Comparison between GLCM and Modified Zernike Moments for Material Surfaces Identification from Photo Images

Fathin Liyana Zainudin[1], Abd Kadir Mahamad[1], Sharifah Saon[1], Musli Nizam Yahya[2]

[1]Faculty of Electrical & Electronic Engineering,
Universiti Tun Hussein Onn Malaysia (UTHM),
84600, Parit Raja, Johor, Malaysia

[2]Faculty of Mechanical and Manufacturing Engineering,
Universiti Tun Hussein Onn Malaysia (UTHM),
84600, Parit Raja, Johor, Malaysia

***Abstract*— Types of materials are one of an important data for research in acoustic engineering. This paper compares methods for extracting texture data of material surfaces for classification. Gray Level Co-occurrence Matrix (GLCM) and modified Zernike moments that is applied for image extraction are tested and compared with back propagation neural network used for classification. These methods are also applied to the Brodatz texture database as a general comparison. The GLCM method shows a good performance and regression, R>0.9 for the Brodatz database while the collected surfaces datasets using GLCM and modified Zernike moments as well as the Brodatz datasets using modified Zernike moments method had only managed an acceptable performance and regression of R>0.8.**

Keywords—texture analysis; texture classification; GLCM; Zernike moments; back propagation Neural Network

I. INTRODUCTION

Absorption coefficient is one of an important feature in acoustic engineering. Different types of materials have different level of sound absorption. A few studies have been conducted in order to test different alternatives for identifying absorption coefficient of material surfaces [1][2]. In order to identify the surfaces absorption coefficient, the material surfaces need to be identified first. For material identification and classification, the surface texture is the key feature used for extraction.

Texture analysis is one of the most important techniques used in image processing and pattern recognition. In texture analysis, the first and most important task is to extract texture features which efficiently embody information about the textural characteristics of the original image [3].These features can then be used for the description or classification of different texture images. Various algorithms have been put forward for texture analysis, such as the Gray Level Co-occurrence Matrix (GLCM) [4], Gabor filtering, wavelet decomposition, modified Zernike moments [5] etc.

Initially, texture analysis was based on the first order or second order statistics of textures. In statistical approaches, textures are considered to be formed by certain random processes. The types of textures were analyzed by studying the statistical properties of the intensity values of pixels or the coefficients of certain filter banks [3]. Haralick et al. calculated second-order grayscale statistics using GLCM and defined the statistical moments as a texture descriptor [7].

To obtain the most accurate and compatible network for material surface identification, a comparison between two different image processing techniques for texture classification, the GLCM and modified Zernike moment has been done in this study.

II. TEXTURE ANALYSIS AND CASSIFICATION

A. Gray Level Co-occurrence Matrix (GLCM)

GLCM method is characterized by its capability of extracting second order statistical texture features when considering the spatial relationship of pixels and has been proved to be a promising method in many image analysis tasks [6]. GLCM method considers the spatial relationship between pixels of different gray levels. The method calculates a GLCM by calculating how often a pixel with certain intensity occurs in relation with another pixel at a certain distance and orientation [7].

Given two sets of sub-cellular localization images under differing experimental conditions, an efficient image feature can be used to evaluate if there is a statistically significant difference, even to the extent that visually indistinguishable images of distinct localizations may be differentiated [8]. The feature sets proposed in the literature include, for instance, morphological data of binary image structures, Zernike moments and edge information [9]. Use of a single technique for the extraction of diverse features in an image usually exhibits limited information description. Features extracted using different techniques can be combined in an attempt to enhance their description capability [9].

B. Modified Zernike Moments

Zernike polynomials were first proposed in 1934 by Zernike [10]. Their moment formulation appears to be one of the most popular, outperforming the alternatives (in terms of noise resilience, information redundancy and reconstruction capability). The use of Zernike moments have been frequently utilized for a number of image processing and computer vision task [11].

The modified Zernike moments in this study was proposed by Sim et. al [5]. For texture analysis, the modified Zernike moments need to go through three different stages: Discrete Fourier Transform (DFT), power spectrum normalization, and the computation of Zernike moments [3].

