Flat image recognition in the process of microdevice assembly

T. Baidyk a,*, E. Kussul a, O. Makeyev b, A. Caballero a, L. Ruiz a, G. Carrera a, G. Velasco a

a Center of Applied Science and Technological Development, UNAM (National Autonomous University of Mexico), Cd. Universitaria, A.P. 70-186, C.P. 04510 Mexico, D.F., Mexico
b National Taras Shevchenko University of Kyiv, Ukraine, 64, Volodymyrska Str., 01033 Kiev, Ukraine

Received 8 January 2003; received in revised form 29 July 2003

Abstract

An image recognition system for use in the assembly of microdevices is developed. The system gives an increase in the assembly process precision. A pin-to-hole insertion task was used to test developed system. The system will be used for assembly of microring-based filters.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Microassembly; Image recognition; Neural classifier; Neural interpolator

1. Introduction

The automation of microassembly is becoming more important in microtechnology with many investigators working in this area (Ofeifer and Dussler, 2002; Yamamoto et al., 2001; Kim and Cho, 1999a,b; Jonathan et al., 2001; Bleuler et al., 2000; Burel et al., 1995; Lee et al., 2001; Fatikow and Rembold, 1996). It is estimated that microassembly with its process steps “handling” and “joining” can make up to 80% of the total production costs (Ofeifer and Dussler, 2002). Automated solutions are necessary to decrease the costs on the one hand and to improve the quality of assembled microsystems on the other hand.

In our opinion there are at most two main approaches to microdevice production. The first is the technology of micro electromechanical systems (MEMS). This technology is developing very fast (Trimmer, 1997; Ohlickers and Jakobsen, 1997; Madni and Wan, 1998) and finds applications, for example, in the field of sensors and actuators, in the field of telecommunications (Eddy and Sparks, 1998; Kuwano, 1996). However to realize all the advantages of this technology it is important to have advanced packaging and assembly technologies (Yamamoto et al., 2001).

The second approach is connected with production of precise microequipment and microfactories.
(Kussul et al., 1996; Baidyk, 2001a,b; Baidyk and Kussul, 2002; Kussul et al., 2002a,b; Friedrich and Vasile, 1996; Ooyama et al., 2000). We call this approach “microequipment technology” (MET) (Kussul et al., 2002b). This approach is connected with a decrease of the size of microequipment. It allows the production of low cost microequipment for production of microdetails and Microsystems. Highly precise assembly processes are very important in this approach too.

The assembly tasks can be very different, depending on the area and the methods of microproduct manufacturing. Traditionally, the assembly process has three stages: (1) approach stage, (2) the contact and search stage, (3) the insertion stage (Whitney, 1982). Our paper is focused on the second and the third stages of pin-to-hole insertion task. These stages usually begin from the search task. The pin can be a rigid or a flexible part (Kim and Cho, 1999a,b). In our investigation we work with a rigid pin. Our task is to determine the coordinates of relative positions of the pin and the hole. Thereafter it is necessary to insert the pin in the hole.

The assembly task can be automated using force sensory information or visual information. To obtain this information different measurement systems have been developed (Kim and Cho, 1999b; Jonathan et al., 2001; Burel et al., 1995; Fatikow and Rembold, 1996). For example, Jonathan et al. (2001) use multiple laser light strips which are projected onto the product and one black-and-white CCD camera which is utilized to record the image of the scene illuminated with the light strips. In this way they can analyze the 2-D image to obtain 3-D information. Others use a stereoscopic microscope, three CCD cameras, a micromanipulator and personal computer (Lee et al., 2001). Our system includes one camera for teleconferences, four lamps, microequipment and a personal computer.

We solve the problem of microfilter assembly (Kussul et al., 2002b). The microfilter contains a large number of microrings inserted in special microchannels. To insert the microring in the microchannel we put it onto a pin. Therefore the position search problem is equal to the conventional pin–hole task.

2. Preliminary considerations

The task of micromechanical device assembly requires the determination of the relative position of microdetails. In the case of the pin–hole task it is necessary to determine the displacements $\Delta X$, $\Delta Y$, $\Delta Z$ of the pin relatively to the hole according to three coordinates $X$, $Y$, $Z$ by using the images obtained with the aid of the TV camera. Since the images, as a rule, are noised this task in its nature is the task of regression. However, the direct application of mathematical methods of regression in this case is not effective since the initial parameter space is too large, and the data base is too small for the mathematical statistics methods application. In order to solve this task it is necessary to carry out the preliminary conversion of initial parameter space into the new space which would make it possible to use known statistical methods. Large experience of such conversions is accumulated in the area of the image recognition. These conversions are designed for the subsequent application of the different classifiers. The existing classifiers can be divided into two large categories: the global classifiers and the local classifiers.

The global classifier assumes the construction of the nonlinear functions determined on the whole converted parameter space. The inequalities are added to these functions. The inequalities divide the space into the regions used for the classification.

The local classifier, as a rule, uses the functions determined on the small part of the parameter space generating the partition of the parameter space on the basis of comparatively simple principles.

An example of the global classifiers is the multilayer perceptron with the back propagation training. The local classifiers include different methods of the nearest neighbor classification (Duda and Hart, 1973), Random Subspace Classifier (Kussul et al., 1994). To the same type of classifiers could be related Radial Bases Functions (Howlett and Jain, 2000), Support Vector Machine (Vapnik, 1999), Modified Cannerva Method (Clarke et al., 1991), etc.

The global classifiers are the closest to the task of regression. However, the experience of the ap-
Application of different classifiers in the task of the image recognition (for example, obtained on the data base MNIST) shows that the best results give the local classifiers (Bottou et al., 1994; LeCun et al., 1998; Belongie et al., 2001, 2002; Cheng-Kin et al., 2002). These results are presented in Table 1. The MNIST data base contains 60,000 handwritten digits in the training set and 10,000 handwritten digits in the test set. We have developed new neural classifier LIRA (LImited Receptive Area classifier). This classifier is a local classifier. It was tested on MNIST data base and showed very good results.

The classifier LIRA gave only 55 errors from 10,000 samples on MNIST data base. In Fig. 1 we present unrecognized symbols.

So, it is interesting to apply the classifier LIRA to solve the regression problem. For this purpose we made attempt to reduce the pin-to-hole task to the handwritten symbol recognition task. There are two possible ways of the local classifier application:

1. to divide the displacements $\Delta X$, $\Delta Y$, $\Delta Z$ into many sections. Each section is to be considered as the output class of the classifier,
2. to divide the displacements $\Delta X$, $\Delta Y$, $\Delta Z$ in the sections as in the item 1 and to use results of the work of the classifier as parametric space for constructing of the regression equations. In this case we create the regression equations as a function of one variable.

In this work we did experimental study of the both methods. The comparison of the obtained results showed that the second method works better.

### 3. The method of microfilter assembly

It is well-known that in the micro domain adhesive forces between parts and micro gripper (in our case between the microring and the pin) make it difficult to realize the microassembly task. It can be easier for the micromanipulator to grip and manipulate an object than to release it afterwards. Gravity only plays a minor role in the microworld, but attractive forces such as electrostatic forces or Van-der-Waals forces are significant. This was one of the reasons to work out a new method of microfilter assembly.

