2011 RO-MAN
20th IEEE International Symposium on Robot and
Human Interactive Communication,
July 31 - August 3, 2011, Atlanta, GA, USA

Histogram based classification of tactile patterns on periodically distributed skin sensors for a humanoid robot

Ze Ji, Farshid Amirabdollahian, Daniel Polani and Kerstin Dautenhahn

Abstract— **The main target of this work is to improve human-robot interaction capabilities, by adding a new modality of sense, touch, to KASPAR, a humanoid robot. Large scale distributed skin-like sensors are designed and integrated on the robot, covering KASPAR at various locations. One of the challenges is to classify different types of touch. Unlike digital images represented by grids of pixels, the geometrical structure of the sensor array limits the capability of straightforward application of well-established approaches for image patterns. This paper introduces a novel histogram-based classification algorithm, transforming tactile data into histograms of local features termed as codebook. Tactile pattern can be invariant at periodical locations, allowing tactile pattern classification using a smaller number of training data, instead of using training data from everywhere on the large scale skin sensors. To generate the codebook, this method uses a two-layer approach, namely local neighbourhood structures and encodings of pressure distribution of the local neighbourhood. Classification is performed based on the constructed features using Support Vector Machine (SVM) with the intersection kernel. Real experimental data are used for experiment to classify different patterns and have shown promising accuracy. To evaluate the performance, it is also compared with the SVM using the Radial Basis Function (RBF) kernel and results are discussed from both aspects of accuracy and the location invariance property.**

I. INTRODUCTION

Research on tactile sensing in robotics has been mainly focusing on specific locations or body parts, such as the end-effector tip, providing feedback to the manipulators for the purpose of grasping objects. In Human Robot Interaction (HRI), there is a demand for more complex forms of interaction. Previous studies of human's tactile behaviours on a humanoid robot have shown [15] that spatial coverage of the sensors needs to be extended in order to detect various unpredictable tactile events. Moreover, tactile sensing in the case of social cognitive interaction plays an important role in conveying information related to the user's intention.

To improve human-robot interaction capabilities, one of the research challenges is the classification of various interaction types, as summarised in [15]. With the current capacitive-based technology, the RoboSkin project [17] has developed a set of flexible sensors that are mounted on a humanoid robot, KASPAR (figure 1).

Temporal and spatial patterns together form a complete descriptor of a touch event. Spatial patterns can be even further divided into global locations and relative local regions. For example, for an event of touching on the robot's hand, the touching type, e.g. finger poking, is described by the local spatial pattern and the robot's hand is its location. However, it would be very inefficient to train the system to learn all possible tactile patterns which would involve many repetitive and duplicated touches, e.g. touching at all locations on KASPAR. This paper explores the effectiveness of a histogram-based feature space for tactile patterns to represent a large volume of data with a relative compact dataset by discarding the global location information. It is expected that, by this, a small amount of training data would provide sufficient information on local spatial tactile patterns. To be specific, due to the periodical pattern of the skin sensors, using similar skin patches over the robot body, a histogram-based method for classifying tactile patterns allows to identify patterns that are invariant in the histogram feature space at the periodical locations. For example, index finger's poking at two periodical positions should produce the same histograms. One advantage of this is that it allows for classification of tactile patterns with limited training data, e.g. obtained only in a small region.

The remainder of this paper is organised as follows. Section 2 describes related work, including the state-of-the-art in tactile interaction and recognition and feature extraction for classification. Section 3 introduces our robot platform and the skin sensors. Section 4 then follows by presenting our proposed method for classifying tactile patterns using a histogram-based technique. The experiments and results are discussed in section 5. The last section summarises this work.

II. RELATED WORK

Tactile interaction and recognition research are mostly motivated by the following three objectives: safety of physical human-robot contact [2], tactile feedback for manipulating objects [21], and tactile Human Robot Interaction (HRI) [4]. In this work we focus on applying tactile sensing to improve HRI capabilities. In this domain, many different robot types have been integrated with tactile sensors, including the RI-MAN robot designed for lifting humans [12], KaMERo, a hard-covered robot, designed for recognising a person's emotional state [10], and small companion robots [23] [20] [7].

For many years, tactile interaction research has attracted considerable attention as a means of emotional communication between humans and robots. Emergence of emotional behaviours through physical interaction between humans and pet robots is one of the early works in this direction [19]. It has also been shown to have a great potential for therapeutic purposes, by adding tactile sensors to companion robots [23]

Z. Ji, F. Amirabdollahian, D. Polani and K. Dautenhahn are with the Adaptive Systems Research Group, School of Computer Science, Universit of Hertfordshire, College Lane, Hatfield AL10 9AB, United Kingdom {z.ji, f.amirabdollahian2, d.polani, k.dautenhahn}@herts.ac.uk

[20]. This is also very close to one of our project targets: to improve the interaction experience between a humanoid robot and children with autism by using skin-like sensors. Design challenges and user requirements for augmenting a humanoid robot with tactile sensors are detailed in [15], [16].

