Research Article

Low-Resolution Tactile Image Recognition for Automated Robotic Assembly Using Kernel PCA-Based Feature Fusion and Multiple Kernel Learning-Based Support Vector Machine

Yi-Hung Liu,[1] Yu-Tsung Hsiao,[1] Wei-Teng Cheng,[1] Yan-Chen Liu,[2] and Jui-Yiao Su[2]

[1] *Department of Mechanical Engineering, Chung Yuan Christian University, Chungli 32023, Taiwan*
[2] *Mechanical and Systems Research Laboratories, Industrial Technology Research Institute, Hsinchu 31040, Taiwan*

Correspondence should be addressed to Jui-Yiao Su; camussu@itri.org.tw

In this paper, we propose a robust tactile sensing image recognition scheme for automatic robotic assembly. First, an image reprocessing procedure is designed to enhance the contrast of the tactile image. In the second layer, geometric features and Fourier descriptors are extracted from the image. Then, kernel principal component analysis (kernel PCA) is applied to transform the features into ones with better discriminating ability, which is the kernel PCA-based feature fusion. The transformed features are fed into the third layer for classification. In this paper, we design a classifier by combining the multiple kernel learning (MKL) algorithm and support vector machine (SVM). We also design and implement a tactile sensing array consisting of 10-by-10 sensing elements. Experimental results, carried out on real tactile images acquired by the designed tactile sensing array, show that the kernel PCA-based feature fusion can significantly improve the discriminating performance of the geometric features and Fourier descriptors. Also, the designed MKL-SVM outperforms the regular SVM in terms of recognition accuracy. The proposed recognition scheme is able to achieve a high recognition rate of over 85% for the classification of 12 commonly used metal parts in industrial applications.

1. Introduction

In an automated assembly line, information of object (e.g., shape and orientation) is necessary in the robotic manipulation. Based on the information received, a robot can assemble the products using the objects or parts in an automated manner. Previously, vision-based sensing technique (e.g., CCD camera) was often applied to recognize the shape and orientation information of objects in an automated manufacturing line. Although this approach can provide good temporal and spatial resolutions of objects, its recognition accuracy is easily affected by the environment factors such as lighting conditions. When a robot is operated in a dark environment, the visual sensing quality becomes poor. On the contrary, the visual sensing approach may suffer from the light reflection when the environment becomes brighter, especially when the objects to be assembled are made of metal. Moreover, the objects are sometimes hidden from the visual sensors during the manipulation. In contrast, tactile sensing is less

sensitive to these conditions. Therefore, tactile image-based object recognition has received increasing attention form researchers and engineers over the past decade [1–7].

When the tactile sensing approach is adopted, a two-dimensional tactile sensing array consisting of multiple sensing elements is attached to a robotic hand or finger. When the robotic finger touches an object, each sensing element in the tactile array measures the contact force or pressure applied on a specific and small area of the object. The pressure values of the sensing elements are then transformed into integer ones within the range of [0, 255], thus forming a pseudoimage in which the gray-level values are the transformed pressure values. Based on the pseudoimage (known as the tactile image), a system can recognize the shape, edge direction, and contour of the object. The recognition results are the inputs to the robot which performs the task of automated assembly.

Previous works mainly solve the problem where the size of object is larger/much larger than the tactile sensing array by way of edge tracking/following [1–8]. First, a system/robot

guides the sensing array to move along the contour of an object. After scanning the whole contour of the object, the system estimates the shape of the object by using the information of the collected tactile images during the contour scanning. In addition to this case, there is another: the object is smaller than the tactile sensing array. In this case, the shape of an object can be identified through the use of one single tactile image. However, due to the following factors, the task of object shape recognition becomes difficult and challenging.

(1) *Low Resolution.* Usually, the size of a tactile sensing array which can be attached to a robotic finger is small, which means that the number of tactile sensing elements in the array is very limited. In other words, the acquired tactile image is of low spatial resolution.

(2) *Diffusion Effect.* In order to project the tactile sensor array, a thin cover (plastic or silica) is usually placed on the array. Although the cover can prevent the tactile array from being damaged when the robotic finger presses an object, the force that the object applies to one sensing element spreads to its neighboring sensing elements simultaneously. The diffusion effect makes the acquired tactile images ambiguous, especially the edges of the objects.

(3) *Fence Effect.* Unlike the sensing elements in a CCD camera where the sensing elements are closely adjacent to each other, there exists a large and noticeable gap between sensing elements in a tactile sensing array. These gaps make the object in a sensing tactile image look like fenced.

Due to the factors above, it is difficult to identify the shape of an object through the tactile image. To achieve a high-reliability automated robotic assembly, it is thus necessary to develop a high-accuracy tactile image recognition scheme. To the end, we propose in this paper a scheme, which is composed of three main layers. Initially, an image preprocessing procedure is performed to enhance the contrast of the tactile images. In this layer, geometric features and Fourier descriptors are first extracted from a given image. The extracted geometric features and the Fourier descriptors form a feature vector, which is high-dimensional and does not necessarily achieve satisfactory recognition accuracy. Kernel principal component analysis (kernel PCA) [9] is a powerful kernel method for pattern representation. It computes higher order statistics among random variables while reducing the data dimensionality, thus being able to achieve the goal of both feature extraction and dimensionality reduction. Kernel PCA has shown success in various pattern recognition problems, such as face recognition [10] and defect inspection [11]. Therefore, in this paper, we apply the kernel PCA to reduce the dimensionality and extract more discriminating features from the feature vector extracted from the tactile images. Finally, in the third layer, support vector machine (SVM) [12] is performed to recognize the shape of the object in the input tactile image. In order to improve the generalization performance of SVM, we introduced the multiple kernel learning (MKL) algorithm [13] to the regular SVM. Regular SVM uses only one single kernel to learn the classifier. Users have to determine the kernel function type and its optimal parameter, which is not only time consuming but also suboptimal because one single kernel may not lead to satisfactory recognition accuracy, especially when the classification problem is complex. Instead of using one single kernel, MKL proposes that an ideal kernel should be a combination of multiple kernels (i.e., base kernels). Based on this idea, MKL trains an SVM with a mixed kernel. In this study, we combine the MKL and the SVM to training a robust classifier for tactile image recognition. Experimental results show that even though the tactile images are of low resolution and suffer from the diffusion and fence effects, the proposed scheme is still able to achieve a high recognition accuracy of over 85% on 12 types of objects.

The rest of this paper is organized as follows. In Section 2, the details of the designed tactile sensing array used in this study as well as the tactile image collection procedure are given. The proposed tactile image recognition scheme is introduced in Section 3. Results and discussion are provided in Section 4. Finally, we conclude this study in Section 5.

2. Tactile Sensor and Image Acquisition

2.1. Material and Manufacturing Process. The piezoresistive layer of sensor is a functional material [14, 15]. It is manufactured by mixing the nanoparticles of carbon and silica into an insulation polymer matrix with high concentration. In this study, we design and fabricate a flexible tactile sensor array by means of screen printing technology. The sensor structure is composed of two $100 \, \mu m$ thick subtracts of Polyethylene Terephthalate (PET) films (double-sided electrode structure) and one adhesion layer. Screen printing process is conducted to deposit the silver ink, piezoresistive ink, and adhesion resin on the PET films, respectively, as illustrated in Figure 1. We also design the raise structure to enhance the sensitivity of the tactile sensor. The raised structure is fabricated on the face sheet of the sensor using a UV curable adhesive and has led to high sensitivity due to stress concentration. Similar to the sensing elements on modern touchscreens, these cells can be used for contact localization and pressure measurement. The developed tactile sensor array contains 100 taxels (i.e., 100 tactile sensing elements) in a 10×10 configuration as shown in Figure 2.