We have developed the technology of automatic assembly of microfilters (the description of the microfilters is presented in (Kussul et al., 2002b)). One of the problems of the microassembly process

<table>
<thead>
<tr>
<th>Methods</th>
<th>% of error number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear classifier</td>
<td>12.0</td>
</tr>
<tr>
<td>Linear classifier (nearest neighbor-NN)</td>
<td>8.4</td>
</tr>
<tr>
<td>Pairwise linear classifier</td>
<td>7.6</td>
</tr>
<tr>
<td>K-NN, Euclidean</td>
<td>5.0</td>
</tr>
<tr>
<td>2-Layer NN, 300 hidden units (HU) (28×28-300-10)</td>
<td>4.7</td>
</tr>
<tr>
<td>2-Layer NN, 1000 HU (28×28-1000-10)</td>
<td>4.5</td>
</tr>
<tr>
<td>2-Layer NN, 1000 HU, [distortions] (28×28-1000-10)</td>
<td>3.8</td>
</tr>
<tr>
<td>2-Layer NN, 300 HU, [distortions] (28×28-300-10)</td>
<td>3.6</td>
</tr>
<tr>
<td>1000 Radial basis function (RBF) + linear classifier</td>
<td>3.6</td>
</tr>
<tr>
<td>40 Principal component analysis (PCA) + quadratic classifier</td>
<td>3.3</td>
</tr>
<tr>
<td>3-Layer NN, 300 + 100 HU (28×28-300-100-10)</td>
<td>3.05</td>
</tr>
<tr>
<td>3-Layer NN, 500 + 150 HU (28×28-500-150-10)</td>
<td>2.95</td>
</tr>
<tr>
<td>3-Layer NN, 300 + 100 HU, [distortions] (28×28-300-100-10)</td>
<td>2.5</td>
</tr>
<tr>
<td>3-Layer NN, 500 + 150 HU, [distortions] (28×28-500-150-10)</td>
<td>2.45</td>
</tr>
<tr>
<td>K-NN, Euclidean, deslant</td>
<td>2.4</td>
</tr>
<tr>
<td>LeNet-1 [16×16]</td>
<td>1.7</td>
</tr>
<tr>
<td>2-Layer NN, 300 HU, [deslant] (20×20-300-10)</td>
<td>1.6</td>
</tr>
<tr>
<td>K-NN, tangent distance, [16×16]</td>
<td>1.1</td>
</tr>
<tr>
<td>Support vector machine (SVM) poly 4</td>
<td>1.1</td>
</tr>
<tr>
<td>LeNet-4</td>
<td>1.1</td>
</tr>
<tr>
<td>LeNet-4/K-NN</td>
<td>1.1</td>
</tr>
<tr>
<td>LeNet-4/Kernel</td>
<td>1.1</td>
</tr>
<tr>
<td>Reduced set SVM poly 5</td>
<td>1.0</td>
</tr>
<tr>
<td>LeNet-5</td>
<td>0.95</td>
</tr>
<tr>
<td>Virtual SVM poly 9 [distortions]</td>
<td>0.8</td>
</tr>
<tr>
<td>LeNet-5 [distortions]</td>
<td>0.8</td>
</tr>
<tr>
<td>SVC-rbf_binary</td>
<td>0.72</td>
</tr>
<tr>
<td>Boosted LeNet-4 [distortions]</td>
<td>0.7</td>
</tr>
<tr>
<td>Shape matching + 3-NN</td>
<td>0.63</td>
</tr>
<tr>
<td>Classifier LIRA</td>
<td>0.55</td>
</tr>
<tr>
<td>SVC-rbf_grayscale</td>
<td>0.42</td>
</tr>
</tbody>
</table>
is the following: the workpiece sticks to the micro manipulator gripper and it is difficult to release the gripper from the workpiece. To resolve this problem we propose the following assembly process (Fig. 2).

The gripper of the assembly device in our case is the needle (1) in the tube (2). The microring is put on the needle and is introduced with the needle into the hole (Fig. 2a and b). After that, the needle is removed and the microring is held in the hole with the tube (Fig. 2c). In the next step the tube with the needle is moved aside and the microring is held in the hole and can not follow the tube (Fig. 2d). After that the tube with the needle are moved up and the microring is held in the hole (Fig. 2e).

Above we mentioned that the important feature of our microequipment is the use of low cost components. Low cost components do not permit us to assemble the device with high precision. To compensate this drawback we have developed an adaptive algorithm of microassembly using a low cost technical vision system. In Fig. 3 the prototype of the visual controlled assembly system is shown.

The main idea of this approach is to replace the stereovision system, which requires two video cameras, with a system based on one TV camera for teleconferences, with a cost of 50–70 dollars, and four light sources. The shadows from the light sources permit us to obtain the 3-D position of the needle with the microring relative to the hole.
4. The method of 3-D pin–hole position presentation by 2-D images

Here we consider the assembly task: to install the pin into the hole (Fig. 4).

For this purpose it is necessary to know the displacements \((d_x, d_y, d_z)\) of the pin tip relative to the hole. It is possible to evaluate these displacements with a stereovision system, which resolves 3-D problems. The stereovision system demands two TV cameras. To simplify the control system we propose the transformation of 3-D into 2-D images preserving all the information about mutual location of the pin and the hole. This approach makes it possible to use only one TV camera.

Four light sources are used to obtain pin shadows. Mutual location of these shadows and the hole contains all the information about the displacements of the pin relative to the hole. The displacements in the horizontal plane \((d_x, d_y)\) could be obtained directly by displacements of the shadows center points relative to the hole center. Vertical displacement of the pin may be obtained from the distance between the shadows. To calculate the displacements it is necessary to have all the shadows in one image. We capture four images corresponding to each light source sequentially (Fig. 5), and then we combine four images. Therefore it is necessary to extract contours and superpose contour images (Fig. 6).

5. Neural classifier for pin–hole position detection

The image in Fig. 6 is similar to an optical characters. It can be interpreted and treated as a symbol (resemblance to letters or numbers). For optical character recognition we developed a Random Threshold Classifier and Random Subspace
Classifier (Kussul and Baidyk, 1994, 2002; Kussul et al., 1994, 2001), which present the modified Rosenblatt perceptron. We adapt the classifier to recognize the position of the pin (Baidyk, 2001a; Baidyk and Kussul, 2002). In this chapter we will describe the structure of our neural random subspace classifier.

The neural classifier includes three neuron layers (Fig. 7). The \(S\)-layer corresponds to the input image, and the \(A\)-layer is a layer of associative neurons. We connect each \(A\)-layer neuron with \(S\)-layer neurons randomly selected not from the whole \(S\)-layer but from the rectangle \((h/C3w)\), which is located in the \(S\)-layer (Fig. 7). The distances \(d_w\) and \(d_h\) are random numbers selected from the ranges: \(d_w\) from \([0, WS/C0w/138]\) and \(d_h\) from \([0, HS/C0h/138]\), where \(WS, HS\) stand for width and height of the \(S\)-layer. Very important parameters of this classifier are the ratios \(w/WS\) and \(h/HS\) which were chosen experimentally. The connections of the \(A\)-layer with the \(S\)-layer do not change during the training.

The excitation of the \(A\)-layer neurons takes place only by the following condition. Every neuron of the \(A\)-layer has \(m\) positive connections with neurons of the \(S\)-layer and \(l\) negative connections with the other \(S\)-layer neurons. Positive connection is activated when the image pixel that corresponds to this neuron has value 1. Negative connection is activated when the image pixel that corresponds to this neuron has value 0. The excitation of the \(A\)-layer neuron takes place when all \(m\) positive connections and \(l\) negative connections are activated.

The \(R\)-layer presents the outputs of the classifier. Every neuron of this layer is connected to all neurons of the \(A\)-layer with trainable connections. The excitation of \(R\)-layer neurons is defined in accordance with the formula:

\[
E_j = \sum_{i=1}^{n} a_iw_{ij},
\]

where \(E_j\)—the excitation of \(j\)-neuron of the \(R\)-layer, \(a_i\)—excitations of \(i\)-neuron of the \(A\)-layer, \(w_{ij}\)—connection weights between \(A\)-layer neuron \(i\) and \(R\)-layer neuron \(j\).

In the neural classifier the neuron from the \(R\)-layer having the highest excitation determines the class under recognition. This rule is used always in the process of recognition. In the training process this rule should be changed. Let the neuron-winner have excitation \(E_w\), its nearest competitor has excitation \(E_c\). If

\[
(E_w - E_c)/E_w < T_E
\]

the competitor is considered as a winner. Here \(T_E\) is the superfluous excitation of the neuron-winner.

As distinct from the Rosenblatt perceptron our neural classifier has only positive connections between the \(A\)-layer and the \(R\)-layer. In this case the training rule is the following:

1. Let \(j\) correspond to the correct class under recognition. During the recognition process we ob-
tain excitations of $R$-layer neurons. The excitation of neuron $R_j$ corresponding to the correct class is decreased by the factor $(1 - T_E)$. After this the neuron having maximum excitation $R_k$ is selected as winner.

2. If $j = k$, nothing to be done.

3. If $j$ does not equal $k$,

$$W_{ij}(t + 1) = W_{ij}(t) + a_i,$$

where $W_{ij}(t)$ is the weight of the connection between the $i$-neuron of the $A$-layer and the $j$-neuron of the $R$-layer before reinforcement, $W_{ij}(t + 1)$ is the weight after reinforcement, $a_i$ is the output signal (0 or 1) of the $i$-neuron of $A$-layer.

$$W_{ik}(t + 1) = W_{ik}(t) - a_i, \quad \text{if } (W_{ik}(t) > 0),$$

$$W_{ik}(t + 1) = W_{ik}(t), \quad \text{if } (W_{ik}(t) = 0),$$

where $W_{ik}(t)$ is the weight of the connection between the $i$-neuron of the $A$-layer and the $k$-neuron of the $R$-layer before reinforcement, $W_{ik}(t + 1)$ is the weight after reinforcement. According to this rule connection weights have only positive values.

We adapt the classifier LIRA for pattern recognition in assembly of micro devices. The experiments were made with the pin, the ring and the hole having the following diameters:

- diameter of the pin 1 mm;
- outer diameter of the ring 1.2 mm;
- inner diameter of the hole 1.25 mm.

6. Neural interpolator for pin–hole position detection

The neural interpolator differs from the neural classifier as follows: the excitations of the output neurons are considered as a set of values of continuous functions $f(dx)$ and $\varphi(dy)$. To determine the functions $f(dx)$ and $\varphi(dy)$ we use a parabolic regression equation, which is built on the base of five values selected from output neuron excitations according to the following rule. As a central point we selected the point with maximal value of the excitation $E_{\text{max}}$ (Fig. 8a). In addition to this point two points to the left and two points to the right of $E_{\text{max}}$ are selected. If point $E_{\text{max}}$ is located at the edge of the point sequence, additional points are obtained as a mirror reflection (Fig. 8b) of the points, which are situated on the other side of $E_{\text{max}}$.

After determination of the functions $f(dx)$ and $\varphi(dy)$ the parameters $dx_0$ and $dy_0$, under which the functions $f(dx)$ and $\varphi(dy)$ have maximal values are determined. The parameters $dx_0$ and $dy_0$ are considered as recognized pin–hole displacements.

The interpolator training algorithm also differs from the classifier training algorithm. In this case the training rule is the following:

The modification of the weights is carried out at every step. They are modified according to the equation:

$$W_{ij}(t + 1) = W_{ij}(t) + a_i \ast (\Delta w_j + \delta w_j),$$

where $W_{ij}(t)$ is the weight of the connection between the $i$-neuron of the $A$-layer and the $j$-neuron of the $R$-layer before reinforcement, $W_{ij}(t + 1)$ is

Fig. 8. Parabolic approximation.
the weight after reinforcement, $a_i$ is the output signal (0 or 1) of the $i$-neuron of the $A$-layer.

$$
\Delta w_j = \frac{1}{(dx_j - dx_c)^2 + \varepsilon} \quad \text{(4)}
$$

$$
\delta w_j = -\frac{1}{(dx_j - dx_0)^2 + \varepsilon},
$$

where $dx_c$—correct pin–hole displacement; $dx_0$—recognized pin–hole displacement, $\Delta w_j$ and $\delta w_j$—positive and negative training factors for the $j$th neuron; $dx_j$—pin–hole displacement which corresponds to the $j$th neuron (Fig. 9). For coordinate $Y$ similar formulas are used.

### 7. Results

To examine the developed method of pin–hole position detection we made a prototype of the microassembly device (Fig. 2).

In this prototype we have measured only two coordinate displacements $(X,Y)$ (Baidyk, 2001b; Kussul et al., 2002a). The experiments with the third coordinate $Z$ will be made in the future. The images were taken from the 441 positions in the plane $X-Y$ around the hole. The positions were located as a $21 \times 21$ matrix and the distance between neighboring positions was 0.1 mm. This distance corresponds to approximately 1.8 pixels in the $X$-coordinate and to approximately 1 pixel in the $Y$-coordinate. An example of the original image, created with four different light sources, is presented in Fig. 10 and after the preprocessing (contour extraction and combination) is presented in Fig. 11.

The investigation of neural classifier has shown the following results. The classifier outputs in the case of neural classifier method contain: 21 classes for $X$-coordinates and 21 classes for $Y$-coordinates (i.e. one $R$-layer has 21 neurons for $X$ and the other $R$-layer has 21 neurons for $Y$, i.e. we have 42 output classes). Every neuron from each $R$-layer has 512,000 connections with the previous layer (i.e. the $A$-layer has 512,000 neurons). In the case of the neural interpolator method we have only two outputs. It is necessary to explain the reasons of the large number of neurons in the $A$-layer.

In the global classifiers, as a rule, it is necessary to decrease the dimension of the converted space since the large dimension of this space needs the large data base for the training. Many of the local classifiers possess the opposite property. The results of their work are improved considerably with increasing of the converted space dimension. Theoretically it is desirable to increase the dimension of this space to obtain the linear separability of the converted space. In the practice it is necessary to take into account also the computation cost. We studied our classifier for the subject of increasing of the converted space dimension both on the MNIST data base and on our data base of the assembly images.

Experimentally was determined the dimension of the converted space which gives the best results in our tasks. The obtained dimension which corresponds to 512,000 parameters does not lead to excessive computation cost, since we use binary parameter space with the sparse coding. The sparse coding means that the quantity of “1” elements in the converted space is small in comparison with the dimension of the space itself. Thus, for instance, for the dimension of the space 512,000, an average quantity of “1” elements in the converted space is equal to 5000–6000. Since all calculations are made only for the “1” elements and “0” elements are ignored the time of calculations remains in reasonable limits.
It is known (Bottou et al., 1994), that the recognition rate may be increased if during the training cycle the images are represented not only in the initial state but also with shifted image positions (so called distortions). We investigated the classifier with 9 distortions.

From 441 images the training and the test sets were formed. The “chessboard” rule was used to divide the initial database into training and test datasets. The initial database was considered as a $21 \times 21$ chessboard. The “black” cells were considered as training samples and the “white” ones as test samples. In Table 2 the results of the classifier investigation are presented.

The number of correct recognitions is presented in the first columns of coordinates $X$ and $Y$ (column numbers 1, 3). Recognition rate (%) is in the second column (column numbers 2, 4). If we accept a pin–hole position error of up to 0.1 mm we obtain a rather high quality of recognition (the right part of the Table 2, columns with numbers 5–8).

The investigation of neural interpolator showed the following results: for coordinate $X$ the results are in Table 3.

For coordinate $Y$ the results are in Table 4.

In Tables 2–4 average results for six independent experiments are presented.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Correct recognition values of the classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>$Y$</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Average number</td>
<td>%</td>
</tr>
<tr>
<td>201.7</td>
<td>91.25</td>
</tr>
<tr>
<td>Average number</td>
<td>%</td>
</tr>
<tr>
<td>220</td>
<td>100</td>
</tr>
</tbody>
</table>
8. Discussion

The equipment for automatic assembly of microdevices must have high precision because the tolerances of microdevice components are very small. To increase the precision of the microequipment we use feedback based on computer vision principles. We propose an image recognition method based on neural network. Two types of neural networks were developed to solve this problem. We call the first type a neural classifier, the second type a neural interpolator. The neural classifier and neural interpolator are used to obtain relative positions of micropin and microhole. The neural classifier gives a finite number of relative positions. The neural interpolator serves as an interpolation system and gives an infinite number of relative positions. The image database of relative positions of microhole and micropin of diameter 1.2 mm was used for the neural classifier and the neural interpolator testing. This database contains 441 images. In this paper the recognition results of the relative positions are presented for the neural classifier and the neural interpolator.

A special prototype was made for the examination of the proposed method. With this prototype 441 images were obtained. The distance between neighboring images corresponds to 0.1 mm displacement of the pin relative to the hole. At the image this displacement corresponds to 1.8 pixels in the $X$-direction and 1 pixel in the $Y$-direction.

The experiments show that the computer vision system can recognize relative pin-hole position with a 0.1 mm tolerance. The neural classifier for this tolerance gives the correct recognition in 100% of cases in $X$ and 86.7% cases in $Y$ (Table 2). The neural interpolator gives for axis $X$ 100% (Table 3) and for axis $Y$—99.86% (Table 4). The neural interpolator also permits us to obtain data for smaller tolerances. For example, for axis $X$ tolerance 0.05 mm it gives 88.6% and for axis $Y$ tolerance 0.05 mm it gives 79.1%. The experiments show that the neural interpolator gives better results in estimation of the pin–hole relative positions.

It is interesting to mention that the 0.05 mm tolerance for $Y$-axis is less than one pixel in the image. In this case, the recognition rate 79.1% shows that the recognition possibilities are not limited by the resolution of one pixel. This result could be explained by the fact that each object in the image contains many pixels and these pixels give much more detailed information than one pixel.

9. Conclusion

A computer vision system for improvement of the microassembly device precision is proposed. Two algorithms of image recognition (neural classifier and neural interpolator) were tested in the task of pin–hole relative position detection.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>The results for $X$-coordinate of interpolator investigation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$ for recognition</td>
<td>$X \pm 0.025$ mm</td>
</tr>
<tr>
<td>$X$</td>
<td>$X%$</td>
</tr>
<tr>
<td>Error</td>
<td>74</td>
</tr>
<tr>
<td>Correct recognition</td>
<td>146</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>The results for $Y$-coordinate of interpolator investigation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$ for recognition</td>
<td>$Y \pm 0.025$ mm</td>
</tr>
<tr>
<td>$Y$</td>
<td>$Y%$</td>
</tr>
<tr>
<td>Error</td>
<td>101</td>
</tr>
<tr>
<td>Correct recognition</td>
<td>119</td>
</tr>
</tbody>
</table>
The neural classifier permits us to recognize the displacement of the pin relative to the hole with 1 pixel tolerance. The neural interpolator permits the recognition of the pin displacement relative to the hole with 0.5 pixel tolerance. The absolute values of detectable displacements depend on the optical channel resolution. In our case, one pixel corresponds to 0.05 mm X-axis displacements and 0.1 mm Y-axis displacements. This precision is sufficient for many cases of assembly processes.

Acknowledgements

This work was supported by CCADET, UNAM project MEC0104, DGAPA UNAM project IN-118799, CONACYT 33944-U, and INTAS-96-2248 of the European Community. We thank Neil Bruce for aid in preparation of the article.

References