The texture image is first transformed by the DFT, where the invariant power spectrum is. The DFT of $F(k_1,k_2)$ for a signal $f(n_1,n_2)$with image size of N_1xN_2 is expressed by

$$F(k_1,k_2) = \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} f(n_1,n_2)e^{-j\frac{2\pi}{N_1}k_1n_1}e^{-j\frac{2\pi}{N_2}k_2n_2} \quad (1)$$

where $0 \le k_1 < N_1, 0 \le k_2 < N_2$.

The power spectrum is then normalized to obtain the scaled invariant. Lastly the invariant are extracted using the modified Zernike basis functions as seen in Equation (2)[5].

$$DMA_{nm} = \frac{n+1}{\pi}\sum_{k_1}\sum_{k_2}\log|F(k_1,k_2)|^2\frac{V^*{}_{nm}(\sqrt{\rho},\theta)}{2\rho}\rho d\rho \quad (2)$$

DMA_{nm} is the discrete modified Zernike moments for the (n,m) basis function that are computed from the normalized power spectrum of an input signal $|F_N(\rho,\theta)|^2$, and * is the complex conjugate.

Equation (3) shows the magnitude of the discrete modified Zernike moment, ZM_i.

$$ZM_i = |DMA_{nm}| \quad (3)$$

where $n \ge 0$, $n \ge m$, and n-$|m|$ is even, and i represents the number of magnitude.

C. Back Propagation Neural Network

Neural Network is a very effective computational tool. It applications are in almost every field of signal processing. Neural networks can be employed in several image processing applications like fingerprint identification, face recognition, cryptography etc. Furthermore, due to its attributes, such as massive parallelism, adaptability, and the inherent capability to handle nonlinear systems, this technique have been widely used in complex nonlinear function mapping, image processing [12], as well as pattern recognition and classification.

Neural network was originally developed to simulate the function of the human brain or neural system. Artificial neural network is basically a massive parallel computational model that imitates the human brain. This method does not really solve problems in a strictly mathematical sense, but it is one method of relaxation that gives an approximate solution to problems. A number of neural network techniques have been used in system identification such as feed-forward network, back-propagation network, Hopfield network and Kohonen network [13].

Multiple layer feed forward neural network or known as multiple layer perceptron is one of the most popular architecture used in neural network. Generally, this architecture consists of three layers; the input layer that receive the input signal, the hidden layer where the data is processed, and the output layer that generates the output signal as the result. Each layer contains neurons where each neuron connected to other neurons via weight connection.

Back-propagation neural network has become the most common techniques for the multi-layer networks. This network is a supervised learning method where the network learns from many inputs that points to the desired output. This method compared the outputs obtained from the network with the desired output. It repeatedly adjusts the weights of the connections in the network until the difference between the actual output vector of the network and the desired output vector are minimized [14]. The word "data" is plural, not singular.

III. EXPERIMENTS

A. Datasets

For this research, five material surfaces were taken from classrooms in Universiti Tun Hussein Onn Malaysia (UTHM). Fig. 1 shows sample of the types of surfaces taken, respectively are concrete walls, wooden walls, floors, doors, and ceilings.

These images were captured using Digital single-lens reflex (DSLR) camera with MICRO Nikkor 105mm 1:2.4 lens. The distances from the camera to the surfaces were fixed at 2 feet and 3 feet. Each image is set to a standard size of 4608x3072 pixels with a resolution of 300dpi. To ensure the accuracy of the system, 421 samples of image surfaces were taken from 2 feet distance and 370 samples of image surfaces were taken from 3 feet distance from five different classrooms in UTHM as shown in Table 1.

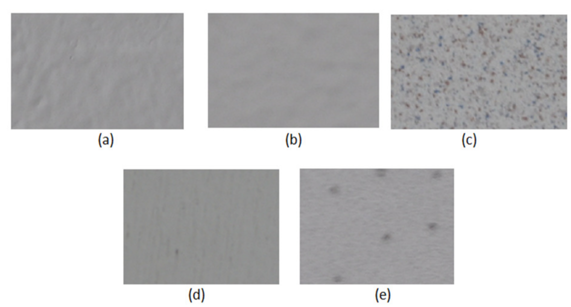

Figure 1. Sample images of (a) concrete wall, (b) wooden wall, (c) floor, (d) door, and (e) ceiling surface

TABLE I. NUMBER OF IMAGES FOR EACH SURFACES

Surfaces	**2ft**	**3ft**
Concrete wall	89	90
Wooden wall	81	93
Floor	80	61
Door	82	65
Ceiling	89	61
Total	421	370

Apart from these datasets, dataset from the Brodatz texture database [15] were also tested for comparison purpose. Five different textures from the Brodatz texture album were selected as shown on Fig. 2. For each Brodatz texture class, 40 images were extracted to produce a total of 200 images.

