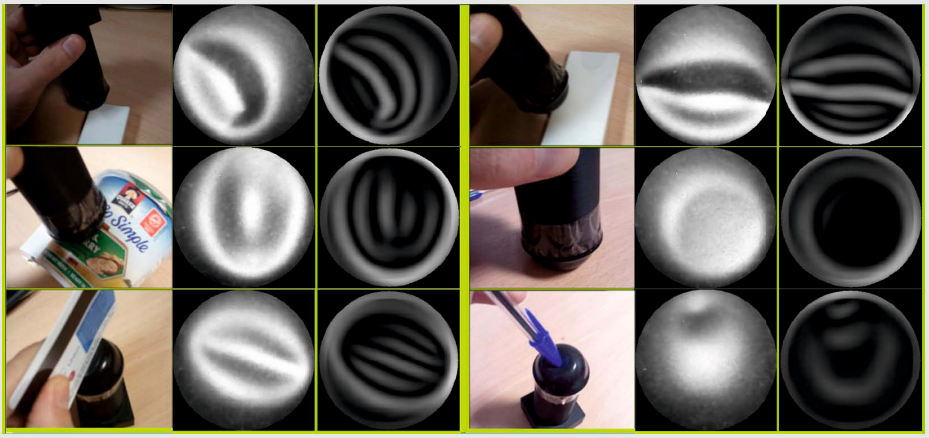

Fig. 2. Examples of occurrences of 6 of the 7 basic tactile shapes (the 7th is "nothing", in Fig. 1) (left columns), and their corresponding shading pattern (middle columns) and the magnitude of its first spatial derivative $D(I)$ (right columns). From the top left, downwards: Corner, Cylinder, Edge, Flat-to-Edge, Flat, Point.

Hu Moments. Hu moments are special combination of central moments which aim to be invariant to rotation, translation and scale (for details see [10]). The implementation used here was the one by [15], who have demonstrated the use of Hu moments in effective feature extraction on edge images for object recognition.

Zernike Moments. A Zernike Moment is the element-wise product of an image with a Zernike polynomial evaluated at the locations of the pixels of the image, rescaled to circumscribe a unit disk.

Definition 1. *Let $m \geq n$ be non-negative integers, and let $0 \leq \phi \leq 2\pi, 0 \leq \rho \leq 1$ define a polar coordinate system. Then the even and odd Zernike polynomials are defined as:*

$$Z_n^m(\rho, \varphi) = R_n^m(\rho)\, \cos(m\varphi) \tag{2}$$

$$Z_n^{-m}(\rho, \varphi) = R_n^m(\rho)\, \sin(m\varphi), \tag{3}$$

Which can be indexed by:

$$Z_j = Z_{n(j)}^{m(j)} \tag{4}$$

Where $m(j), n(j)$ are Noll's indices (See Table 1) of Zernike polynomials [17], and

$$R_n^m(\rho) = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k\,(n-k)!}{k!\,((n+m)/2-k)!\,((n-m)/2-k)!}\,\rho^{n-2\,k} \tag{5}$$

Table 1. First ten Noll indeces [17] to compose a linear sequence of Zernike polynomials

j	1	2	3	4	5	6	7	8	9	10
n(j)	0	1	1	2	2	2	3	3	3	3
m(j)	0	1	-1	0	-2	2	-1	1	-3	3

Now, the d^{th} Zernike Moment of an image M is given by:

$$Zer_d(M) = \left| \sum_{i,j \in \{i^2+j^2 \leq n^2/2\}} M(i,j) Z'_d(i,j) \right| \tag{6}$$

Where,

$$Z'_d(i,j) := Z_j \left(\frac{\sqrt{(i^2+j^2)}}{\frac{\sqrt{2}}{2}n}, arctan\left(\frac{j-n/2}{i-n/2}\right) \right) \tag{7}$$

PCA and Zernike-PCA. In the third encoding, the orientation of each image was computed (from central moments) and the image was rotated so as to regularize its orientation. Then PCA was performed on the vectorised images. The fourth encoding, Zernike-PCA, was simply applying PCA to the Zernike Moments of all images. In both of these, the dimensionality reduction matrix was computed on training data and used for both the training dataset and the testing dataset.

Scaling (Vectorized). For the fifth encoding, image orientations are regularised first, then images are resized by averaging pixel intensities, into a much smaller resolution (up to 13 by 13 pixels, from an original resolution of 300 by 300). The resulting images are vectorized, so for example, a 13-by-13 image, is converted into a 1-by-13^2 vector, by concatenating the pixel columns.