Regarding classification of tactile patterns, many attempts have been made on extracting temporal changes as the feature for temporal event classification. An efficient touch pattern recognition system using a Temporal Decision Tree (TDT) method is introduced in [10]. This method is implemented with an array of 9 capacitive sensors and one piezoelectric accelerometer, which are integrated on the KaMERo robot. The classification mainly deals with temporal patterns, based on the TDT method, and also introduces the consistency index mechanism to minimise the false classification chance during real-time interaction. The features that are used to distinguish single data samples are the basic force level and contact area. Another example is the small companion robot, Huggable [20], using the temporal difference as one feature. The neural network is then fed with the feature data for classification.

One early work focusing on spatial pattern classification uses the neural network with the Fourier spectrum space as the feature space [22]. In addition, there are also related work in other domains, such as hand gesture recognition on multi-touch screens. For example, the Support Vector Machine (SVM) is used for classifying hand gestures using detected contours as features [24].

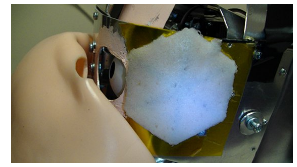
Iwata et al. introduce a method for designing an identification system for human-robot contact states based on tactile recognition, named PIFACT [8]. Similarly, an interesting application for estimating a person's position is presented in [11], which maps the locations and postures of a person with the tactile sensors statistically.

Slightly different from the above work, many attempts have also been made to use robot manipulators to explore environments. For a specific purpose of estimating parameters of an object's orientation and location, a Baysian approach is proposed in [14]. The Bag-of-Features model, inspired by image categorisation, is employed for recognising objects with a sensitive manipulator [18].

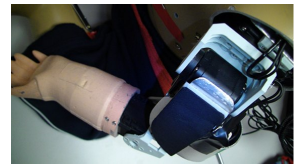
Tactile patterns are similar to grey-scale patterns in digital images, which are represented as grids of pixels. For digital images, the BoF method, which is employed in this work, transforms local features into the histogram form, which has been proved effective in applications of image categorisation [13]. This approach decomposes the whole image into several small local features, which are normally chosen by detecting local salience, and discards their global relations. The local features are called codewords and the whole set of these codewords is a structured sequence, termed as codebook. The idea of using histogram in the codebook domain is based on the assumption that a certain category of images must contain some local features, e.g. the natural scene category would contain more features such as sky, green trees, or similar objects.

Feature extraction from large dimensional data is also relevant to this work, which deals with large-scale distributed tactile sensors. Despite the similarities between skin-like sensors and grey-scale images, there are some differences between the two areas in the data structure and the fundamental nature of the data. Not many attempts have been made in this direction.

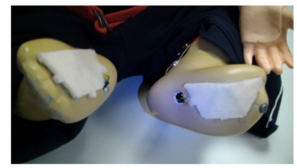
(a) The humanoid robot, KASPAR



(b) One sensor on the face



(c) The left arm and hand covered by sensors



(d) Two sensors on both feet

Fig. 1: The humanoid robot, KASPAR and some skin sensors

III. The robot and skin sensors

A. The robot, KASPAR

In our work, a child-sized humanoid robot, KASPAR, is used as the platform. The main motivation of using this robot is based on the previous observation of tactile interactions using this robot with children with autism. The robot (figure 1), as detailed in [5], has been shown as a suitable interactive platform for children with autism, for its minimally expressive face and appropriate number of degrees of freedom.

B. Skin sensors

KASPAR is mounted with skin sensors at several locations, which are the hands, arms, cheeks, feet, and torso respectively, as depicted in figure 1. The skin sensors are based on the capacitive sensing technology, with a layer of foam on top, that can effectively reduce the sensitivity and enable it to distinguish a range of pressures, which are caused by the elastic effect of the top foam layer [17] [1]. The sensors are constructed using a number of triangular patches, each with 12 sensors, denoted taxels here. The triangle patches are made of flexible printed circuit boards (PCB), that can be deformed to wrap on slightly curved surfaces. Figure 2

[1]The sensors are fabricated by the Italian Institute of Technology (IIT), as part of the RoboSkin consortium

shows one of such sensors and its taxel distribution for one part of the sensors on KASPAR is depicted in figure 3. More information of the sensors can be found in [17]. In total, on Kaspar, there are 94 triangles, hence, 1128 taxels. The sampling rate is 50 Hz with the sensor acquisition system.