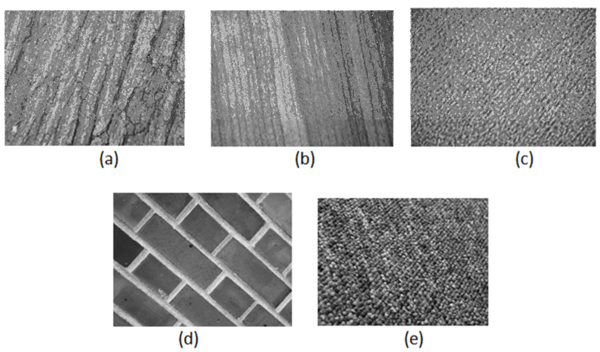

Figure 2. Sample textures of (a) bark, (b) wood, (c) glass, (d) brick, and (e) carpet from the Brodatz datasets.

B. Method Implemented

1) GLCM Implementation: The traditional gray co-occurrence matrices were computed, using a distance of 1 and angles of θ = 0°, 45°, 90° and, 135°. A set of 13 textural features such as contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, autocorrelation, maximum probability, sum average, sum variance, and sum entropy were calculated as described by Haralick [7] for each angle. Therefore, each feature extraction provides four values based on the setting and the average value is considered for this study.

2) Modified Zernike Moments Implementation: The amplitudes of the moments were experimentally calculated using the modified Zernike moments method [5]. For a better retrieval results for geometrically transformed textures, this study also includes the mean, P_0 and AC power, P_{AC} features.

$$P_0 = \frac{\sum_{n_1}\sum_{n_2} f(n_1,n_2)}{N_1 N_2} \tag{4}$$

$$P_{AC} = \frac{\sum_{n_1}\sum_{n_2}\left(f(n_1,n_2)-P_0\right)^2}{N_1 N_2} \tag{5}$$

3) Back-propagation Neural Network Implementation: The features extractions were then fed to the back-propagation Neural Network. The datasets were arranged randomly and then divided into 3 sets; 60% is for training, 20% is for validation, and the other 20% is for testing the network. Validation data is important to make sure overfitting did not occur. Testing data used only for testing the final solution in order to confirm the predictive power of the network. Training data must be more than validation and testing data to ensure that each type of texture is trained. The learning algorithm chosen was Levenberg–Marquardt because it is faster and more efficient. This algorithm is the combination of the excellent local convergence properties of the Gauss-Newton near a minimum with the consistent error decrease provided by gradient descent far from solution. This algorithm is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization that was designed to approach second-order training speed without having to compute the Hessian matrix. To obtain the optimum network, a trial and error scheme was conducted from 2 to 15 hidden nodes and the best network performance is selected.

IV. RESULTS

Table 2 shows the best performance for testing data of each dataset. Mean Squared Error (MSE) for each training, validation, and testing data are calculated and observed. From the result gained, the Brodatz dataset performed the best as the extracted features have much difference between them. It shows that GLCM perform better for images that have different and higher contrast between the neighbouring pixels. The MSE for the modified Zernike moments methods are similar between all datasets, shows that even with the high contrast of dissimilarity for each class in the dataset did not play a role in increasing the performance of the network.

Table 3 and Fig. 3 shows the regression plot that is the relationship between the outputs of the network and the targets. The test result only shows that Brodatz dataset using GLCM method managed to reach R >0.9 value while the rest only reach R>0.8. It shows that the collected surfaces datasets still need other proper processes to get a good fit.