4.3 Encoding Evaluation

Each of these encodings was applied to a training dataset of 175 images, labelled from 1 to 7, corresponding to the tactile shapes they represented (see Fig. 2). Each encoding will produce a different set of data clusters. Good encodings will result in clusters which are spatially conglomerate: vectors corresponding to images of equal label will be close together and those with different labels will be far apart. One way of measuring this property is the Davies-Bouldin index in L^2 [6], defined below. Lower values of this index represent more distinctive clusters.

Definition 2. *Let $d(a, b)$ represent the euclidean distance metric. Let X be a set of vectors of dimension d, partitioned into k disjoint clusters, $X = \bigcup_{i=1}^{k} X_i$. Let c_i be the centroid of cluster X_i. The Davies−Bouldin index is given by:*

$$D = \frac{1}{k} \sum_{i=1}^{k} \max_{j:i \neq j} \left(\frac{\sigma_i + \sigma_j}{d(c_i - c_j)} \right) \tag{8}$$

Where,

$$\sigma_i := \sqrt{\frac{1}{|X_i|} \sum_{x \in X_i} d(x - c_i)^2} \tag{9}$$

Classifiers Cross Validation. As a second way of judging the suitability of a particular encoding is to train a supervised classifier given the known labels and to test their accuracy at predicting the labels of the encoded data. The measure used here is the 5-fold cross validation accuracy, defined as the average percentage of correct classifications performed by a given classifier trained with $\frac{4}{5}$ of the labelled data and tested on the remaining $\frac{1}{5}$ of the data. The process is repeated 5 times so that all data is used for testing. This method was applied to the following classifiers:

- Nearest Neighbor classifier
- Artificial Neural Network with a single 7 neuron hidden layer, trained using backpropagation.
- A group of seven binary Support Vector Machines (one per label) used in conjunction, arbitrarily choosing the largest label id, if more than one returned a positive classification.
- A simple Naive Bayes classifer, using Kernel Density Estimation (KDE) [18], [22].

For the implementation of these four algorithms, and for the simulations described in this paper, MATLAB[1] was used.

5 Results

Seven basic tactile shapes were defined: Corner, Cylinder, Edge, Flat-to-Edge, Flat, Nothing, and Point. Using the new sensor images were manually captured, resulting in 70 sample frames of each one (see Fig. 2). Data was split: 175 images were used for training (selecting the optimum encoding vector size), and the remaining 175 images for validation. Each one of the encodings defined in Section 4.2 was applied to each training image and the magnitude of its discrete derivatives (as described in Section 4.1). Then two tests were performed: cluster evaluation and classifier evaluation.

[1] MATLAB©, Statistics Toolbox and Neural Network Toolbox Release 2013b, The MathWorks, Inc., Natick, Massachusetts, United States.

5.1 Cluster Evaluation

First, the Davies-Bouldin index was computed on the training data data (175 images) to find the optimum number of components to use in each encoding (number of principal components, number of zernike polynomials, etc.). This parameter (number of components) is then fixed and the Davies-Bouldin index is computed on the remaining 175 images (the validation dataset). Table 2 shows the result. Zernike moments combined with PCA seem to produce the most distinct clusters under this criteria. Cluster formation using the new sensor seems superior with respect to the TacTip using this measure. This is possibly due to the fact that papillae displacements mean that small perturbations in the object surface translate into significant non-linear changes in the image.

Table 2. Davies−Bouldin index (described in Section 4.3) computed for the clusters resulting from the different encodings. They represent the distinctiveness of a cluster, smaller numbers represent better defined clusters.

Encoding	Applied to Image (Our sensor)	Applied to Image (TacTip sensor)	Applied to D(Image) (Our sensor)	Applied to D(Image) (TacTip sensor)
Hu Moments	5.3	10.4	5.1	13.2
Zernike M.	2.0	2.5	1.9	3.8
PCA	2.6	5.9	1.8	5.4
ZernikePCA	1.5	2.6	1.4	2.9
Scale (Vect.)	37.1	37.9	10.4	1378.8

5.2 Classifier Evaluation

Each one of the classifiers described in Section 4.3 is now trained. Using 20 iterations of randomized 5-fold cross validation on the training dataset the optimal vector sizes for each encoding and classifier are obtained. Then, the process is repeated on the validation dataset, but only using these optimum vector sizes. Figure 3 shows the accuracy of each encoding/classifier pair.

Zernike PCA applied directly to the image outperforms other encodings in general. In terms of classifiers, Nearest Neighbor is the overall best for both sensors, reaching an accuracy on the validation dataset of 96.4%. It must be born in mind that Nearest Neighbor classifiers using cross validation are prone to data twinning (bias if similar data are present in a dataset). To reduce the effects, a small value for k (5) was used in k-fold cross validation, together with randomisation and multiple trials; furthermore, separate dataset were used for training and validation. Nevertheless, if data twinning is likely to be an issue in further applications, it may be advisable to use Naive Bayes (KDE).