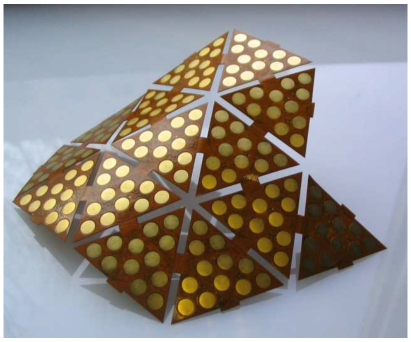


Fig. 2: A skin-sensor patch prototype

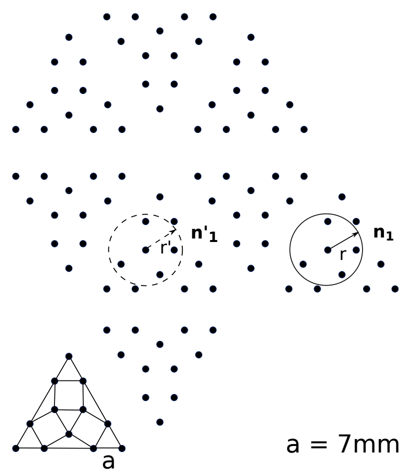


Fig. 3: Taxel layout of part of the right hand sensor

IV. Histogram-based tactile pattern classification

This section introduces the classification algorithm for tactile patterns. There are two main stages in our approach, namely the codebook generation and the classification stage. The main focus of this paper is on the codebook, which is used to construct histograms of local features of tactile patterns. With the histograms as structured features, the classification is then performed using the Support Vector Machine (SVM) with the intersection kernel [1].

The basic idea of using histograms to represent tactile patterns is as follows. With a simple example of greyscale images, the greyscale values in the image can be transformed into the histogram form. Figure 4 shows two examples of the histograms of two greyscale patterns, ellipse and cross, respectively. The difference of the two histograms can be seen clearly.

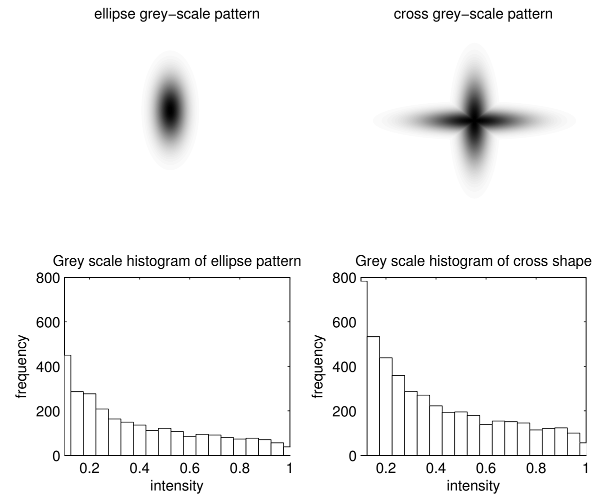


Fig. 4: Two grey-scale patterns and their grey-scale histograms

However, this feature space will not be reliable to discriminate patterns. For example, a larger ellipse in the figure may have a histogram more similar to that of the cross pattern, because both are generated with the Gaussian distribution in this example. The main limit is that the greyscale histogram representation abandons all spatial structure information by only using single pixels. Inspired by the Bag-of-Feature (BoF) method mentioned earlier, we employ local neighbourhood regions as feature units, and each bin in the histogram is not a single grey value, but the greyscale intensity distribution of the neighbour pixels. The intensity distribution over these local neighbour pixels retain relative spatial features. With the same patterns above, we can segment local square regions of the same size. Figure 5 shows some typical local square regions, segmented from the original patterns to represent their local features. It can be seen that places like corners and edges in the above examples can produce distinctive features that can be useful in producing higher variations in the histograms of the two patterns. Here, these local features are the codewords and they altogether form a codebook. The codebook is a structured set of encodings of local features or codewords. The histogram is a statistical representation of the codewords. The histogram of those local features can be used as the feature space for classification. One advantage of using this feature space is the location-independence for pattern classification. In other words, the same pattern at different locations can produce identical histograms.

A. Codebook descriptor of local features

The above example shows that, with digital images, it is possible to segment grids of pixels of the same dimension to construct the codebook. As the skin sensors are not distributed in uniform grids, the local neighbourhood sensors do not have a unified local geometrical structure (figure 3). Therefore, if we segment local features as used for images,

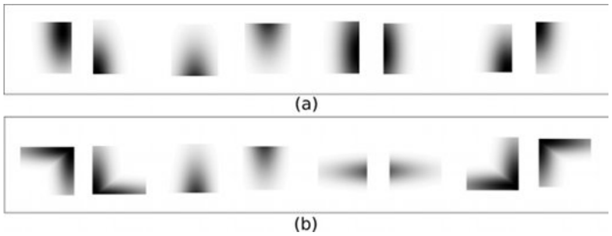


Fig. 5: Some local features extracted from two grey-scale patterns of (a) ellipse and (b) cross

the similarity between local features cannot be measured, due to the mismatched taxels.

This section introduces the construction of the codebook for the tactile data. Mainly, the codebook is a two-layer descriptor, as illustrated in figure 6. A codeword is defined as an encoding of pressure values of neighbour taxels. Each local neighbourhood structure has a number of codewords as its sub-layer representing different pressure values. The next two sub-sections detail the two layers respectively:

- local neighbourhood structures;
- codewords at the sub-layer.

1) Local neighbourhood structure: Although the tactile sensors are not uniformly distributed as grid pixels in images, the layout is periodical that local neighbourhood structures are repetitive periodically. The purpose of this stage is to allow pairwise comparison of tactile data for the periodical local regions.

