Research Journal of Applied Sciences 2 (12): 1255-1260, 2007
ISSN: 1815-932X

Ceramic Wall Tile Quality Classification Training Algorithms Using Statistical Approach

[1]D.O. Aborisade and [2]T.S. Ibiyemi
[1]Department of Electronics Engineering, Ladoke Akintola University of Technology, Ogbomoso, Oyo-State, Nigeria
[2]Department of Electrical Engineering, University of Ilorin, Kwara-State, Nigeria

Abstract: This study describes a ceramic wall tiles surface quality control classification training algorithms. The algorithm employed statistical approach based on Bayes decision functions and minimum distance techniques for classification. The measured feature vectors of the training tile samples are used by the algorithm to generate 2-D display of the classifier in feature space. Euclidean distance between each of the training tile samples and the test tile sample in the feature space is computed by the algorithm. The test sample is assigned the class to which it is closet. Many experiments were conducted using different number of defective test tile samples ranging from 50-150 samples. The classification of these defective samples for three sets gave an average of 1.45% error rate. This classification performance is better than human operator within the shorted time taken by the machine.

Key words: Ceramic wall, tile, algorithms, training, statistical approach

INTRODUCTION

The surface quality of ceramic wall tiles becomes more and more importance as its demand in the market increased rapidly. The quality of a wall tile is decided by the function of several features of its occurring surface defects such as cracks.

In most cases, three different classes of tiles are considered in ceramic tiles industry. Tiles are classified based on the complex combination of defect features. Figure 1 shows some example images of defective ceramic wall tiles. Figure 1(a-d), shows second class tile images. There exist a few but still acceptable cracks. Figure 1 (e-h) has a lots of unacceptable cracks on the surface and is been considered as a third class or waste.

Traditionally, the surface quality of ceramic wall tiles is inspected manually by human experts that keep an eye on a production line moving at a rate of

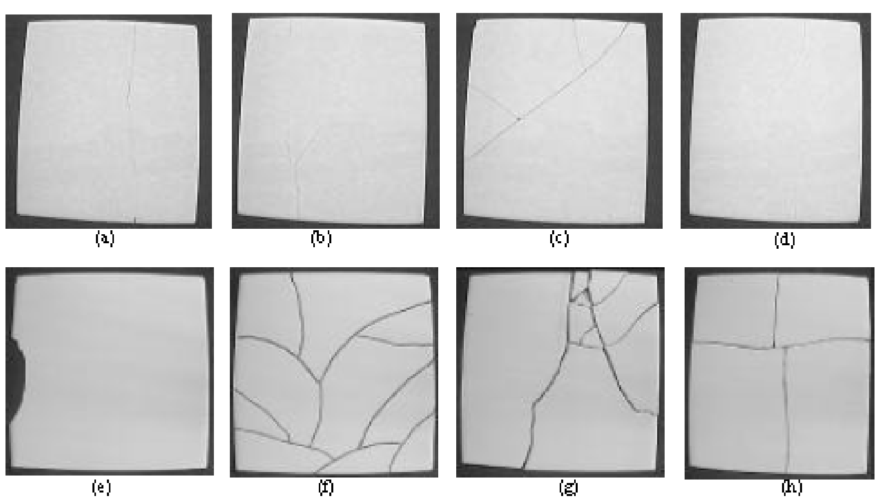

Fig. 1: Example of defective ceramic wall tile images

Corresponding Author: D.O. Aborisade, Department of Electronics Engineering, Ladoke Akintola University of Technology, Ogbomoso, Oyo-State, Nigeria