2.2. Characteristics of Tactile Sensor Cells. The pressure-piezoresistivity characteristics of the proposed sensor were measured by a customized instrument developed in LabVIEW environment which includes a pressure chamber, multifunction switch/measure unit (Agilent 34980A), and a National Instruments data acquisition (NI DAQ) card. In the calibrating process, the sensor was placed in the chamber, subjected to a static uniform load (to make sure each cell faced the same pressure). The pressures in chamber were controlled by a LabVIEW interface, and the measured data was scanned by Agilent 34980A and recorded via the NI DAQ card. Figure 3 shows the measuring device and a measuring result of one tactile sensor cell. The pressure range
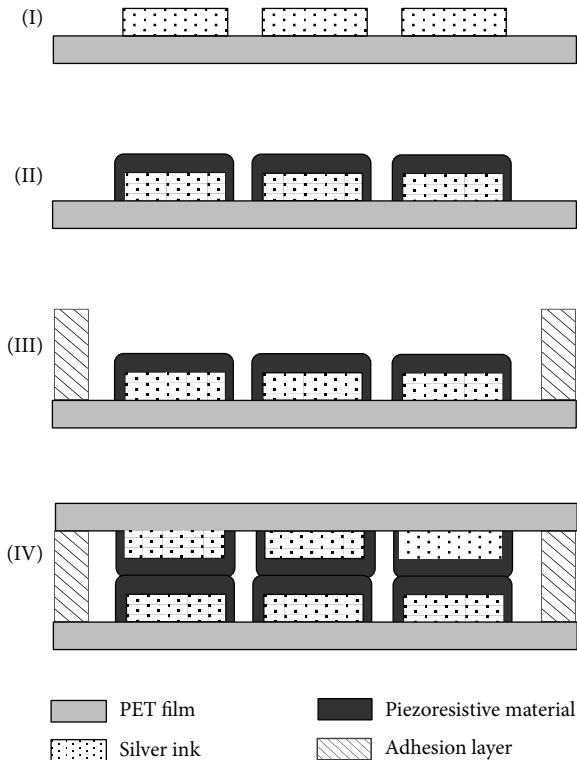
FIGURE 1: Fabrication processes of tactile sensor arrays using screen printing technology. (I) Print the row and column electrodes on the PET films, respectively. (II) Print the piezoresistive material. (III) Bottom PET film with adhesion resin. (IV) The top and bottom PET films are laminated into a large area tactile array sensor.

of the tactile sensor was measured from 10 to 580 psi and the linear relationship between pressure and conductivity in each cell is acceptable.

2.3. Design of Experiment for Image Acquisition. To determine the characteristic of the sensor cell, the experimental setting in Figure 4 was used. It mainly consists of a controlled linear actor using a linear motion stage in conjunction with a high precision servo motor and a load cell. The testing object is placed between the linear actuator's indenter and the tactile sensor array. By positioning the indenter, a defined force can be applied to the testing object, due to the defined flexibility of the load cell. For each sensing element, the sensor material changes its resistance itself as the normal stress changes. The resistivity is measured using a data acquisition (DAQ) module for signal transduction. Each sensing element then provides signals carrying information about the local value of the normal stresses, thus forming a 10×10 tactile image of the contacting object. In this paper we use the tactile image as input to a pattern recognition system, which should classify the contacting object according to its image features.

In addition, to justify the stress applied to collect appropriate tactile images, five different loads (1 kgf, 2 kgf, 3 kgf, 4 kgf, and 5 kgf) were generated by the indenter. Raw images of a bar shape object with fixed cover under the five different loads are shown in Figure 5. The experimental results show

the following: (1) while applying small force (1~3 kgf), the tactile images of the testing object are usually incomplete, due to the uneven press on the contact surface; (2) on the other hand, due to the elastic cover layer, the tactile data present high distortion while applying a too heavy load. According to the observations above, the force of 4 kgf is a suitable one for collecting tactile images.

The contact behavior is mainly determined by the surface flatness and roughness between two objects. That is, local stress will be concentrated on the first contacting area, and this phenomenon will lead to fragment of tactile image. To avoid this phenomenon, we place an elastic cover on a tactile sensor as buffer layer. Several commercially available cover layers with similar hardness were examined under the loading of 4 kgf, as shown in Figure 6. The experimental results show that thicker cover decreases both spatial and force resolutions. However, if the cover is too thin, the tactile image will be more fragile. Accordingly, the 0.9 mm thick silicone sheet was chosen as the cover in this study. Even though the chosen cover is more suitable than others, the diffusion effect mentioned in Section 1 would still happen.

Moreover, the tactile sensor array is fabricated on a flexible film (20 mm × 20 mm × 250 μm) and contains 100 sensing elements. Each element has a 1.3 mm × 1.3 mm sensing area and a 0.7 mm width spacer in each direction. The layout of the designed tactile sensing array is shown in Figure 7. The spacer refers to the gap between elements. As mentioned in Section 1, such a large spacer results in the fence effect.

2.4. Object Types. In this study, 12 mental objects with different shapes and sizes are designed as the testing objects. Samples of the designed objects are shown in Figure 8, and their descriptions are listed in Table 1. These objects are the parts commonly used in the manufacturing and are smaller than the tactile array. When an object is placed on the tactile array, a 4 kgf force is applied on the object. Notice that the cover is placed between the object and the tactile array. When the corresponding tactile image is acquired, the same object is reput on the tactile array with a slightly different position and direction (orientation) in order to get another image for the same object. By repeating this procedure 20 times, we collect 20 images for each object. Therefore, in this study, the number of classes is 12, and for each class we prepare 20 tactile images. Figure 9 displays examples of the tactile images of the 12 different objects.