We segment local regions with the distance criterion. With a taxel as the centre of a local region, its neighbouring taxels are those with a distance lower than a pre-defined radius r (figure 3). The selection of r is related to the number of neighbouring taxels. A neighbourhood structure can be formulated as:

$$\mathbf{n_{t_c}} = (t_c, n_1, n_2, \dots, n_M)^T \ subject \ to \ d(n_i, t_c) \le r \tag{1}$$

where $\mathbf{n_{t_c}}$ is the vector of neighbour taxels of the centre taxel t_c, $d(n_i, t_c)$ is the distance between the neighbour taxel n_i and t_c, and it satisfies the condition of pre-defined distance limit r. The neighbour taxels are sorted in the clockwise order.

As discussed above, the taxel layout is in a periodical pattern and local neighbourhood structures are repetitive periodically. Figure 3 illustrates an example of two identical neighbourhood structures $\mathbf{n_1}$ and $\mathbf{n_1'}$ with the same radii r and r' centred at two taxels located at periodical positions. Therefore, altogether, there are only a limited number of unique local neighbourhood structures. The whole set of local neighbourhood structures together form the first-layer of the codebook, as illustrated in figure 6.

With KASPAR's sensors, in practice, to minimise the number of local structures, we use a less strict rule that two local neighbourhood structures are considered identical if they are similar enough. As an example, with one of the KASPAR's sensor patches (figure 3), 27 local structures are created, when $r = 8mm$. Each taxel, with it as the centre of its local neighbourhood, is then mapped to a corresponding local structure at the first layer in the codebook. Figure 6 illustrates the mapping relationship between the taxels and their local neighbourhood structures, where there are N neighbourhood structures, and every one is mapped to one or more taxels as the local centres.

2) Sub-layer features using vector quantisation: Rather than using encodings of local features for square regions as in figure 5, we create encodings of features for every local structure defined in the last stage. This layer is the codebook containing encoding models for every local structure. As an example, in figure 6, 4 sub-layer encodings for each local region are used as the codewords. It can be seen that, similar to the local features in figure 5, these sub-layer encodings represent some typical pressure distributions.

However, how to create optimal encodings is a challenge. For this purpose, we use the vector quantization (VQ) method, which is a popular approach to find an encoding of vectors that reduces the expected distortion [6].

The k-means clustering implementation is used here for the VQ. We first prepare a large number of data samples by touching the sensors randomly. These recorded data are used to generate the sub-layer features or codewords. With all local neighbourhood structures created earlier, the pressure values of the neighbour taxels are fed into the k-means clustering method. The centroid of each cluster is used to represent that cluster. Given k as the number of clusters for each set of neighbour taxels, the complete codebook can be constructed with a dimension of $k \times N$, where N is the number of codewords in the first layer (figure 6). The optimal value of k should be selected with consideration of both efficiency and accuracy.

3) Creating histograms: With the codebook created above, we can now transform a tactile data sample into the histogram feature space. The basic idea is to find the closest codeword to the pressure values of the current neighbour taxels. The closest codeword will be incremented by 1 in the histogram. With all taxels processed iteratively, a complete histogram of local features can be created. The process is introduced in algorithm 1.

Algorithm 1 Transforming data sample into histogram

for all taxels, $i = 1$ to S **do**
  $\mathbf{n_{tc}}, l \leftarrow$ first layer codeword of current $taxel_i$ {$\mathbf{n_{tc}}$ is the neighbour taxels, and l is its index in the first layer codebook (figure 6)}
  $\mathbf{p} \leftarrow$ pressure values of $\mathbf{n_{tc}}$ and $taxel_i$
  $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_k} \leftarrow$ corresponding sub-layer codewords of $\mathbf{n_{tc}}$
  $x \leftarrow \arg\min_j ||\mathbf{v_j} - \mathbf{p}||, j = 1, 2, \dots, k$
  $\mathbf{h}((l-1)*k+x) \leftarrow \mathbf{h}((l-1)*k+x)+1$ {$\mathbf{h}$ is the histogram}
end for

Now, any tactile pattern can be transformed into the histogram space. Recall that it is expected that the histogram