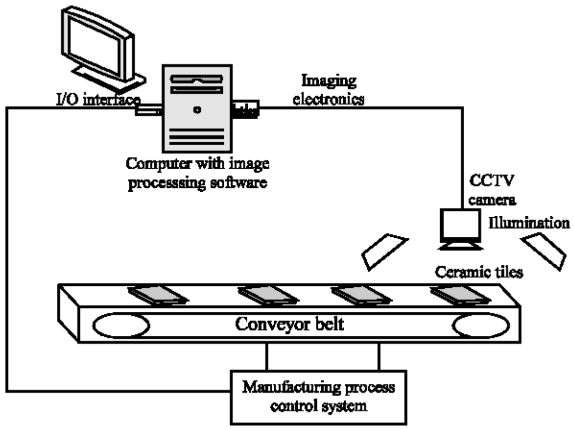

Fig. 2: Imaging system

40-100 items per min (Mital *et al*., 1998). This is a high speed to meet in real-time manual industrial inspection for a single person. It causes many problems, such as subjective results, incomplete inspection and deciphering of low-resolution defects (Sun *et al*., 2003). Recently, electronic tile classification systems have been introduced to solve these problems (Aborisade, 2005; Kopardekar *et al*., 1993; Hon-Son *et al*., 1984; Chin and Harlow, 1982). In such systems the tile surface image is first captured by charge-coupled device, CCD, camera as shown by the imaging setup in Fig. 2. After the image thresholding, some features are identified from each sample. The system makes the decision on the class unto which the tile belongs according to the measured features. The system performance is acceptable for first class tiles. However, for second class and third class tiles, the number of misclassified samples is large, approximately 45% in misclassification ratio. Due to this poor performance in classification, the tiles should then be reevaluated by human experts.

In this study, we present a ceramic wall tile surface classification training algorithms based on more advance statistical approach. The proposed training algorithm achieves significantly higher performance when it is compared to the traditional classification training algorithms. The proposed training algorithm is to be adapted in actual ceramic wall tile factories for everyday usage.

FEATURES FOR CLASSIFICATION

The classification method defined in this study is based only on visual data. Therefore, the classification features should be selected from visually perceptible parameters. The shape and location of the crack defects are the main features that human visual system detects. Therefore, we focused on detecting the crack defects by defining the shape of each crack with as few parameters as possible. The location of the crack is not a critical factor for classification, since the classification should be invariant of rotation and translation of the tile. So the location of the crack can be ignored in the definition as one of the shape measures.

For classification the crucial point is the distribution of the cracks on tiles surface. The crack distribution on a ceramic tile is a random process. It is impossible to have same crack distribution for two arbitrary ceramic tiles. The irregularity of a crack comes from the fact that, it may split into many branches and the thickness of the crack changes through these branches. A branch is defined as the region where we can assume a uniform thickness of a crack. Therefore, rather than defining the exact crack shapes (Haralick and Shapiro, 1992) and classifying the tiles based on these definitions, one should look for an approximate shape definition for classification. The approximation would not lead to misclassification since the main criteria is not the shape of a crack but the distribution of cracks.

The classification could be based on thickness of each branch, area of each branch, length, equivalenght diameter and extent of the branches in a crack. If the number of branches and a measure of the size of each branch can be defined, these will be sufficient criteria for comparing two cracks. The methods for calculating each of these feature vectors.

Area: The pixel area of the interior of the crack. Computed as the total number of pixels inside and including, the crack boundary.

Major axis length: The pixel distance length between the major axis end points. Computed as

$$\text{Major axis length} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \qquad (1)$$

Where, (x_1, y_1) and (x_2, y_2) are the major axis endpoints.

Minor axis width: The pixel distance length between the minor axis endpoints. Computed as

$$\text{Minor axis length} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \qquad (2)$$

Where, (x_1, y_1) and (x_2, y_2) are the minor axis endpoints.

Equivalent diameter: The diameter of a defect circle with the same area as the region. Computed as

$$\text{Equivalent diameter} = \sqrt{4 * \frac{\text{Area}}{\pi}} \qquad (3)$$

Extent: The proportion of the pixels in the bounding box that are also in the region. Computed as the area divided by area of the bounding box that surrounds the defect.
Where,

Bounding box area = Major axis length×Minor axis length

In an attempt to select the best feature vector for classification, feature variance for each category of crack is computed as

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2 \qquad (4)$$

Where, N is the number of training cracks samples that were measured, $\bar{x}$ is the mean of the features measurement and x_i is the i th actual measurement.

It is hard to determine the discriminatory features from the feature variance as the features that have large variance also have large means. Hence the best discriminating features is determined by computing the distance between the means of the two classes normalized by the variances;

$$V_{ij} = \frac{\left|\bar{x}_i - \bar{x}_j\right|}{\sqrt{\sigma_i^2 + \sigma_j^2}} \qquad (5)$$

The smaller this value is the nearer will be the class means and the worse the feature would be for deciding between the classes. Thus, the feature vector for a crack are arranged in form of a pattern vectors of the form

$$x = (x_1, x_2, \ldots, x_n)^T$$

CLASSIFICATION AND DECISION ALGORITHMS

For the classification of a set of ceramic wall tiles, in this section we described a decision algorithm used to determine a distance measure for the feature vectors defined above. The statistical formulation of the algorithms is implemented by Bayes decision functions (Julius and Rafael, 1974)

$$d_i(x) = p(C_i / X), \quad i = 1, 2, \ldots, M \qquad (6)$$

Where, M is the number of classes.

For each class we defined a random classification variable, r_i (X), with the following properties;

$$r_i(X) = \begin{bmatrix} 1 & \text{if } X \in C_i \\ 0 & \text{otherwise} \end{bmatrix} \qquad (7)$$

With the assumption that $E\{r_i(X)\} = E\{p(C_i/X)\}$, probability density function is written as a linear approximations

$$p(C_i/X) \approx w_i'X \qquad (8)$$

Where, $w = (w_1, w_2, \ldots, w_{n+1})'$ and $X = (x_1, x_2, \ldots, x_n, 1)'$ are called weight vectors and the augmented pattern, respectively.

In order to determine the weight vector belonging to the training patterns of classes of the tiles, we used increment-correction algorithm. The algorithm at kth iterative step is written as

$$w_i(k+1) = w_i(k) + \alpha_k X(k)\,\mathrm{sgn}\left\{r_i\left[X(k)\right] - w_i'(k)X(k)\right\} \qquad (9)$$

Where, $w_i(k)$ is the weight vector estimates at kth iterative step and α_k is the correction factors. However, using the definition of the sgn function which is defined by

$$\mathrm{sgn}\left(w_i^! X\right) = \begin{bmatrix} 1 & \text{if } w_i^! X > 0 \\ -1 & \text{if } w_i^! X \le 0 \end{bmatrix} \qquad (10)$$

weight vector belonging to pattern classes C_1 and C_2 is expressed in the equivalent form

$$w_i(k+1) = \begin{cases} w_i(k) + \alpha_k X(k) & \text{if } w_i'(k)X(k) < r_i\left[X(k)\right] \\ w_i(k) - \alpha_k X(k) & \text{if } w_i'(k)X(k) \ge r_i\left[X(k)\right] \end{cases} \qquad (11)$$

The weight vector at k = 100 for class C_1 and C_2 is computed as

$$w = \begin{pmatrix} 0.1127 \\ 0.1576 \\ 0.9462 \end{pmatrix} \text{ and } \begin{pmatrix} -0.0236 \\ 0.0116 \\ 0 \end{pmatrix} \qquad (12)$$

With the computed values of the weight vector, the equivalent decision boundary of the two-class characterized by a single prototype is determined as

$$d(x) = d_1(x) - d_2(x) = 0.1363x_1 + 0.1460x_2 + 0.0538 = 0 \qquad (13)$$

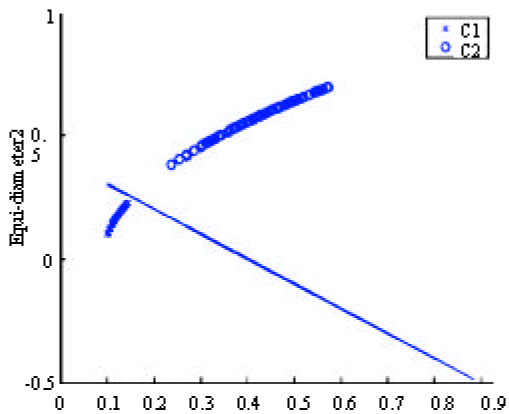

Fig. 3: 2-D feature space rep. of training data with decision surface

The entire classifier system based on the decision function is implemented in a computer with a high speed of computation. The system employed supervised, minimum distance techniques for classification. Details of the classification process follow:

Let $\{X^i, i= 1, 2, \ldots, N\}$ denote a set of N prototypes of class C_i samples of ceramic wall tile and let Z denote feature vector of an unknown tile sample to be recognized. Each sample is a concatenation of its component feature attributes. $X_r{}^i$ and Z_r refer to these components which is normalized as appropriate for the feature;

$$f \in F = [a(\text{area}), e(\text{equivalent diameter})]$$

With each prototype X^i, a ceramic tile identifier $I(X^i)$ is associated.

Comparison between Z and prototype vector X of each training class is performed using a distance measure. The weighted Euclidean distance between the test tile sample and each of the prototype samples is computed by

$$D \equiv D(Z, X^i) = \sum_{r \in F} \omega_r D(Z, X_r^i) \qquad (14)$$

Where, w_r is the feature-weighting coefficients.
Z is assigned the class to which it is closet.

$$Z = \begin{cases} C_i \text{ if } D(Z, X^i) < D(Z, X^j) \text{ for all } j \neq i \\ C_j \text{ if otherwise} \end{cases} \qquad (15)$$

In order to address the issue of accepting or rejecting pattern vector of an unknown tile which are not in the database, a more useful decision rule is proposed. The decision algorithm involves computing and storing threshold T (two dimensional decision surface given by Eq. 13) between the two training classes. The 2-D feature space representation of the training data with respect to the threshold is depicted in Fig. 3. The decision function is computed by:

$$\delta(Z) = \begin{cases} C_i, \text{ accept if } D(Z, X^i) \leq T \\ \text{otherwise, reject} \end{cases} \qquad (16)$$

THE DATABASE OF CERAMIC TILE SAMPLE IMAGES

A detailed discussion of the construction of the ceramic wall tiles database can be found in Aborisade (2005). Briefly, the database consists of 200 training pattern vector data for the two tile classes. The tiles images were collected in a laboratory setting so as to minimize the amount of preprocessing that is necessary in order to eliminate complicating effects, such as tilt, rotation, shifting, scaling and changes in illumination. For each class the training pattern vector is characterized by two measurements, giving a total of 200 training feature measures per class in 2-D feature space. A numbers of pattern vector were also collected and used to test the two class system.

Both the database and test pattern vector of the ceramic tile images were snapped at a dimension of 82×115 and stored as 8-bit gray scale. Some of the data test images are shown in Fig. 1.

EXPERIMENTAL RESULTS

The training pattern vector, X, consists of 200 training data sets for the two tile classes in the database. For training, X consists of two extracted feature measures, X_1 and X_2 in two-dimensional feature space. X_1 consists of 100 area feature measures and X_2 consists of 100 equivalent diameter measures, respectively per class. The system is trained to plot at each location in the feature space the coordinate (X_1, X_2). To segment the training set into two classes, decision surface described in Eq. 13 is plotted as shown in Fig. 3.

After training, the system was tested for correct classification performance. To see how the performance scales with increasing number of test samples with real and synthetic defects in the database, the performance of the system is looked at as the number of test tile samples with unknown classification varied from 50-150 samples.