As can be observed from these examples, the spatial resolution of the tactile image is extremely low, and it is very difficult to discriminate between objects by observation. For example, the object in the last image (i.e., the sixth image) of the first row and the one in the last image of the second row are originally different: the former is a solid hexagon, while the latter is a solid circle. However, due to the low resolution and the aforementioned diffusion and fence effects, the two different objects in the two images look very similar and are thus difficult to discriminate. Therefore, a robust recognition scheme is required. In the following, we introduce our recognition scheme in detail.
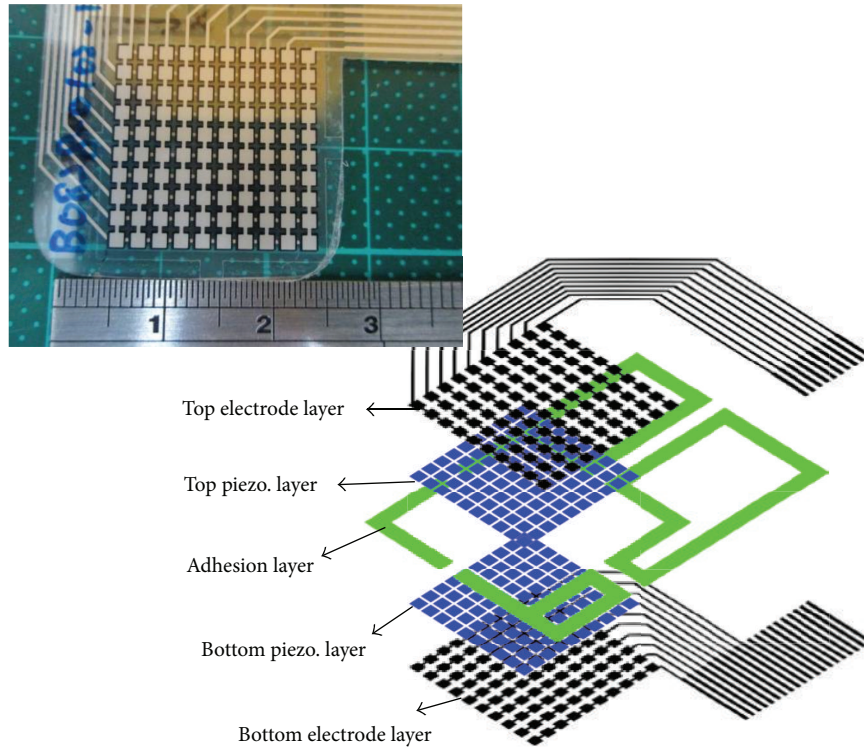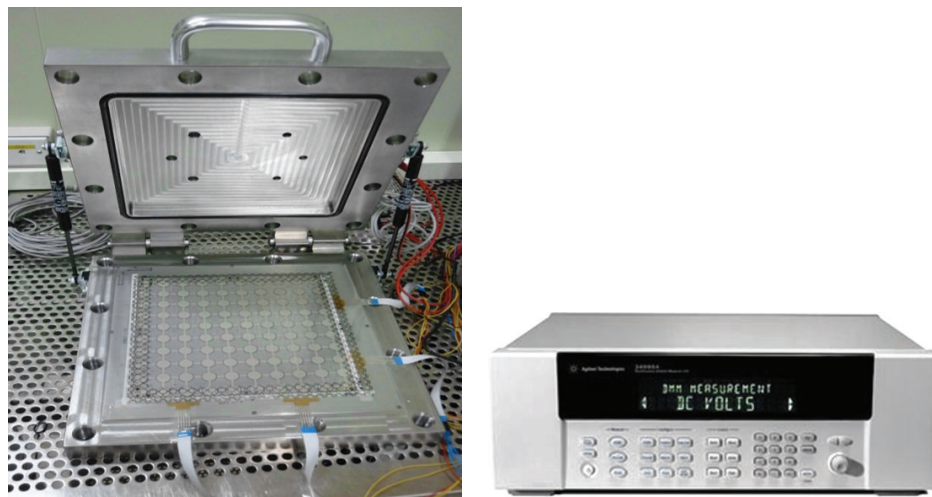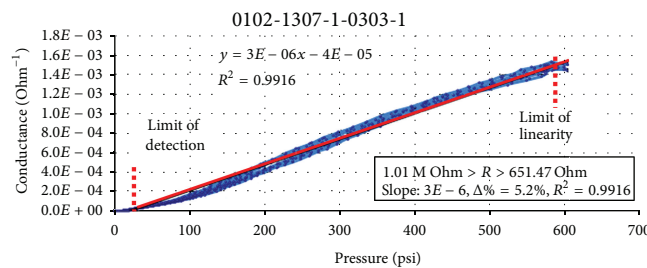
Top electrode layer ←

Top piezo. layer ←

Adhesion layer ←

Bottom piezo. layer ←

Bottom electrode layer ←

FIGURE 2: Tactile sensor array was fabricated on a flexible film (20 mm × 20 mm × 250 μm) and contains 100 sensor cells (1.3 mm × 1.3 mm).

(a)

(b)

0102-1307-1-0303-1

$y = 3E - 06x - 4E - 05$

$R^2 = 0.9916$

Limit of detection

Limit of linearity

1.01 M Ohm > R > 651.47 Ohm
Slope: $3E - 6$, $\Delta\% = 5.2\%$, $R^2 = 0.9916$

(c)

FIGURE 3: Sensor array calibration device and the multifunction switch/measure unit. (a) The pressure chamber. (b) Agilent 34980A. (c) The pressure-piezoresistivity characteristics of one cell (position (3, 3)).
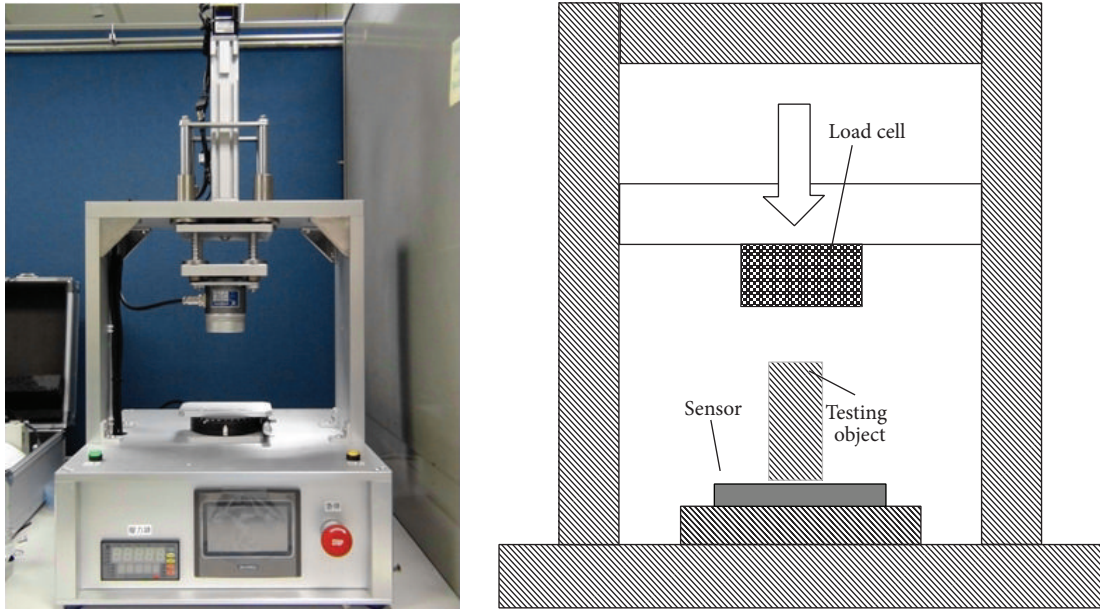
FIGURE 4: Pressure testing machine and experimental setting.

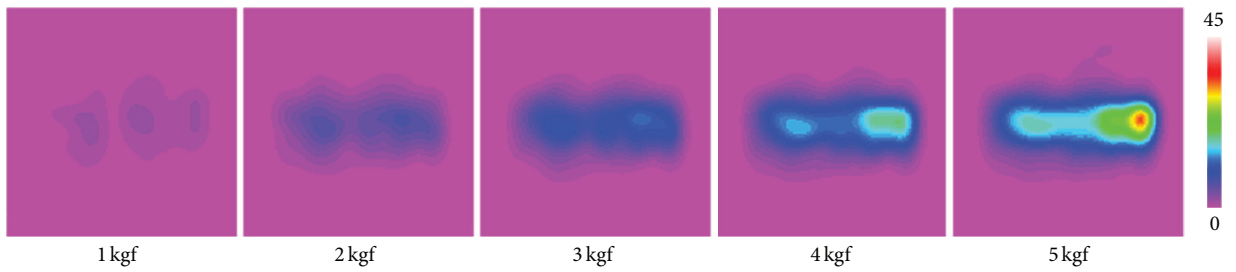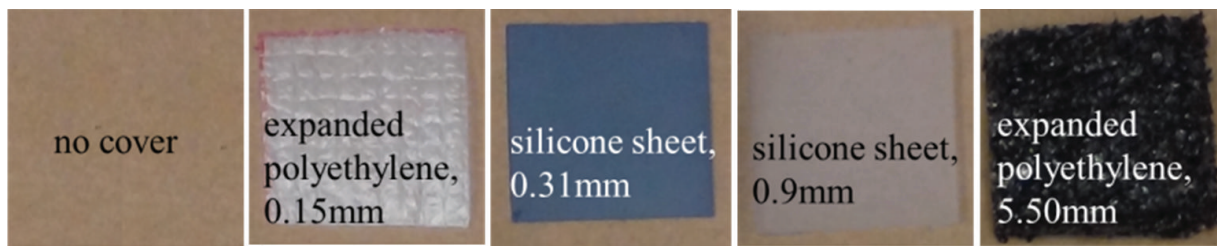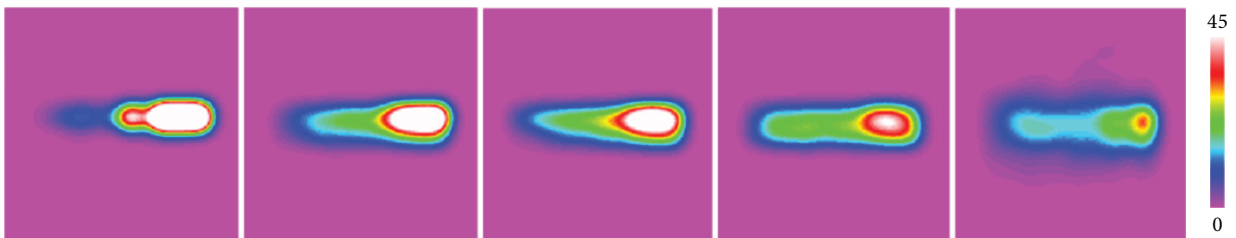| 1 kgf | 2 kgf | 3 kgf | 4 kgf | 5 kgf |

FIGURE 5: Raw images of a bar shape object with fixed cover under various loads 1 kgf~5 kgf.

(a)

(b)

FIGURE 6: Raw images of a bar shape object with various covers under fixed load 4 kgf.
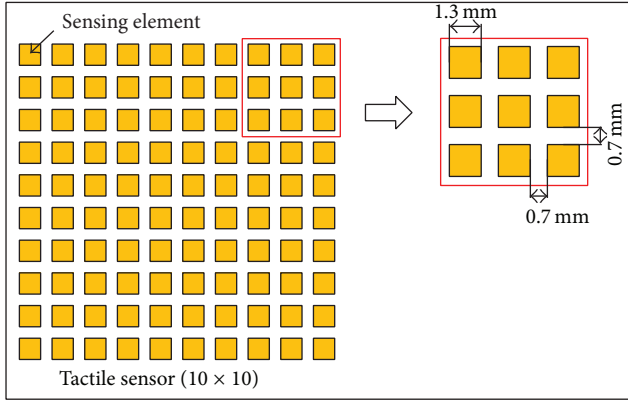
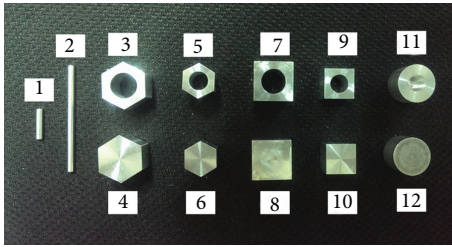Figure 7: Layout of the designed tactile sensor array.

Figure 8: Samples of the 12 objects to be recognized in this study.

Table 1: Descriptions of the 12 objects.

Object number	Description
1	Bar shape with 10 mm length
2	Bar shape with 35 mm length
3	Hexagon with flat size of 13 mm and a Φ8 mm hollow hole
4	Solid hexagon with flat size of 13 mm
5	Hexagon with flat size of 10 mm and a Φ6 mm hollow hole
6	Solid hexagon with flat size of 10 mm
7	Square with flat size of 13 mm and a Φ8 mm hollow hole
8	Solid square with flat size of 13 mm
9	Square with flat size of 10 mm and a Φ6 mm hollow hole
10	Solid square with flat size of 10 mm
11	Φ13 mm circle with a Φ8 mm hollow hole
12	Solid Φ13 mm circle

3. Proposed Tactile Image Recognition Scheme

3.1. Layer 1: Image Preprocessing. Each tactile image is originally a pixel matrix of 10×10. One example is shown as Figure 10(b). In order to increase the spatial resolution, each image is resized to a 33×33 image by linear interpolation. The resized image of Figure 10(b) is shown as Figure 10(c). As can be observed from Figure 10(c), the resized image has a low contrast. Therefore, Gamma correction [15] is further applied to enhance the contrast of each resized image (see Figure 10(d)). However, the gray levels of pixels near the edge of object and the ones of few isolated pixels inappropriately become higher. To eliminate such noises, a statistical filtering method is performed. Let μ and sd be the mean and the stand deviation of the gray levels of the image. Then, the gray level of one pixel is replaced by zero if the gray level of this pixel is below the threshold T, where $T = \mu - 2 \times$ sd. The image after the statistical filtering-based noise removal is shown as Figure 10(e). Finally, the processed gray-level image is transformed to a binarized image through Otsu thresholding [16]. It should be noticed here that when the force is applied to the test object, the force is not necessarily uniformly distributed over the tactile sensing array. Therefore, directly performing the thresholding on the entire image may not lead to an ideal binarized image. If the applied force is highly nonuniformly distributed over the contact surface between the tactile sensing array and the test object, only part of the object will appear after thresholding the entire image according to our preliminary test on the images collected. Thus, our strategy is to partition the image into $n \times n$ subimages with equal size and then perform the Otsu thresholding on each subimage independently. We found that when n is set as 2, the highest recognition accuracy can be obtained. The binarized results of 1×1, 2×2, and 3×3 are shown in Figures 10(f), 10(g), and 10(h), respectively.

3.2. Layer 2: Kernel PCA-Based Feature Fusion

3.2.1. Geometric Features. Two kinds of geometric features are extracted from each binarized image: area and edge-to-mean variance (called variance hereafter). Area denotes the number of pixels labeled as 1 in the binarized image. To compute the variance, we first detect the edge points and the centroid of the object within one binarized image and then compute the distance between each edge point and the centroid. Finally, the variance of the computed distances is calculated.