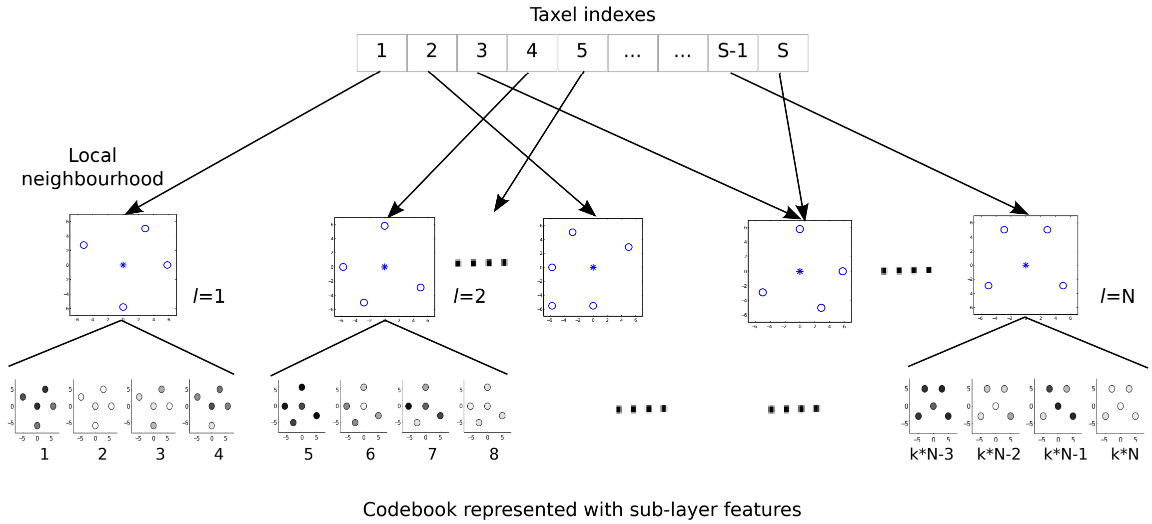


Fig. 6: Structure of the two-layer codebook. There are N local neighbourhood structures and k codewords or encodings for each local neighbourhood structure. In total, there are $N \times k$ codewords.

representation of a tactile pattern should be invariant at periodical locations. For example, index finger's poking at two different but periodical positions should produce the same histogram representation, because they can produce the same local pressure distributions at the same local neighbourhood taxels. This is also the aim of this work, allowing classification of tactile patterns with limited training data, e.g. obtained from a cyclic region.

B. Classification with SVM

As specified earlier, the main purpose of this work is to evaluate the effectiveness of using histograms to represent tactile patterns. For this purpose, we measure the performance with classification accuracy of different touch types.

As histograms lie in a probability-like space, the Euclidean distance or the generalised Minkovski distance is not a suitable metric here. The intersection metric has been widely used for histogram similarity measurement and considered as an appropriate metric for histogram space [9]. Due to its non-linear feature, a suitable classifier for the above introduced feature space should be able to handle non-linear data. The Support Vector Machine (SVM) [3] together with various kernel tricks provides a convenient mechanism for discriminative analysis in both linear and non-linear spaces. The intersection kernel is designed for histogram features and has been proved to be a Mercer's kerne [1]. Therefore, we use SVM for classifying the types of touch pattern.

A brief introduction of SVM is given below. The basic idea is to construct a hyperplane or a set of hyperplanes in a high dimensional space that separates classes by maximising the distance or margin between classes. Let $(\mathbf{x_i}, y_i)$ be a set of pairs of data and the corresponding classes, e.g. $y_i \in \{-1, 1\}$ for two classes. A hyperplane can be written as

$$\mathbf{w} \cdot \mathbf{x} - b = 0 \tag{2}$$

where $\cdot$ denotes the dot product. The vector $\mathbf{w}$ is perpendicular to the hyperplane. Data points should be separated by two hyperplanes, that can be formulated with the following constraint:

$$y_i(\mathbf{w} \cdot \mathbf{x_i} - b) \geq 1 \qquad for\ all\ 1 \leq i \leq n \tag{3}$$

The distance between the two hyperplanes is $2/||\mathbf{w}||$. If the data is to be separated by the two hyperplanes, these need to be selected in a way that no data point would fall between the two planes. Thus the distance between the planes needs to be maximised under the constraint above (equation 3). The above problem leads to the maximisation of

$$\begin{array}{c} L = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j (\mathbf{x_i^T x_j}) \\ subject\ to\ \alpha_i \geq 0 \ \ and \ \ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{array} \tag{4}$$

where α_i are the Lagrange multipliers to solve the above optimisation problem.

For non-linear data, it is usually incorporated with kernel machines, which allow the algorithm to fit the maximum-margin hyperplane in a feature space. As a general form, the inner product $\mathbf{x_i^T x_j}$ in equation 4 can be replaced with a kernel function, $K(\mathbf{x_i}, \mathbf{x_j})$. Popular kernel functions include the Polynomial kernel and Radial Basis Function (RBF).

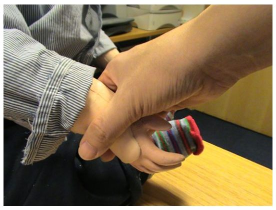
As mentioned earlier, naive Minkovski distance is not a suitable metric here, because histograms lie in a probability-like space. Hence, here we use the SVM with the intersection kernel [1], which is defined as:

$$K(A, B) = \sum_{i=1}^{m} \min\{a_i, b_i\} \tag{5}$$

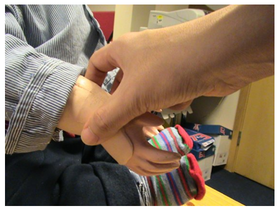
where A and B are two histograms for comparison, and a_i and b_i are the ith bins of the two histograms respectively. Since the numbers of taxels in A and B are the same, $\Sigma_i^m a_i = \Sigma_i^m b_i$, this satisfies the normalisation condition for using the intersection kernel.

(a) Poking with index finger

(b) Larger contact area with index finger



(c) Whole hand grasp



(d) Gripping with thumb, index, and middle fingers

Fig. 7: Four examples of touch types

Ultimately, the SVM here acts as a classifier, given a point, here a touch histogram, it identifies whether this point belongs to either sides of its classification planes.

V. Experiments and discussion

In this section, we evaluate the performance of the proposed method with real experimental data. It is then followed by discussion and future work in the last section.

A. *Experiment setup*

As mentioned before, the purpose of the work is twofold:

- to evaluate if histograms of local features hold sufficient information to discriminate between different patterns;
- to evaluate the histogram invariance of tactile patterns at periodical locations.

For these purposes, we compare our proposed histogram-based features with a general approach using the original data. The well established method, SVM with the Gaussian or Radial Basis Function (RBF) kernel [3], is employed here for comparison. With this, we feed the original tactile data into the SVM directly, and the parameter of the Gaussian function is selected optimally with a grid search. A 4-fold cross-validation is used here for both approaches.

To compare with the SVM with RBF, which is suitable for location specific patterns, we limit our experiment to only the right hand part of the robot. However, despite of this constrained experiment, we have to emphasise that our approach can be in fact useful for larger sensor patches and can be generic to other periodical patterns as well.

We test with the following four basic types of touch on the robot: 1) poking with the index fingertip, 2) bar shape (contact with a full finger), 3) gripping with three fingers (thumb, index, and middle fingers), and 4) gripping with the whole hand. Figure 7 illustrates these touch types.

Among the above 4 types of touch, the first two types present higher variations in locations, while the other two gripping types are more location specific. In other words, poking can be conducted at many different locations. Therefore, for the SVM with RBF using the raw data, it is difficult to obtain sufficient training data in practice. On the other hand, the SVM with RBF method can be more suitable for the two gripping types.

Regarding the setting of the proposed method, there are two parameters that need to be decided in advance: the distance limit r for local structures and the resolution k of vector quantisation. Here, we use empirical values for $r = 8mm$ and $k = 8$ with the consideration of both accuracy and efficiency.

The data are collected by doing the above four types of contact on the KASPAR's right hand for a few minutes each with as many gestures as possible (in total, 1419, 1712, 1697, and 2318 samples for the four types of touch respectively). Because the sensor data are sampled at 50 Hz and the touch events are temporally extended, each touch event has been sampled multiple times, causing many duplicated and similar patterns in the same data set. To have a fair measurement of the performance, we reduce the chance of using too many repeated data for testing the classification accuracy, by resampling the data again at lower sampling rates.

The touch events vary in their duration lengths, which are difficult to be prepared identically in practice. However, based on the initial analysis of the data, the durations are generally shorter than $1s$. Especially, the poke and full finger touch events present much shorter time lengths (e.g. $< 0.5s$) and present higher variances, and the gripping patterns are relatively more stable and last longer. Therefore, we compare two resampled data sets, which are prepared by resampling the original data at every 50 and 10 epochs or $1s$ and $0.2s$ correspondingly, and this results in 2% and 10% of the total data effectively. One purpose of using a smaller training dataset is to simulate the condition, when training data are only prepared from limited places, instead of everywhere on the robot body. In this case, it is expected that the histogram-based method should outperform the general template-based classification, especially for patterns to be identified independent of locations. The data are then normalised before processing. Table I summarises the classification percentages for each touch type individually, where the labels of the touch types are: 1) poking, 2) a full finger touch, 3) grasping with three fingers, and 4) grasping with whole hand respectively. Each is shown with two resampled subsets, as described above.

	Resampled Percentage	SVM + Intersection	SVM + RBF
1	2%	83.3%	61.0%
	10%	95.6%	92.9%
2	2%	70.9%	52.4%
	10%	91.4%	84.4%
3	2%	96.3%	90.5%
	10%	97.1%	97.3%
4	2%	96.6%	92.6%
	10%	97.9%	97.9%

TABLE I: Summary of experimental result.

When using 10% of the original data set, duplicated tactile patterns are reduced to a reasonable level, but still exist, because the average time of touch events is generally longer than half a second. However, keeping similar patterns is particularly necessary for the SVM with RBF for a reasonable measurement of performance. Both the histogram-based approach and the raw data using the SVM with RBF show high accuracy in classifying touch types. This indicates that the histogram representation retains sufficient information for classifying these data even when quite coarse encodings are used for the vector quantisation. On the other hand, although it seems that the histogram-based method has no advantage here, it is notable that this dataset simulated the case when training data is adequately prepared and sampled. Also, it only includes data obtained from the right hand sensor patch, as the SVM with RBF is tied to the data from the same sensor taxels used in obtaining its raw data. For example, the SVM with RBF method using the raw data is not able to classify the poking patterns trained on other parts with testing data from the hand sensors, due to the mismatch of taxel dimensions. This is while our histogram based approach allows to classify touch across the robot body and all sensor patches regardless of the patch used for training.