TABLE II. BEST TESTING MSE FOR EACH EXPERIMENTS DATASETS

Datasets	**GLCM**	**modified Zernike moments**
Surfaces 2 feet distance	0.0247	0.0285
Surfaces 3 feet distance	0.0257	0.0221
Brodatz	0.0021	0.0257

TABLE III. REGRESSION FOR EACH EXPERIMENT DATASETS

Datasets	**GLCM**	**modified Zernike moments**
Surfaces 2 feet distance	0.8258	0.8572
Surfaces 3 feet distance	0.8457	0.8187
Brodatz	0.9883	0.8437

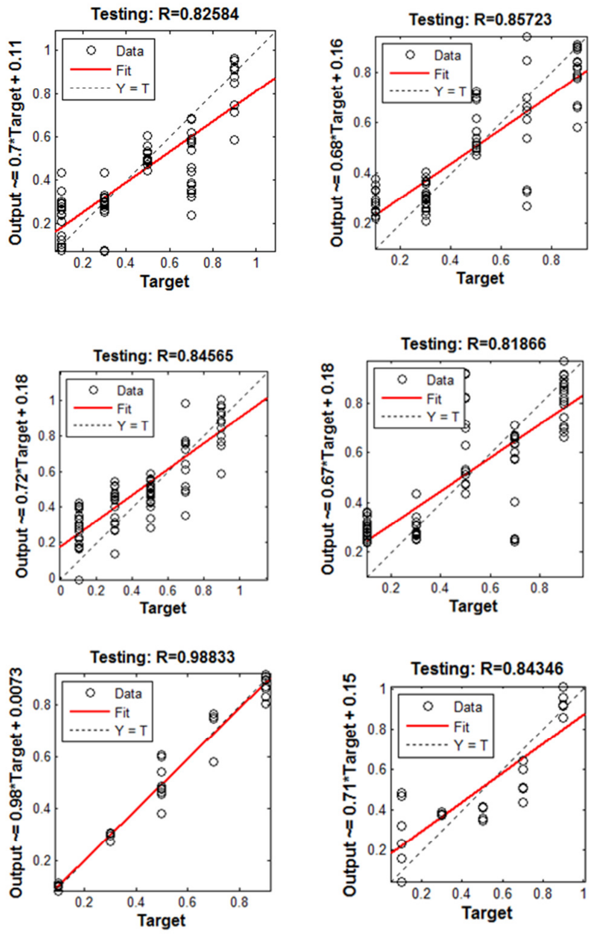

Figure 3. Regression Plot for (a) 2ft distance surfaces dataset using GLCM, (b) 2ft distance surfaces dataset using modified Zernike moments, (c) 3ft distance surfaces dataset using GLCM, (d) 3ft distance surfaces dataset using modified Zernike moments, (e) Brodatz dataset using GLCM, and (f) Brodatz dataset using modified Zernike moments.

V. CONCLUSIONS

The two methods for extracting features of image data had been tested. The GLCM method where different contrasts of neighbouring pixels are considered shows a higher performance than the modified Zernike moments. With the high variation of brightness at the time and place for collecting data, as well as the existing paint used to coat the concrete walls, wooden walls, and doors make it harder to capture the real texture of the surfaces.

For future experiments, modified Zernike moments with a higher number of amplitudes using different number of order and magnitude should also be tested, as well as using different other methods such as Gabor filters or wavelets, or combination of different extraction techniques. Other neural network algorithm such as Radial Basis Function (RBF) neural network, recurrent neural network (RNN), or probabilistic neural network (PNN) can also be implemented to test and investigate the performance.

ACKNOWLEDGMENT

The financial support received from the Research Acculturation Grant Scheme (RAGS): Vot R014, Office for Research, Innovation, Commercialization and Consultancy Management (ORICC), University Tun Hussein Onn Malaysia is gratefully acknowledged.

REFERENCES

[1] Yahya M. N., Otsuru T., Tomiku R., Okuzono T., NagasawanT. and Oda T. "Practical Subsystems To Identify Room Dimensions And Material Surfaces Using Photograph Image For Room Acoustic Simulation" AIJ Journal of Technology and Design, 39(6), 2012, pp. 6601-606.

[2] Hodgson M. and Scherebnyj K. "Estimation of the Absorption Coefficients of the Surfaces of Classrooms" Applied Acoustics, 67, 2006, pp. 936–944.

[3] Sujatha B., VijayaKumar V., Harini P. "A New Logical Compact LBP Co-Occurrence Matrix for Texture Analysis" International Journal of Scientific & Engineering Research, 3(2), 2012.

[4] Bridle J.S. "Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition, Neurocomputing" Algorithms, Architectures and Applications, 1989, pp. 227-236.

[5] Sim D.G., Kim H.K., Park R.H. "Invariant Texture Retrieval using Modified Zernike Moments" Image and Vision Computing, 22, 2004, pp. 331-342.

[6] Zhang B. & