Neural Network Diagnosis for Visual Inspection in Printed Circuit Boards

A. Fanni, M. Lera, E. Marongiu, A. Montisci

DIEE Dipartimento di Ingegneria Elettrica ed Elettronica, Universita' di Cagliari
Piazza d'armi
09123 Cagliari, ITALY
fanni@diee.unica.it

Abstract

In this paper we present an Automatic Optical Inspection system to diagnose Printed Circuit Boards mounted in Surface Mounting Technology. The diagnosis task is handled as a classification problem with a neural network approach. The Printed Circuit Board tested images are preprocessed by means of several methods to reduce the amount of data to feed to the neural networks. We compare the results obtained in the diagnosis for all methods. The Automatic Optical Inspection system seems to be a good solution in an industrial application because of the low cost, very fast diagnosis and easiness to set-up and handle.

I. Introduction

The problem of operator diagnostic system usability is not much discussed in the scientific research but it is very important because the final user of such kind of systems usually does not know anything about diagnosis theory. To apply a theoretical diagnostic method to a real system it need to render the decision-making process clear to the operator who has to know very well the system under test but not the diagnostic theory process. Generally this implies to make an automatic diagnostic system process. The problem depends on both the implemented method and the system under test. From this point of view in this work we present a method to develop an Automatic Optical Inspection (AOI) system for Printed Circuit Board (PCB) using Surface Mount Technology (SMT). Automatic Optical Inspection plays a very important role in the automatic production process of Printed Circuit Boards. The advances in computers, image processing, pattern recognition and Artificial Intelligence have resulted in better and cheaper equipment for industrial visual inspection in the electronics industry, especially in the Surface Mount Technology.

Traditionally, PCB inspection has been performed manually or via electrical testing but the complexity of the circuits is continuously growing and the current trend is toward miniaturization of the components, so these previous methods seem not to assure high quality.

In automatic industrial inspection we can distinguish two important classes: electrical/contact methods and non-electrical/non contact methods. The contact methods have many limitations: high cost, low speed inspection and sometimes they do not detect potential defect such as linewidth or spacing reductions; on the other hand non-contact methods are improving the diagnostic capabilities in terms of speed and tasks. In the last 20 years a variety of algorithms for the automatic visual/optical inspection have been reported. Moganti et al. (Moganti et al. 1996) proposed a survey of the algorithms based on the nature of the information to be treated: referential approaches, non-referential approaches and the hybrid approaches. The referential comparison approaches execute a comparison between the image of the board under test and a reference image stored in a database. The non-referential approach is more complicated and involves the recognition of circuit features to compare with the reference features. The hybrid approaches are the combination of these methods.

In this paper we will present a PCB-AOI system based on neural networks. The algorithm we implemented could be classified as referential approach in the Moganti’s scheme even though it is not possible to put it in one of the subclasses proposed. The neural approach has been proposed by several authors (Jagannathan, Balakrishnan and Poppletwell 1991) differing for ancillary methods used such as fuzzy rule-based classification (Ko et al. 1998), or rule-based Expert System (Bartlett et al. 1988). Sometimes these AOI systems are time consuming or they need complicate illumination sources as many lasers, or sophisticated lighting design (Ko et al. 1998), (Loh and Lu 1999), or many CCD cameras (Gallegos et al. 1996), and sometime it could be very complicated to acquire good images. In contrast our system is particularly simple and cheaper, and presents many advantages. In fact, it needs just one CCD camera, moreover it doesn't need to place the PCB onto a precision X-Y table but it is enough to move slightly the CCD camera with a small precision system. This allows locating the PCB board in an automatic conveyor line. The illumination system must be uniform and diffuse. The main advantage of our system is its flexibility and transparency. In fact it can be easily
modify for a new kind of board to inspect, with low cost
for the training of the human operator in terms of time
and money. It does not need that the operator knows
neural network or the theory of Fourier transform or
learning strategy.
In this paper we will present results on defective solder
joints on an SMT-PCB but the AOI could detect many
more potential defect.
We concentrated our attention to analyze just the
interesting part of the board, for, i.e., the solder joints
area. This modular approach allows us to save
computational costs and computer memory. The system
acquires the image of the PCB under test, and operates a
pre-processing to reduce the amount of data maintaining
the right information for the diagnosis. The feature
extraction improves the speed of the inspection. We
develop two different kind of pre-processing techniques:
a) Fast Fourier Transformation FFT and b) Haar
Transformation of the images.
This work is oriented to study the application of neural
network in the quality control in the electronics
manufacturing industries, but it is very important to
underline that it is possible to use this procedure in a very
different kind of industries. It is worth to note that there is
a big effort to find practical applications of these
classification methods.
This paper is organized as follows. In section II the
diagnostic approach is developed as a pattern
classification task using neural networks. In section III we
describe the database and in section IV we discuss the
image features extraction. In section V a complete
description of the neural network algorithm used for the
diagnosis is presented. In section VI the experimental set-
up is presented. In section VII the results are reported.
Conclusion will end the paper.

II. Diagnosis Approach

The diagnosis process works as a pattern recognition
system where the patterns are the preprocessed images of
board areas. Classically, a pattern recognition system is
composed of three modules:
• a transducer which acquires data on a physical device;
• a feature extractor which reduces the data
dimensionality by computing certain features or
properties.
• a classifier which makes the final decision on the state
of device.
In this paper is presented a AOI system which is a
particular pattern recognition system, where the physical
device is a Printed Circuit Board (PCB), the transducer is
an image acquisition device and the classifier is a neural
network.
The purpose of the diagnostic system is to automatically
detect a set of defects, which can be recognized by means
of a visual inspection.
The architecture of the proposed diagnostic system is
shown in Fig. 1. The system consists of two procedures,
retrieved by an image databank. The databank is obtained from defect images detected on generic circuits, which have the same mounting technology and use the same components of the tested circuit.

The databank stores the images of the interesting part of the circuits such as solder joints, components or assembly defects as insufficient soldering and component lack. This image set is useful for every situation in the sense that we can build it independently of the specific circuit under test. If the layout of a new board is known, we can obtain an image of the circuit without defect (golden board) as a combination of the databank images. In the same way it is possible obtain an image of the faulty board just combining the image of the new board without defect with the images of others defects stored in the database. This procedure is totally automatic, the task of the operator is just to fix the defect set which the system have to diagnose, so this job need just some information about the process but not about the diagnostic system.

For this reason the data bank must contain the images of all kind of useful components and all kind of relevant defects originating from the assembly. The diagnostic system builder can create the image set, leaving the databank open to the updating.

IV. Image Processing Techniques

The Image Processing Module takes into account all the operations in order to obtain the images in a suitable form for neural network feeding (Pratt 1978). The need of pre-processing the image is due to the following assumptions:

- A diagnosis requires a set of measurements; therefore it is subject to tolerance issues and repeatability problems. The measured quantities have to be robust with respect to noise and significant in order to distinguish different defects.
- As previously stated, the neural networks are robust and accurate for decision problem solving whereas a well-suited pattern configuration set is presented. Since it is not possible to investigate for defects in the whole board as it is acquired, it is necessary to select the regions of interest and to process each region independently. Moreover, the region selection is important in order to reduce the displacement error due to uncertainty of board positioning in the apparatus.

For Surface Mounted Device (SMD), the selection of the region of interest is based on a template matching technique. This technique consists in searching a reference image (the template) in the whole image, in our case the image of the circuit board. The first step of this search is to determine in the whole image the region of interest where certainly there is the template zone. We choose a region who is a compromise between the dimension of the template and the uncertainty we have during the acquisition process and the mounting process. In any case this region shouldn't be too big to improve computational costs. Then the template is compared with any windows we can extract from the image with the template size. The comparison is simply the calculations of the Euclidean distances between the template and the windows selected. When we found the windows with the minimum distance from the template we have a reference point in the acquired image to determine the position of the regions of interest for the diagnosis.

![Fig. 2a) Region of interest 2b) template](image1)

![Fig. 2.c) Error Function](image2)
opaque), and slight variation of production line. Therefore it is fundamental to select the specific information which is more invariant with respect to these phenomena and that is useful to detect a kind of defect.

With particular reference to the solder joint inspection, three main types of feature extraction techniques have been used: statistics, edge detection and reversible linear transforms methods. Since images are in grayscale, they can be expressed as matrix of integers bounded in the [0–255] interval. The statistics considered are mean value and variance. The main advantage of these methods is the fast calculation and the invariance with respect to positioning, but they are highly dependent on lighting variations and contrast balancing.

The second group of feature extraction methods is based on edge detection techniques: the main advantage of these techniques is the invariance with respect to lighting variations; moreover, they can be easily transformed as positioning invariant. Fig. 3 reports the original images (a) and the images obtained applying a Sobel filter (b) for two examples, of a good and a poor solder joint.

As it can be seen an open joint does not present an edge in correspondence of the upper part of the image. Therefore it should be possible to discriminate the solder joints by counting the "edge" white pixels in the upper part of the binary images. It is important to note that this method is highly dependent on threshold tuning, which needs a-posteriori knowledge of the lighting and acquisition parameters. Moreover, albeit the images are small, the state of the art edge detection algorithms are time consuming and not applicable for online massive data processing.

The third group of methods for feature extraction takes into account the coefficients of a linear transform. A general 2D linear reversible transform can be defined as:

$$F(k_1, k_2) = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} f(n_1, n_2)A(n_1, n_2; k_1, k_2)$$

where $f(n_1, n_2)$ is the $(N_1 \times N_2)$ input matrix, $A(n_1, n_2; k_1, k_2)$ is the transform kernel.

With particular reference to the solder joint inspection problem, the Fast Fourier Transform (FFT), with a $\sin(.)$ kernel function, and the Haar Transform (HT), with a piecewise constant value kernel, were used.

![Fig. 3 Sobel Filtered images](image)

The Fig. 3 reports the original images (a) and the images obtained applying a Sobel filter (b) for two examples, of a good and a poor solder joint. As it can be seen an open joint does not present an edge in correspondence of the upper part of the image. Therefore it should be possible to discriminate the solder joints by counting the "edge" white pixels in the upper part of the binary images. It is important to note that this method is highly dependent on threshold tuning, which needs a-posteriori knowledge of the lighting and acquisition parameters. Moreover, albeit the images are small, the state of the art edge detection algorithms are time consuming and not applicable for online massive data processing.

We use the linear transform to obtain the helpful information for the diagnosis from the images, this fact reduces the matrix dimensions.

The Fig 4a) and 4b) show the images of the solder joints and the FFT and the HT. The element in the upper-left side in the transform plot is the zero frequency component, moving towards downright we have the increasing frequency components. Then in Fig. 4b) we can see the transform values we normalized and gray level converted. The most significant frequency components have the period similar to the image objects. In our case these objects are the solder joints reflected light. The good solder joint is very flat and has a good reflection of light. The poor solder has less reflected light. Then taking in to account few frequency components we can distinguish the good solder from the poor solder because the reflected light has a regular dimension. The linear transform allows us the reduction of the amount of data from 248x386 pixel, image dimension, to 28, that is the number of FFT components we selected and 31, that is the number of HT components. In the Fig.4 we can see the selected frequency component located by means of the white line. The two transform operators work in different way. It is possible to extract from the FFT module the shape information, on the other hand the phase diagram gives us the details position in the image. This is an advantage because if we consider only the module diagrams the Transformation is independent of the position uncertainty. These uncertainties cannot be completely removed. The FFT limit is their kernel functions which are define in the whole domain so it needs take in to account many frequency components to describe a discontinuity. The HT kernel functions are not equal to zero only in a limited field and this field becomes...
smaller as the frequency order increases. For this reason the HT could be very efficient to describe discontinuous functions but the HT are sensitive to the position uncertainty because they have only a diagram, not two, modules and phase, like FFT. The linear transforms considered yield a new matrix of coefficients of the same dimension of the original image. In fact, the FFT yields two matrixes, one for intensity and one for phase information, but only the amplitude matrix is useful for our scope. The phase matrix stores the relative position of the object, while the amplitude matrix is responsible for the general shape of the 2D object. Not all the coefficients are suitable for diagnosis purposes: since a solder joint is characterized as a low variation 2D shape, the high frequency coefficients are not necessary. Therefore we considered only the coefficients with low frequency information: as shown in table 1, we took the up-left triangle with exclusion of the DC coefficient, which is proportional to the mean value and does not produce any useful additional information. For the HT, with the same criterion we selected 16 coefficients, excluding the DC component.

V. Neural Model

As proposed in most of the literature (Ryu and Cho 1996), (Jagannathan, Balakrishnan and Poppletwell 1991) we use a Multi Layer Perceptron (MLP) with only one hidden layer of neurons, with sigmoid activation function and Backpropagation learning with generalized delta rule (Fanni et al. 1999), (Fanni et al. 1994). The number of input neurons is equal to the patterns dimension, which depends on the images dimension and on the pre-processing procedure adopted. The number of output nodes depends on the defect coding, which is stated by the system designer. The network training consists of modifying the connection weights so that the network associates the corresponding code to each input pattern belonging to the training-set. To ensure the training convergence it is fundamental to avoid that equal patterns are associated to different codes. This could happen when the searched defect is very similar to the corresponding image without defects, due to low CCD resolution, or when the feature extractor looses important information. We use *early-stopping* technique (Bishop 1995) to avoid overfitting. This consists in measuring, during the training, the error with respect to an independent set of patterns, called validation-set, and in stopping the training when this error reaches a minimum. In Caruana (Caruana 1996) is shown that, if early-stopping is used, the number of neurons in the hidden layer may vary without appreciably influencing the network performances, provided it is sufficiently, but not excessively, large. The results of our simulations, not reported in this paper, seem to confirm this general rule. The validation-set we use for the *early-stopping* is independent from the training set. We construct it by adding random vectors to the training patterns. The random vector norm depends on the uncertainty in the pattern constructing, which is previously measured in several tests. The same method is adopted to construct a set of different training patterns for the same defect in order to increase the neural networks robustness with respect to the noise.

VI. The experimental set-up

As described in the diagram system reported in Fig. 1, the PCB optical inspection system has the following parts:

1. X-Y precision mechanical positioning system for the image acquisition device;
2. Image acquisition device (CCD camera or Digital camera);
3. Illumination system
4. PC computer for integrating the movements of the scanning stage, the image acquisition, the image pre-processing, the implementation of the neural network and the post-processing phase.

The diagram in Fig.4 shows the inspection process modules we identified:

- **The Motion Control module**
  This module supervises all the displacements, which allows putting the image acquisition device in the correct position. It should be possible to choose the dimension of the region of interest to acquire, modifying in the right way the zooming of the objective. We chose to move the image acquisition device rather than the board in order to obtain a faster and more flexible system, independent from the production line. In this way the AOI become very useful not only in electronic industry but in a wide range of applications.

- **The Image acquisition module**
  This module supervises the image acquisition. We though it is better to use a CCD camera, although it is more expensive than a digital camera, because it is easier to manage during the positioning procedure due to the image display on line.

- **The Illumination System**
  This module is very simple because the illumination requirements are not so complicate: it needs a diffuse, uniform, constant light.

- **The Diagnosis Module**
  The diagnosis module supervises many tasks of the inspection process. It includes: the feedback positioning routine which allows to put near the correct position using reference points (fiducials); the preprocessing routine reduces the images size extracting the features useful for the diagnosis; the diagnosis routine which is a neural network and the post processing routine.

- **The Task Module Manager (TMM)**
  This module supervises the whole acquisition and inspection process. An integrated environment has to be developed to control at any time image acquisition, signal conditioning devices, instrument control interfaces,
motion control, diagnosis and image processing and the communication device.

- **The Database**

  In this module all the information and data useful for most part of the routines of the inspection process are stored. It contains the fiducials points for the correct positioning of the PCB, the reference images, and the illumination conditions. In the database there are also the models of the board components.

- **The Graphic User Interface (GUI).**

  A very simple GUI overviews all the process steps. It should be easy to handle and modify.

  The system works in this way: it is possible to put the AOI on line so the automatic PCB conveyor line can put the board under test directly in the test area. This position is close to the right one but it is necessary to move the image acquisition device to catch the good one to acquire the image. We can obtain it by means of the precision mechanical positioning system. The repeatability is about 50 µm over the two axes X and Y. This repeatability is good enough to have very suitable images of the interesting part compare to the reference ones. The zooming and the imaging system resolution should be at least 20 x 40 pixels for each solder joint in order to have images ready for the neural network diagnosis.

  There is a feed back control system to adjust the position using the fiducials point information retrieved from the database images. The first image we acquire is just to recognize the fiducials points and with small displacements of the camera the system corrects the wrong position, then it is possible to acquire the images. In the database there are also information about the particular kind of board, the area where there are the solder joints to inspect, the illumination conditions and the image of the "golden board", which is the reference for that kind of board.

### VII. Results

In this section the results are shown, with particular reference to the FFT and HT methods.

A test phase is organized in two different diagnosis scenarios: the first one is based on two diagnosis classes, "go" (a good solder joint) and "No-go" (an unacceptable open joint). The second scenario has one more class, an intermediate "warn" class (open joint with more than 50%...
of well-positioned soldering paste). The latter class is very interesting for re-factoring process. The knowledge of an acceptable but non-optimal soldering paste distribution can be used as a starting point for statistical defect analysis, which will improve the soldering process. For the first test case (go/No-go diagnosis), both the transform methods showed good performance, but the HT based diagnosis yielded slightly better results. It is important to calculate the estimated false alarm probability \( P_{fa} \), a “Go” solder joint classified as “No-go”, and the a posteriori missed alarm probability \( P_{ma} \), a “No-go” solder joint classified as “Go”. In this way it is possible to associate the cost (economic loss) for a missed alarm (which implies the board replacement in whatever point of production line) and for a false alarm (which implies a human inspection of the board). Since the cost of a missed alarm can be very high, the system has to be tuned in order to limit this event: both of the diagnostic systems yield a \( P_{ma} = 0 \), which is an optimum result, but the HT diagnosis system shows better performance yielding a lower \( P_{fa} \). Tables 2.a and 2.b show a detailed comparison between the two diagnosis approaches.

### Table 2.a

<table>
<thead>
<tr>
<th>Real state</th>
<th>Go</th>
<th>Warn</th>
<th>No-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go</td>
<td>92,5%</td>
<td>6,7%</td>
<td></td>
</tr>
<tr>
<td>No-go</td>
<td>0,0%</td>
<td>100,0%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2.b

<table>
<thead>
<tr>
<th>Real state</th>
<th>Go</th>
<th>Warn</th>
<th>No-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go</td>
<td>94,3%</td>
<td>5,7%</td>
<td></td>
</tr>
<tr>
<td>No-go</td>
<td>0,0%</td>
<td>100,0%</td>
<td></td>
</tr>
</tbody>
</table>

Tables 2 The Bold numbers represent classification results. \( P_{ma} \) in bottom-left cell; \( P_{fa} \) in top right cell.

Concerning the second test case (Go/Warn/No-go diagnosis), Tables 3.a and 3.b report poor performances for “warn” class diagnosis, because is not possible define clearly the boundary between the warn and the other two classes. This is not a real problem since it is used only for process quality control and this does not imply a highly expensive risk associated to misclassification of warn class. Even in this case, the HT diagnosis method is better than the FFT one. If we decide to reduce the three class decision problem to a two class problem by grouping “Go” and “Warn” classes, it is possible to compute \( P_{fa} \) and \( P_{ma} \) and it is evident (as reported on Tables 4.a and 4.b) that we obtain optimal results with HT diagnosis while the FFT is less accurate but however better than the Go/No-go test case.

### Table 3.a

<table>
<thead>
<tr>
<th>Real state</th>
<th>Go</th>
<th>Warn</th>
<th>No-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go</td>
<td>96,9%</td>
<td>4,1%</td>
<td>0,0%</td>
</tr>
<tr>
<td>Warn</td>
<td>58,3%</td>
<td>33,3%</td>
<td>8,3%</td>
</tr>
<tr>
<td>No-go</td>
<td>0,0%</td>
<td>0,0%</td>
<td>100,0%</td>
</tr>
</tbody>
</table>

### Table 3.b

<table>
<thead>
<tr>
<th>Real state</th>
<th>Go</th>
<th>Warn</th>
<th>No-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go</td>
<td>98,2%</td>
<td>1,8%</td>
<td>0,0%</td>
</tr>
<tr>
<td>Warn</td>
<td>54,5%</td>
<td>45,5%</td>
<td>0,0%</td>
</tr>
<tr>
<td>No-go</td>
<td>0,0%</td>
<td>0,0%</td>
<td>100,0%</td>
</tr>
</tbody>
</table>

Tables 3. Numbers outside the main diagonal are the percentages of misclassification; bold cells represent \( P_{ma} \). FFT is less accurate than HT in classifying “warn” class.

### Table 4.a

<table>
<thead>
<tr>
<th>Real state</th>
<th>Go+Warn</th>
<th>No-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go+warn</td>
<td>99,3%</td>
<td>0,7%</td>
</tr>
<tr>
<td>No-go</td>
<td>0,0%</td>
<td>100,0%</td>
</tr>
</tbody>
</table>

### Table 4.b

<table>
<thead>
<tr>
<th>Real state</th>
<th>Go+Warn</th>
<th>No-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go+Warn</td>
<td>100,0%</td>
<td>0,0%</td>
</tr>
<tr>
<td>No-go</td>
<td>0,0%</td>
<td>100,0%</td>
</tr>
</tbody>
</table>

Tables 4. Results for two classes case.

### VIII. Conclusion

We have shown an AOI system neural network based for the diagnosis of PCB. The neural network approach allows having diagnostic systems very easy to handle in the set-up and in the diagnosis phases. Th good results obtained with a low cost prototype underline the robustness of the system and the method effectiveness.

### Acknowledgements

This work was partially supported by RAS, Regione Autonoma della Sardegna, (Project "Programma di Ricerca Applicata e Innovazione Tecnologica della RAS). It is now continuing in collaboration with the industrial partner STEP S.r.l., Macchiareddu, Cagliari, Italy
References