On the contrary, the performance of the SVM with RBF method is significantly affected, when a small amount of training data (2%) is used, while the histogram-based method performs better. As expected, this is because that the smaller dataset does not contain adequate patterns to train the classifying system, so that some areas on the skin sensor are not covered by the training data. This is the main limit of classification using the raw data and also our motivation to explore the effectiveness of this work. The histogram-based method, in this case, outperforms the other, especially for touch types of poking and one full finger touch. This is as expected and shows the feasibility in classifying patterns invariant at periodical locations. The poorer performance of the full-finger touch is also partially because of the imperfect data collection. To conduct a touch on KASPAR, it takes a short period of time to complete a full finger touch (type 2 here) process, while a part of this touch pattern at the start may produce poking-like patterns (type 1 used here). The similarity between the two touch types, types 1 and 2, seem to effect the classifier as its lower sampling rates, which further support our reasoning for confusion caused between the two due to re-sampling. With the grasping patterns, the difference is not very large, because most grasping gestures are quite limited to certain locations and present less variation in locations.

B. Discussion

Optimal selection of the neighbouring radius r is related to the performance and efficiency of this approach. A larger r would contain more neighbour taxels, and vice versa. The value of it should be decided with the consideration of the minimum size to acquire sufficient local details or features. When $r = 0$, the local structure becomes a single taxel. The histogram is then identical to the grey scale histogram, as in figure 4. However, this results a very flexible statistical feature space that is independent of locations or orientations. The other extreme is when r is larger than the size of the sensor patch. The invariance of local features vanishes completely. All local neighbourhood features become the same as raw data in the global map. In this case, the vector quantization step with the k-means clustering will operate on the original data and lose all properties of location-independence.

The main limitation of this approach is that it discards the order of local regions that do not guarantee uniqueness for pattern matching. In other words, it can be ambiguous in recognising patterns with similar local features but in different orders. In fact, this is a trade-off of two factors: efficiency and accuracy. This can also relate to the selection of the parameter r and k as discussed before. Larger r and k will ensure covering more details, but consequently may also reduce the classification accuracy due to the over-fitting trend in generating the codebook. This approach is rather generic in the sense that the flexible selection of r can model patterns at various levels of detail.

Besides, the work in this paper has presented its periodical location invariance. However, feature invariance is usually discussed in other aspects as well, such as rotation invariance and scale invariance. In fact, rotation invariance can be similarly achieved for every $\pi/3$ by modifying the criteria used for grouping local neighbour taxels, because of the rotational periodical property of the taxel structure. However, scale invariance is not of interest in this work because of the nature of tactile interaction, where there is no scale difference for any particular object in its measured pattern.

VI. Summary

The main purpose of this work was to implement a classification method of tactile patterns that can improve the human-robot tactile interaction capabilities. A large number of tactile sensors have been integrated on a humanoid robot. This paper introduced a novel histogram-based method to classify tactile patterns. It was expected that the histogram representation of a tactile pattern should be invariant at the periodical locations, due to the periodical nature of the sensor layout. For example, index fingers poking at two periodical positions should produce the same histogram representation. One motivation of this work is that it allows classification of tactile patterns with limited training data, e.g. obtained only in a cyclic region.

The method is a two-stage algorithm, which consists of the codebook generation and classification stages, respectively. To generate the codebook, because of the periodical sensor layout, we first extract local neighbourhood structures as the bases of the first layer of the codebook. Next, the pressure distributions over the local neighbourhood features are clustered for vector quantisation to create the sub-layer codewords. Classification is then performed with the histogram representation based on the codebook. The Support Vector Machine (SVM) with the intersection kernel is used for classifying the histogram-based features.

Experimental data are obtained from the robot skin hardware and used to demonstrate the performance of this method, which has shown promising results. The proposed method has been compared with a standard classification method, the SVM method with the RBF kernel using the raw data, and has shown its effectiveness in classifying location invariant patterns.

Despite of some limitations, we have evaluated several common touch types in our application and believe this proposed approach is suitable in most occasions. Also, although the experiment shown in this paper is only limited to one of our robot sensors, it can be generic and efficient to be used on much larger sensors and can be also used with other types of periodical patterns. In summary, the method has shown the suitability of achieving our initial motivation: using the histogram-based feature space to represent patterns that are invariant at periodical locations.

Our future work will investigate optimal selection of the local neighbourhood size r and the number of sub-layer encodings k. We intend to identify frequent touch types used in interaction with children with autism, and train the classifier for selecting the correct touch as an event detection used for action planning in the context of robot assisted play for children with autism.

VII. Acknowledgement

The research is supported by the European Commission's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 231500 – RoboSkin. We would like to thank Dr Ben Robins for valuable discussions on human-humanoid tactile interaction.

References

[1] A. Barla, F. Odone, and A. Verri, "Histogram intersection kernel for image classification," in *Proc. Int. Conf. Image Processing ICIP 2003*, vol. 3, 2003.

[2] A. Bicchi, M. A. Peshkin, and J. E. Colgate, *Handbook of robotics*. Springer, 2008, ch. Safety for Physical Human-robot interaction, pp. 1335 – 1346.

[3] N. Cristianini and J. Shawe-Taylor, *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.