3.2.2. Fourier Descriptors. To compute the Fourier descriptors, the boundary extraction algorithm [17] is performed to find the Cartesian coordinates of the sequential boundary pixels of the object in an image. Examples of the images after the boundary extraction are shown in Figure 11.

Suppose that the Cartesian coordinates of the boundary pixels of an object are $(x[m], y[m])$, $m = 1, 2, \ldots, L$, where L denotes the number of boundary pixels. The Fourier series expansion of the boundary pixels is as follows:

$$x[m] = \sum_{n=-\infty}^{\infty} a[n] e^{jn\omega_0 m}, \tag{1a}$$

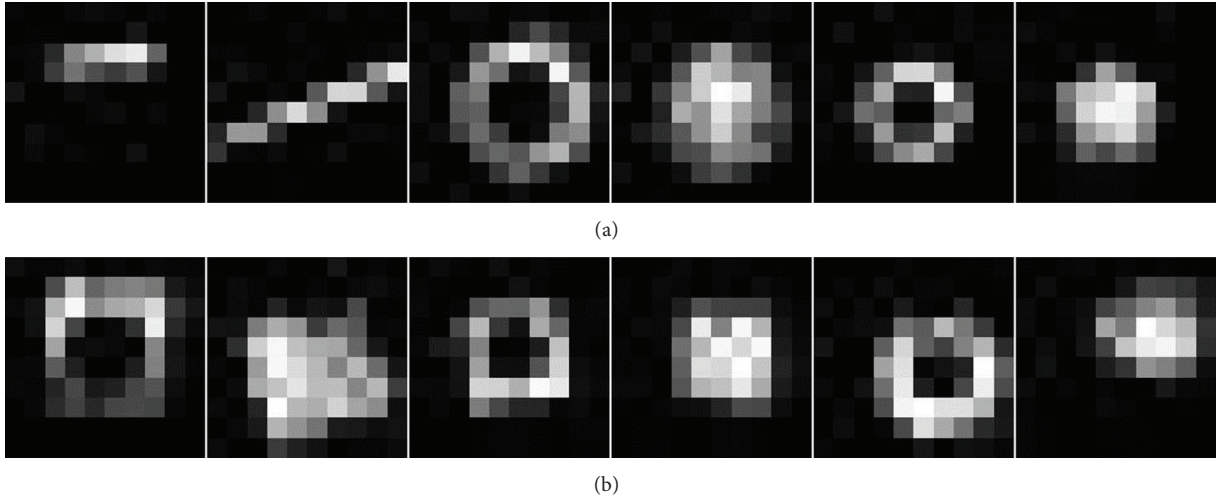$$y[m] = \sum_{n=-\infty}^{\infty} b[n] e^{jn\omega_0 m}, \tag{1b}$$

(a)

(b)

FIGURE 9: Examples of the tactile images. The images in the first row of this figure are the tactile images of the first six objects (class 1–class 6), respectively. The second row displays the examples of the tactile images of class 7–class 12, respectively. Each image is a 10-by-10 gray-level matrix.

(a)   (b)   (c)   (d)

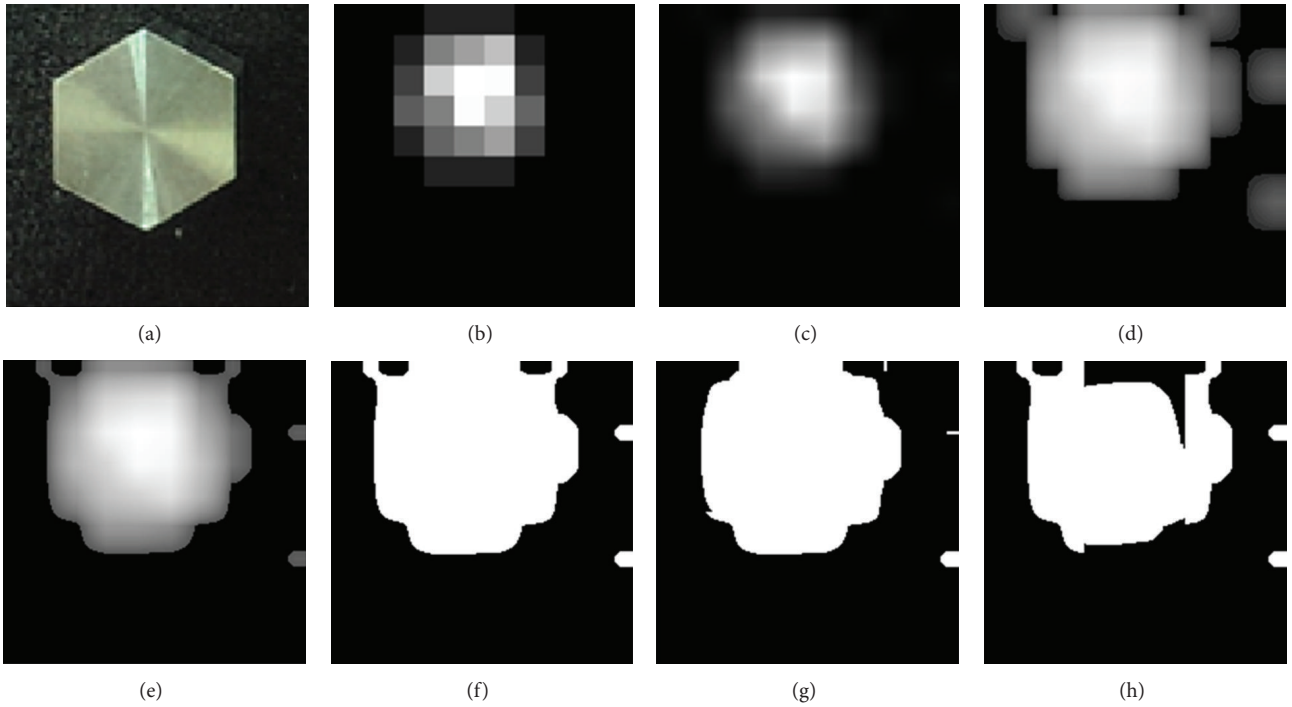

(e)   (f)   (g)   (h)

FIGURE 10: An illustrative example for the image preprocessing stage, where (a) is the testing object, (b) is the corresponding 10-by-10 tactile image, (c) is the resized 33-by-33 image, (d) is the image after Gamma correction-based contrast enhancement, (e) is the image after noise reduction, and (f)–(h) are the binarized images with 1×1, 2×2, and 3×3 image partitioning-based thresholding.

where $\omega_0 = 2\pi/L$ and $a[n]$ and $b[n]$ are Fourier coefficients:

$$a[n] = \frac{1}{L} \sum_{m=1}^{L} x[m] e^{-jn\omega_0 m}, \quad n = 1, \ldots, L, \tag{2a}$$

$$b[n] = \frac{1}{L} \sum_{m=1}^{L} y[m] e^{-jn\omega_0 m}, \quad n = 1, \ldots, L. \tag{2b}$$

The Fourier descriptors $s[n]$ are given by

$$s[n] = \frac{r[n]}{r[1]}, \quad n = 1, \ldots, L, \tag{3}$$

where $r[n] = \sqrt{|a[n]|^2 + |b[n]|^2}$. The Fourier descriptors are translation-, rotation-, and scaling-invariant. Actually, not all of the L Fourier descriptors are required: only the first n descriptors are necessary and perform the best. After
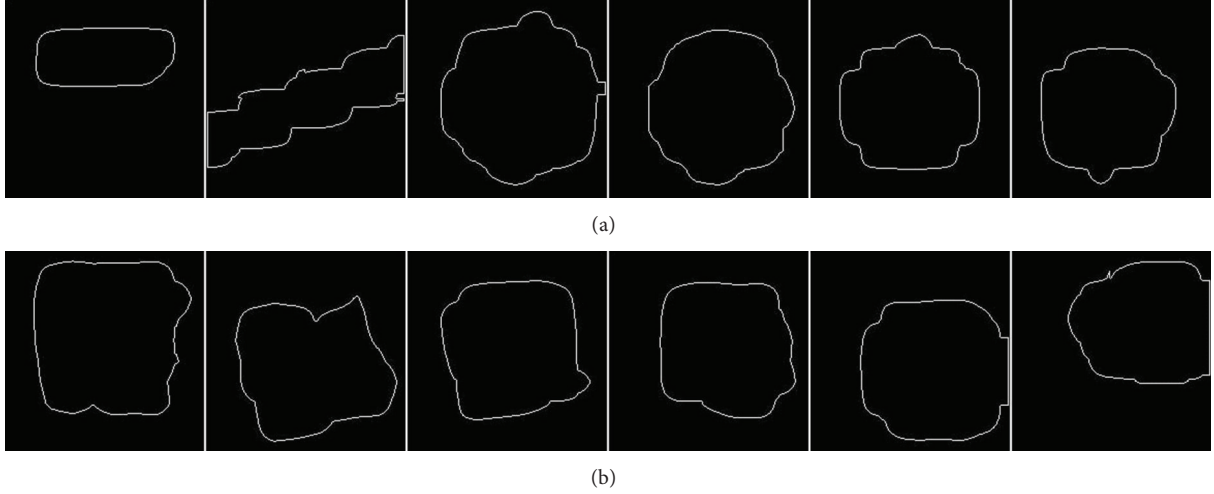
(a)

(b)

FIGURE 11: Examples of boundary extraction result. The corresponding gray-level tactile images are displayed in Figure 9.

the feature extraction, each tactile image is represented by a vector of $n + 2$ dimension, in which two are geometric features and the rest are the extracted Fourier descriptor. To facilitate the following illustration, the feature vectors are simply called data hereafter.

3.2.3. Kernel PCA-Based Feature Transformation.

The kernel PCA feature fusion consists of a training phase and a testing phase. Suppose that there is a set of training data $\mathbf{x}_i \in R^m$, $i = 1, \ldots, M$. Kernel PCA maps the training data into a higher-dimensional feature space F using a nonlinear mapping $\phi : R^m \rightarrow F$, where $m = n + 2$, and then centers these mapped data such that they have a zero mean: $\sum_{i=1}^{M} \phi(\mathbf{x}_i) = \mathbf{0}$ (see [9] for a detailed derivation of the data centering method in the feature space). In the feature space, kernel PCA solves the following eigenvalue problem:

$$\lambda \mathbf{v} = \Gamma \mathbf{v}, \tag{4}$$

where $\mathbf{v} \in F$ are eigenvectors associated with nonzero eigenvalues λ and $\Gamma = 1/M(\sum_{i=1}^{M} \phi(\mathbf{x}_i)\phi^T(\mathbf{x}_i))$ is the mapped-data covariance matrix. By introducing the kernel function: $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_i)$, the dual problem of (1a) and (1b) is as follows:

$$M\lambda \mathbf{a} = \mathbf{K}\mathbf{a}, \tag{5}$$

where $\mathbf{K} : K_{ij} \equiv K(\mathbf{x}_i, \mathbf{x}_j)$ is a $M \times M$ kernel matrix and $\mathbf{a} = (a_1, \ldots, a_M)^T$ is the eigenvector associated with $\lambda_l \neq 0$ and is subject to the normalization condition $\|\mathbf{a}\|^2 = 1/\lambda$. Solving the eigenvalue problem expressed as (5) yields M eigenvectors $\mathbf{a}^k$, $k = 1, \ldots, M$. However, we select only the first d leading eigenvectors as the basis for transformation. The number of chosen eigenvector should be smaller than the number of total features and the number of training data, that is, $d < m$ and $d < M$, and the optimal number of the eigenvectors should be experimentally determined. By doing so, the goal of dimensionality reduction can be achieved.

After the eigenvector selection, the training phase of kernel PCA is completed.

In test phase, the projection of testing data $\mathbf{x} \in R^m$ onto the kth eigenvector $\mathbf{v}^k$ is computed by

$$z^k = \mathbf{v}^k \cdot \phi(\mathbf{x}) = \sum_{i=1}^{M} a_i^k K(\mathbf{x}_i, \mathbf{x}), \quad k = 1, \ldots, d, \tag{6}$$

where a_i^k is the ith component of the kth eigenvector $\mathbf{a}^k$, and z^k is the nonlinear principal component of $\mathbf{x}$ corresponding to the nonlinear mapping ϕ. The d nonlinear principal components constitute a vector $\mathbf{z} = (z^1, \ldots, z^d)^T$, which is the nonlinear fusion of the geometric features and the Fourier descriptors. In this study, the Gaussian function $K(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2/2\sigma^2)$ is chosen as the kernel, where σ is a user-specified kernel parameter and can be optimized by using a cross validation procedure.

3.3. Layer 3: Multiple Kernel-Based SVM Classification

3.3.1. SVM.

Given a training set $\{\mathbf{z}_i, y_i\}$, $i = 1, \ldots, M$, where $\mathbf{z}_i \in R^d$ are training data and $y_i \in \{-1, +1\}$ are class labels, SVM maps the data into a higher-dimensional feature space and then finds an optimal separating hyperplane (OSH) which maximizes the margin of separation and minimizes the training errors simultaneously, which can be formulated as the constrained optimization problem as

$$
\begin{aligned}
\text{Minimize} \quad & \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{M} \xi_i \\
\text{subject to} \quad & y_i\left(\mathbf{w}^T\phi(\mathbf{z}_i) + b\right) - 1 + \xi_i \geq 0, \quad \forall i \\
& \xi_i \geq 0, \quad \forall i,
\end{aligned}
\tag{7}
$$

where $\mathbf{w}$ denotes the weight vector of the hyperplane, b is the bias of the hyperplane, ξ_i are slack variables representing training errors, and C is a penalty weight. The value of C needs

to be specified in prior. Introducing the Lagrangian to (7) yields the dual problem:

$$\text{Maximize} \quad \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i \alpha_j y_i y_j K\left(\mathbf{z}_i, \mathbf{z}_j\right)$$

$$\text{subject to} \quad 0 \leq \alpha_i \leq C, \quad \forall i \tag{8}$$

$$\sum_{i=1}^{M} \alpha_i y_i = 0,$$

where α_i are Lagrange multipliers. The training data for which $0 < \alpha_i \leq C$ are called support vectors (SVs). The class label for a test data $\mathbf{z}$ is computed by the decision function:

$$D\left(\mathbf{z}\right) = \text{Sign}\left(\sum_{\mathbf{z}_i \in \text{SV}} \alpha_i y_i K\left(\mathbf{z}_i, \mathbf{z}\right) + b_o\right), \tag{9}$$

where b_o is the optimal bias of the OSH, which can be calculated by taking any support vectors whose corresponding Lagrange multipliers satisfy $0 < \alpha_i < C$ into the Kuhn-Tucker conditions. If $D(\mathbf{z}) > 0$, $\mathbf{z}$ is classified as a positive data or negative data otherwise.

3.3.2. Multiple Kernel SVM. MKL is a data-driven learning algorithm which learns kernel from the given training data [13]. It assumes that an ideal kernel is a linear combination of predefined base kernels:

$$K\left(\mathbf{x}_i, \mathbf{x}_j\right) = \sum_{k=1}^{N} d_k K_k\left(\mathbf{x}_i, \mathbf{x}_j\right), \tag{10}$$

where K_k are base kernels, d_k are kernel combination weights, and N is the number of chosen base kernels. MKL-SVM solves the following optimization problem:

$$\text{Minimize} \quad \sum_{k=1}^{N} \frac{1}{2} \frac{\|\mathbf{w}_k\|^2}{d_k} + C \sum_{i=1}^{M} \xi_i$$

$$\text{subject to} \quad y_i\left(\sum_{k=1}^{N} \mathbf{w}_k^T \phi_k\left(\mathbf{z}_i\right) + b\right) - 1 + \xi_i \geq 0, \quad \forall i,$$

$$\xi_i \geq 0, \quad \forall i, \tag{11}$$

$$\sum_{k=1}^{N} d_k = 1, \quad d_k \geq 0, \forall k.$$

The dual problem of (11) is written as

$$\text{Maximize} \quad \sum_{\alpha, \gamma}^{M} \sum_{i=1}^{M} \alpha_i - \gamma$$

$$\text{subject to} \quad 0 \leq \alpha_i \leq C, \quad \forall i,$$

$$\sum_{i=1}^{M} \alpha_i y_i = 0, \tag{12}$$

$$\frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i \alpha_j y_i y_j K_k\left(\mathbf{z}_i, \mathbf{z}_j\right) \leq \gamma, \quad \forall k.$$
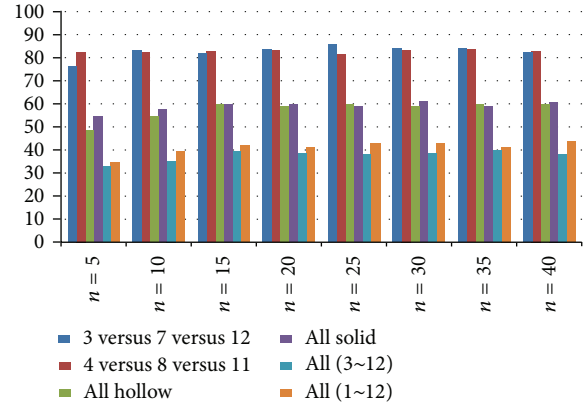

FIGURE 12: k-NN recognition accuracies of Fourier descriptors among different class combinations.

Finally, for a test data point $\mathbf{z}$ its class label is determined by the MKL-SVM decision function:

$$f\left(\mathbf{z}\right) = \text{Sign}\left(\sum_{i=1}^{M} \sum_{k=1}^{N} \alpha_i y_i d_k K_k\left(\mathbf{z}_i, \mathbf{z}\right) + b\right). \tag{13}$$

In this paper, we solve the optimal values of α_i, d_k, and b by using the Simple MKL by Rakotomamonjy et al. [18] for it adopts the reduced gradient method to solve the MKL optimization problem, which is computationally cheaper than other MKL solvers.

4. Results and Discussion

In this section, we first test the recognition accuracies of the geometric features or properties (GP) and Fourier descriptors (FD) using a simple classifier, that is, the k-nearest neighbor (k-NN) classifier, to find the optimal number of FDs, where k is set as 3. A ten-run twofold cross validation procedure is performed to test the recognition accuracy on the data set. The data set contains 240 data belonging to 12 classes, and each class has 20 data. The results are shown in Figures 12–14. Notice that (3 versus 7 versus 12) means that the number of objects to be classified is 3, including class 3, class 7, and class 12. Also, "all hollow" means that the objects to be classified are the ones with hollow holes as shown in Figure 8, including class 3, class 5, class 7, class 9, and class 11, totally 5 classes. On the other hand, "all solid" contains 5 different classes, including class 4, class 6, class 8, class 10, and class 12. "All (1~12)" means that the number of classes is 12 (from class 1 to class 12).

It can be seen from Figure 12 that different class combination results in different accuracy. For example, in the case of "All solid," FD with $n = 40$ gives a recognition accuracy of 60.8%. For the same case, GP results in recognition accuracy of 74.45% (see Figure 13). When the two kinds of features are combined, a higher recognition accuracy of 84.3% occurs, as shown in Figure 14. Moreover, if the objective is to classify all types of objects, that is, the case of All (1~12), the recognition accuracy of FD with $n = 40$ is only 43% (see Figure 12), and the GE gives only the recognition accuracy
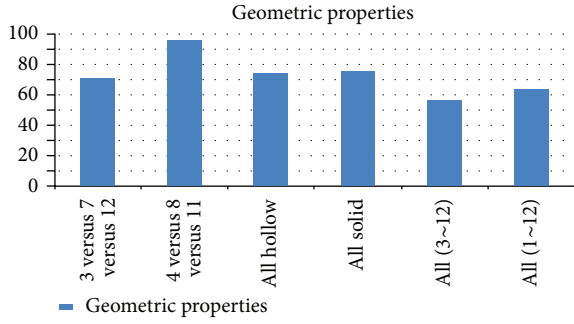
FIGURE 13: k-NN recognition accuracies of geometric features among different class combinations.
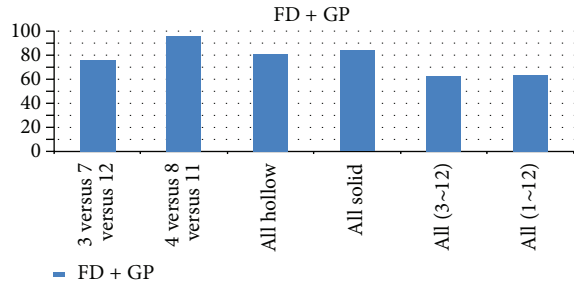

FIGURE 14: k-NN recognition accuracies of geometric features and Fourier descriptors among different class combinations.

of 63.3 (see Figure 13). However, when the two kinds of features are combined, the accuracy can be enhanced to 68.69%, as indicated in Figure 14. These comparisons show that the combination of GE and FD is able to achieve higher recognition accuracy than any of them. In addition, we can also observe from Figure 12 that the recognition accuracy saturates as n reaches 20. Moreover, when $n = 40$, FD gives the best result. Therefore, we set $n = 40$ in the following experiments.