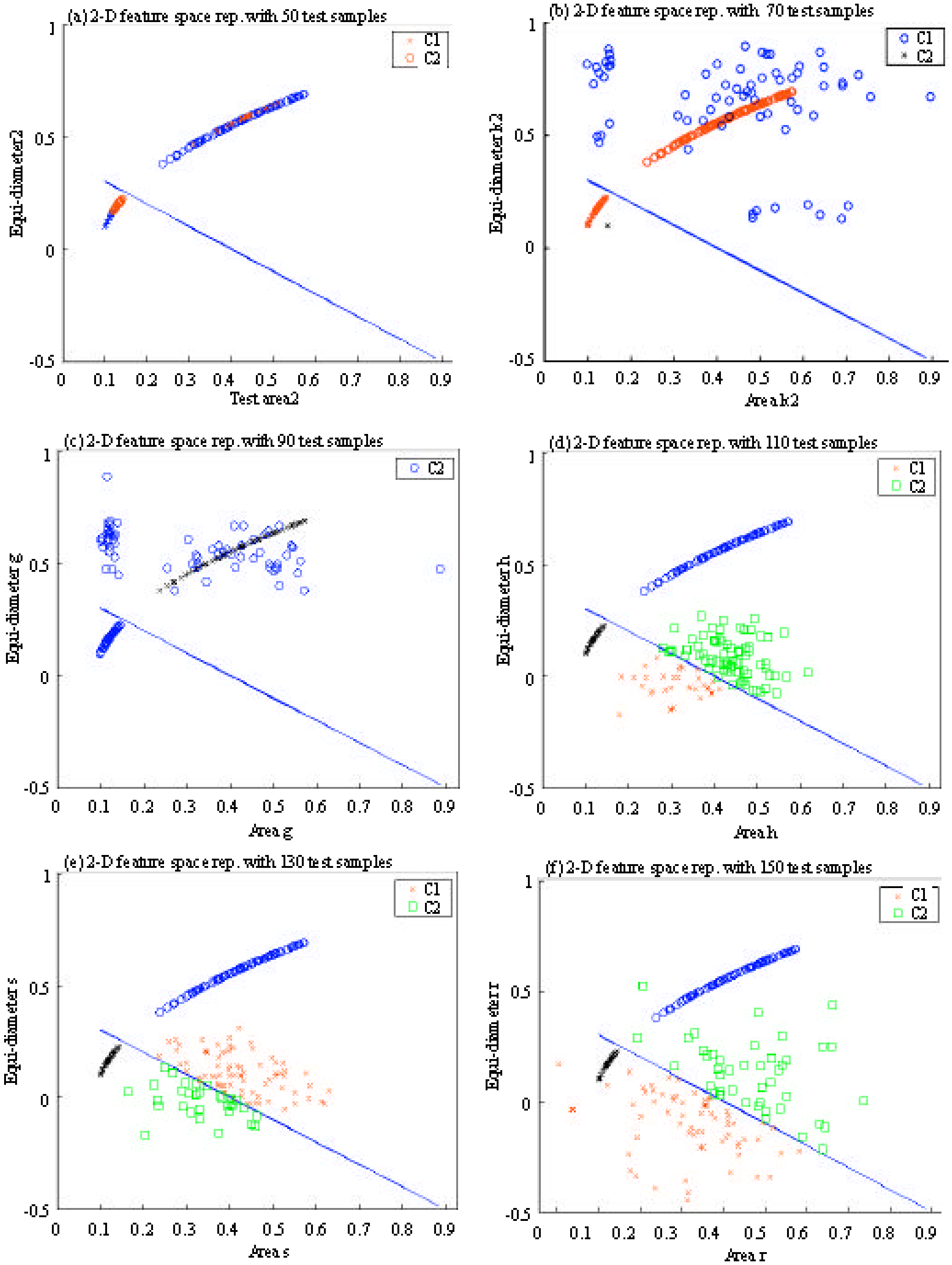

Fig. 4: 2-D feature space rep. of testing data with decision surface

Figure 4 shows the results obtained for the 2-D feature space representation of the testing data. Table 1 shows the correct classification experimental results for the test ceramic wall tiles samples. Note that each number in Table 1 is an average over 3 different trials of the experiment. In each trial, a different database was randomly chosen.

The results show that the algorithm scales well and for the arbitrary test tile samples gives an average classification error rate of 1.45%. Figure 5 shows how the

Fig. 5: Training time (in minutes) against the number of defect samples in the database

Table 1: Correct classification results for various number of ceramic wall tiles in the database

No. of test tile samples	Class, C_1	Class, C_2	Error	Train time (mm)
50	50.00	50.00	0	9.2
70	1.43	98.60	0	13.1
90	0.00	100.00	0	19.3
110	30.00	66.40	3.6	24.0
130	22.30	74.60	3.1	28.6
150	43.30	54.70	2.0	33.6

training time scales as a function of the number of training tile data sample in the database. Clearly, the training scales linearly with the number of tile sample in the database.

CONCLUSION

An efficient classification training algorithm for the quality control of a ceramic wall tile has been presented. We have applied statistical approach in the derivation of the proposed classification algorithms using 2-D feature space representation. Evaluation results on a limited number of test tile samples ranges from 50-150 samples seemed very promising and demonstrate the effectiveness of the algorithm in the ceramic wall tile classification hence, the algorithm eliminate subjectivity in the rejection/ acceptance decisions. The average error rates of 1.45% produced using the algorithm show a distinct advantage over human operator. This was primarily due to the inclusion of distance information in the 2-D feature space. Finally, the proposed algorithm was detailed. It assures that nearest neighbor error count of the entire data set is exactly preserved in the condensed data set.

REFERENCES

Aborisade, D.O., 2005. A Development of Computer Vision For Real-Time Industrial Inspection. Ph.D thesis, University of Ilorin.

Chin, R.T. and C.A. Harlow, 1982. Automated Visual Inspection: A Survey. IEEE. Trans. Pattern Anal. Machine Intell., 4: 557-573.

Haralick, R.M. and L.G. Shapiro, 1992. Computer and Robot Vision, Addison-Wesley.

Hon-Son Don, King-Sun Fu, C.R. Liu and Wei-Chung Lin, 1984. Metal Surface Inspection Using Image Processing Techniques. IEEE Transactions of System, Man and Cybernetics, Vol. SMC-14, No.1.

Julius, T.T. and C.G. Rafael, 1974. Pattern Recognition Principles, Addision-Wesley.

Kopardekar, P., A. Mital and S. Anand, 1993. Manual, Hybrid and Automated Inspection Literature and Current Research. Integrated Manufacturing Sys., 4: 18-29.

Mital, A., M. Govindaraju and B. Subramani, 1998. A Comparison between Manual and Hybrid Methods in parts Inspection. Integrated Manufacturing Sys., 9: 344-349.

Sun, H., K. Xu and J. Xu, 2003. Online Application of Automatic Surface Quality Inspection System to finishing line of Cold Rolled Strips. J. Uni. Sci. Technol., 10: 38-41.