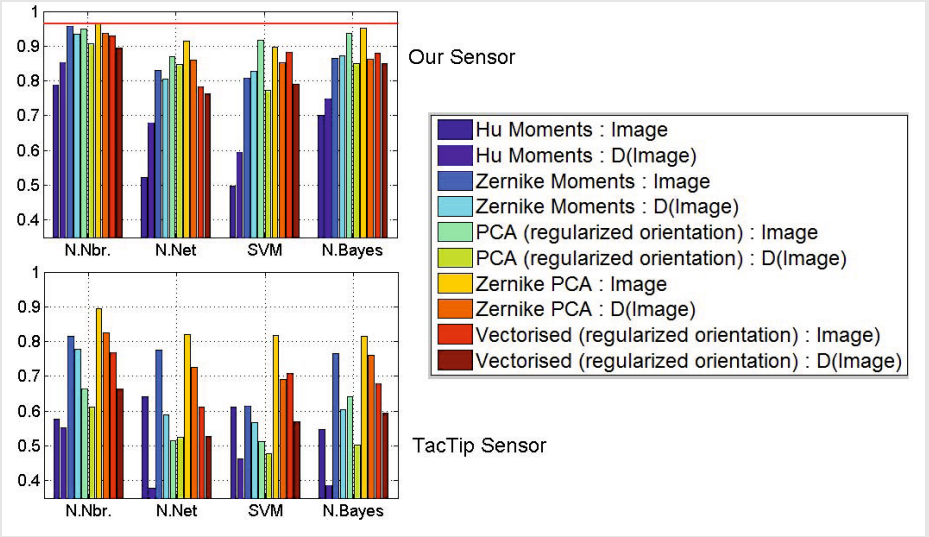

Fig. 3. Randomized 5-fold cross validation accuracy for the 7 basic tactile shapes (higher is better, 1 is 100% perfect recognition). Input set of 175 labelled tactile images, corresponding to 7 clusters. Comparison between our sensor and the TacTip, using four different encodings as classified by four different supervised algorithms.

There is no significant difference between the performance of any encoding/classifier pairing when comparing their use on the image and on its derivative. This may be due to the fact that the discrete derivative only loses base intensity information, which is a single degree of freedom over images which are 90000-dimensional. The accuracy achieved with our sensor is slightly higher to the one with the TacTip, for these particular choices of encodings and classifiers. Once again, the non-linearity introduced by papillae is potentially a factor, and so the comparison is by no means exhaustive in scope.

6 Conclusions

This paper presented a novel, simple and inexpensive tactile sensor based on shading resulting from the deformation of a rubber membrane. Various encodings were tested on the input images and on their discrete derivatives. For each encoding, the accuracy of a selection of classifiers was tested, by performing tactile shape recognition. The new sensor is capable of distinguishing between these shapes, the most accurate encoding is Zernike Moments combined with PCA, applied directly to the input image. The most accurate classifier is Nearest Neighbor, which reaches a classification accuracy of 96.4%. Our sensor performed slightly better than the TacTip in these tests, which is remarkable considering the simplicity of our sensor's design. However it must be stressed that other approaches may very well favor the TacTip. The discrete localization of the papillae

may be a disadvantage in linear encodings, but it can be an advantage in general, as it is more resilient to image noise and less dependent on calibration of camera parameters. Only pattern recognition was discussed in this paper, it may be of interest to use "shape from shading" [9] to reconstruct the exact shape of the deformed hemisphere. Further work should also focus on this sensor's potential for object recognition.

Acknowledgments. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), UK. We would like to thank Bristol Robotics Lab[2] for lending us the TacTip sensor.