Next, we test the proposed recognition scheme (combination of the kernel PCA-based feature fusion and MKL-SVM) and compare the proposed scheme with other combinations. Similarly, the 10-run twofold cross validation is performed to optimize the parameters of the methods. For kernel PCA, the parameters to be optimized include the kernel parameter and the number of eigenvectors. The parameters of SVM are the penalty weight C and the kernel parameter σ. For MKL-SVM, not only the penalty weight C needs to be adjusted, but also the based functions need to be determined in advance. Choosing a set of good kernels as the base kernels is crucial to the MKL-SVM. Accordingly, two kinds of frequently used kernel functions are adopted as the base kernels in this study: Gaussian function and the polynomial function:

$$K_G(\mathbf{x}, \mathbf{y}) = \exp\left(-\|\mathbf{x} - \mathbf{y}\|^2 / 2\sigma^2\right),$$

$$K_P(\mathbf{x}, \mathbf{y}) = \left(1 + \mathbf{x}^T \mathbf{y}\right)^p,$$

(14)

where p is the power of the polynomial kernel. In order to get a good combination of kernels, the range of σ is

TABLE 2: Comparison of recognition accuracies among different methods.

Feature	Classifier	Recognition accuracy (in %)
FD + GE	k-NN	68.69
FD + GE	SVM	76.17
Kernel PCA-based feature fusion	SVM	82.13
Kernel PCA-based feature fusion	MKL-SVM	85.54

set as wide as possible: $\sigma = \{2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^{0}, 2^{1}, 2^{2}, 2^{3}, 2^{4}, 2^{5}, 2^{6}, 2^{7}, 2^{8}, 2^{9}, 2^{10}\}$. Also, the values of p are in the set of $\{1, 2, 3, 4, 5, 6\}$. There are totally 16 Gaussian kernels and 6 polynomial kernels used in the MKL-SVM. Therefore, the number N of base kernels is 22 in this experiment. In addition, since the objective of this study is to classify 12 different objects, we only consider this case in this experiment. Since there are 12 classes to be classified, the binary classifier SVM and MKL-SVM needs to be extended to multiclass classifiers. In this study, the one-against-one method combined with the voting strategy [19] is adopted for this purpose. Finally, the cross validation results are listed in Table 2.

As can be seen from Table 2, when feature FD + GE is used, SVM largely improves the recognition accuracy. However, the accuracy of 76.17% is still unacceptable when the method is applied to automated robotic assembly. When the kernel PCA is further applied to FD + GE, namely, the kernel PCA-based feature fusion, the recognition accuracy is significantly improved from 76.17% to 82.13%, which demonstrates the validity of the proposed feature fusion scheme in improving the tactile image recognition. Further, when the classifier is replaced by MKL-SVM, an accuracy improvement of 3.41% (85.54% − 82.13%) is observed. Although this difference appears to be small, the error reduction ratio is large: 3.41/(100 − 82.13) = 19.08%. Therefore, we can conclude that MKL-SVM is more suitable than the widely used SVM for tactile image classification.

5. Conclusion

In this paper, we have presented a recognition scheme for solving the difficult tactile image recognition problem, which plays a critical role in automated robotic assembly. The proposed kernel PCA-based feature fusion technique largely improved the recognition accuracy of the frequently used geometric featured and Fourier descriptors, and the multiple kernel learning (MKL)-based SVM can perform much better than the regular SVM in terms of object recognition through the use of tactile image. Experimental results have indicated the effectiveness of the proposed recognition scheme in tactile image recognition. Nevertheless, there remain several worth-studying issues that may further improve the current results. For example, other types of kernels can be included in the MKL-SVM to gain better kernel combination, which will be our future work.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] T. C. Phung, Y. S. Ihn, J. C. Koo, and H. R. Choi, "Edge identification of a small object through a low-resolution tactile sensor array," *International Journal of Precision Engineering and Manufacturing*, vol. 11, no. 2, pp. 247–254, 2010.

[2] N. Chen, R. Rink, and H. Zhang, "Efficient edge detection from tactile data," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, pp. 386–391, August 1995.

[3] G. M. Krishna and K. Rajanna, "Tactile sensor based on piezoelectric resonance," *IEEE Sensors Journal*, vol. 4, no. 5, pp. 691–697, 2004.

[4] K. Suwanratchatamanee, M. Matsumoto, and S. Hashimoto, "Robotic tactile sensor system and applications," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 3, pp. 1074–1087, 2010.

[5] Z. Pezzementi, C. Reyda, and G. D. Hager, "Object mapping, recognition, and localization from tactile geometry," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '11)*, pp. 5942–5948, Shanghai, China, May 2011.

[6] Z. Pezzementi, E. Plaku, C. Reyda, and G. D. Hager, "Tactile-object recognition from appearance information," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 473–487, 2011.

[7] U. M. Hernandez, T. J. Dodd, L. Natale, G. Metta, T. J. Prescott, and N. F. Lepora, "Active contour following to explore object shape with robot touch," in *Proceedings of the IEEE World Haptics Conference (WHC '13)*, pp. 341–346, Daejeon, Republic of Korea, April 2013.

[8] Z. Pezzementi, E. Jantho, L. Estrade, and G. D. Hager, "Characterization and simulation of tactile sensors," in *Proceedings of the IEEE Haptics Symposium*, pp. 199–205, Waltham, Mass, USA, March 2010.

[9] B. Schölkopf, A. Smola, and K. Müller, "Nonlinear component analysis as a Kernel Eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[10] K. I. Kim, K. Jung, and H. J. Kim, "Face recognition using kernel principal component analysis," *IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 40–42, 2002.

[11] Y. H. Liu, C. K. Wang, Y. Ting et al., "In-TFT-array-process micro defect inspection using nonlinear principal component analysis," *International Journal of Molecular Sciences*, vol. 10, no. 10, pp. 4498–4514, 2009.

[12] V. N. Vapnik, *Statistical Learning Theory*, Adaptive and Learning Systems for Signal Processing, Communications, and Control, Springer, New York, NY, USA, 1998.

[13] M. Gönen and E. Alpaydın, "Multiple kernel learning algorithms," *Journal of Machine Learning Research*, vol. 12, pp. 2211–2268, 2011.

[14] K. Weiß and H. Worn, "The working principle of resistive tactile sensor cells," in *Proceedings of the IEEE International Conference on Mechatronics and Automation (ICMA '05)*, pp. 471–476, Niagara Falls, Canada, August 2005.

[15] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, New York, NY, USA, 2nd edition, 2003.

[16] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[17] Y. H. Liu, *Feature analysis and classifier design and their applications in computer vision and data mining [Ph.D. thesis]*, National Taiwan University, Taipei, Taiwan, 2003.

[18] A. Rakotomamonjy, F. R. Bach, S. Canu, and Y. Grandvalet, "SimpleMKL," *Journal of Machine Learning Research*, vol. 9, pp. 2491–2521, 2008.

[19] Y. H. Liu and Y. T. Chen, "Face recognition using total margin-based adaptive fuzzy support vector machines," *IEEE Transactions on Neural Networks*, vol. 18, no. 1, pp. 178–192, 2007.

Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Submit your manuscripts at
http://www.hindawi.com

Hindawi

Journal of
Complex Analysis

Journal of
Mathematics

Mathematical Problems
in Engineering

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

International Journal of
Stochastic Analysis

Journal of
Optimization