[4] M. R. Cutkosky, R. D. Howe, and P. W. R., *Handbook of Robotics*. Springer, 2008, ch. Force and tactile sensors, pp. 455 – 476.

[5] K. Dautenhahn, C. L. Nehaniv, M. L. Walters, B. Robins, H. Kose-Bagci, N. A. Mirza, and M. Blow, "Kaspar - a minimally expressive humanoid robot for human-robot interaction research," *Applied Bionics and Biomechanics*, vol. 6, no. 3&4, pp. 369 – 397, September 2009.

[6] R. Gray, "Vector quantization," *ASSP Magazine, IEEE*, vol. 1, no. 2, pp. 4 – 29, apr 1984.

[7] M. Hayashi, T. Yoshikai, and M. Inaba, *Sensors: Focus on Tactile Force and Stress Sensors*. InTech, 2008, ch. Development of a Humanoid with Distributed Multi-axis Deformation Sense with Full-Body Soft Plastic Foam Cover as Flesh of a Robot, pp. 319–324.

[8] H. Iwata and S. Sugano, "Human-robot-contact-state identification based on tactile recognition," vol. 52, no. 6, pp. 1468–1477, 2005.

[9] F.-D. Jou, K.-C. Fan, and Y.-L. Chang, "Efficient matching of large-size histograms," *Pattern Recogn. Lett.*, vol. 25, pp. 277–286, February 2004.

[10] Y.-M. Kim, S.-Y. Koo, J. G. Lim, and D.-S. Kwon, "A robust online touch pattern recognition for dynamic human-robot interaction," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 3, pp. 1979–1987, 2010.

[11] T. Miyashita, T. Tajika, H. Ishiguro, K. Kogure, and N. Hagita, "Haptic communication between humans and robots," in *Robotics Research*, ser. Springer Tracts in Advanced Robotics, S. Thrun, R. Brooks, and H. Durrant-Whyte, Eds. Springer Berlin / Heidelberg, 2007, vol. 28, pp. 525–536.

[12] T. Mukai, M. Onishi, T. Odashima, S. Hirano, and Z. Luo, "Development of the tactile sensor system of a human-interactive robot "RI-MAN"," *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 505–512, 2008.

[13] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," in *European Conference on Computer Vision*. Springer, 2006.

[14] A. Petrovskaya, O. Khatib, S. Thrun, and A. Y. Ng, "Bayesian estimation for autonomous object manipulation based on tactile sensors," in *Proc. IEEE Int. Conf. Robotics and Automation ICRA 2006*, 2006, pp. 707–714.

[15] B. Robins, F. Amirabdollahian, Z. Ji, and K. Dautenhahn, "Tactile interaction with a humanoid robot for children with autism: A case study analysis involving user requirements and results of an initial implementation," in *RO-MAN, 2010 IEEE*, Sept 2010, pp. 704 – 711.

[16] B. Robins and K. Dautenhahn, "Developing play scenarios for tactile interaction with a humanoid robot: A case study exploration with children with autism," in *Social Robotics*, ser. Lecture Notes in Computer Science, S. Ge, H. Li, J.-J. Cabibihan, and Y. Tan, Eds. Heidelberg: Springer Berlin, 2010, vol. 6414, pp. 243–252.

[17] A. Schmitz, M. Maggiali, L. Natale, and G. Metta, "Touch sensors for humanoid hands," in *Proc. IEEE RO-MAN*, 2010, pp. 691–697.

[18] A. Schneider, J. Sturm, C. Stachniss, M. Reisert, H. Burkhardt, and W. Burgard, "Object identification with tactile sensors using bag-of-features," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems IROS 2009*, 2009, pp. 243–248.

[19] T. Shibata, T. Tashima, and K. Tanie, "Emergence of emotional behavior through physical interaction between human and robot," in *Proc. IEEE Int Robotics and Automation Conf*, vol. 4, 1999, pp. 2868–2873.

[20] W. D. Stiehl, J. Lieberman, C. Breazeal, L. Basel, L. Lalla, and M. Wolf, "Design of a therapeutic robotic companion for relational, affective touch," in *Proc. IEEE Int. Workshop Robot and Human Interactive Communication ROMAN 2005*, 2005, pp. 408–415.

[21] J. Tegin and J. Wikander, "Tactile sensing in intelligent robotic manipulation – a review," *Industrial Robot: An International Journal*, vol. 32, no. 1, pp. 64–70(7), 2005.

[22] M. Thint and P. P. Wang, "Tactile feature extraction and classification with connectionist models," in *Proc. IJCNN Int Neural Networks Joint Conf*, 1990, pp. 497–502.

[23] K. Wada, T. Shibata, T. Asada, and T. Musha, "Robot therapy for prevention of dementia at home – results of preliminary experiment," *Journal of Robotics and Mechatronics*, vol. 19, no. 6, pp. 691–697, 2007.

[24] Y. Yuan and K. Barner, "An active shape model based tactile hand shape recognition with support vector machines," in *Proc. 40th Annual Conf. Information Sciences and Systems*, 2006, pp. 1611–1616.