References

1. Allen, P.K.: Integrating vision and touch for object recognition tasks. The International Journal of Robotics Research 7(6), 15–33 (1988)
2. Barron-Gonzalez, H., Prescott, T.: Discrimination of social tactile gestures using biomimetic skin. In: IEEE International Conference on Robotics and Automation, Karlsruhe, Germany (2013)
3. Chorley, C., Melhuish, C., Pipe, T., Rossiter, J.: Development of a tactile sensor based on biologically inspired edge encoding. In: International Conference on Advanced Robotics, ICAR 2009, pp. 1–6 (2009)
4. Chorley, C., Melhuish, C., Pipe, T., Rossiter, J.: Tactile edge detection. In: 2010 IEEE Sensors, pp. 2593–2598 (2010)
5. Dahiya, R., Mittendorfer, P., Valle, M., Cheng, G., Lumelsky, V.: Directions toward effective utilization of tactile skin: A review. IEEE Sensors Journal 13(11), 4121–4138 (2013)
6. Davies, D.L., Bouldin, D.W.: A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-1(2), 224–227 (1979)
7. Decherchi, S., Gastaldo, P., Dahiya, R., Valle, M., Zunino, R.: Tactile-data classification of contact materials using computational intelligence. IEEE Transactions on Robotics 27(3), 635–639 (2011)
8. Gorges, N., Navarro, S., Goger, D., Worn, H.: Haptic object recognition using passive joints and haptic key features. In: 2010 IEEE International Conference on Robotics and Automation (ICRA), pp. 2349–2355 (2010)
9. Horn, B.K.P., Brooks, M.J. (eds.): Shape from Shading. MIT Press, Cambridge (1989)
10. Hu, M.K.: Visual pattern recognition by moment invariants. IRE Transactions on Information Theory 8(2), 179–187 (1962)
11. Jamali, N., Sammut, C.: Majority voting: Material classification by tactile sensing using surface texture. IEEE Transactions on Robotics 27(3), 508–521 (2011)
12. Johnsson, M., Balkenius, C.: Sense of touch in robots with self-organizing maps. IEEE Transactions on Robotics 27(3), 498–507 (2011)
13. Kohonen, T.: Self-organized formation of topologically correct feature maps. Biological Cybernetics 43(1), 59–69 (1982)
14. Liu, H., Song, X., Bimbo, J., Seneviratne, L., Althoefer, K.: Surface material recognition through haptic exploration using an intelligent contact sensing finger. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 52–57 (2012)

[2] `www.brl.ac.uk`

15. Mercimek, M., Gulez, K., Mumcu, T.V.: Real object recognition using moment invariants. Sadhana 30(6), 765–775 (2005)
16. Navarro, S., Gorges, N., Worn, H., Schill, J., Asfour, T., Dillmann, R.: Haptic object recognition for multi-fingered robot hands. In: 2012 IEEE Haptics Symposium (HAPTICS), pp. 497–502 (2012)
17. Noll, R.J.: Zernike polynomials and atmospheric turbulence. Journal of the Optical Society of America 66(3), 207–211 (1976)
18. Parzen, E.: On estimation of a probability density function and mode. Annals of Mathematical Statistics 33, 1065–1076 (1962)
19. Pezzementi, Z., Plaku, E., Reyda, C., Hager, G.: Tactile-object recognition from appearance information. IEEE Transactions on Robotics 27(3), 473–487 (2011)
20. Ratnasingam, S., McGinnity, T.: A comparison of encoding schemes for haptic object recognition using a biologically plausible spiking neural network. In: 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3446–3453 (2011)
21. Roke, C., Melhuish, C., Pipe, T., Drury, D., Chorley, C.: Deformation-based tactile feedback using a biologically-inspired sensor and a modified display. In: Groß, R., Alboul, L., Melhuish, C., Witkowski, M., Prescott, T.J., Penders, J. (eds.) TAROS 2011. LNCS, vol. 6856, pp. 114–124. Springer, Heidelberg (2011)
22. Rosenblatt, M.: Remarks on some nonparametric estimates of a density function. The Annals of Mathematical Statistics 27(3), 832–837 (1956)
23. Schneider, A., Sturm, J., Stachniss, C., Reisert, M., Burkhardt, H., Burgard, W.: Object identification with tactile sensors using bag-of-features. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009. pp. 243–248 (2009)
24. Schopfer, M., Ritter, H., Heidemann, G.: Acquisition and application of a tactile database. In: 2007 IEEE International Conference on Robotics and Automation, pp. 1517–1522 (2007)
25. Sinapov, J., Sukhoy, V., Sahai, R., Stoytchev, A.: Vibrotactile recognition and categorization of surfaces by a humanoid robot. IEEE Transactions on Robotics 27(3), 488–497 (2011)
26. Soh, H., Su, Y., Demiris, Y.: Online spatio-temporal gaussian process experts with application to tactile classification. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4489–4496 (2012)
27. Taddeucci, D., Laschi, C., Lazzarini, R., Magni, R., Dario, P., Starita, A.: An approach to integrated tactile perception. In: 1997 IEEE International Conference on Robotics and Automation, vol. 4, pp. 3100–3105 (1997)
28. Weiss, K., Worn, H.: The working principle of resistive tactile sensor cells. In: 2005 IEEE International Conference on Mechatronics and Automation, vol. 1, pp. 471–476 (2005)
29. Winstone, B., Griffiths, G., Pipe, T., Melhuish, C., Rossiter, J.: TACTIP - tactile fingertip device, texture analysis through optical tracking of skin features. In: Lepora, N.F., Mura, A., Krapp, H.G., Verschure, P.F.M.J., Prescott, T.J. (eds.) Living Machines 2013. LNCS, vol. 8064, pp. 323–334. Springer, Heidelberg (2013)
30. Zernike, V.: Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode. Physica 1(7-12), 689–704 (1934)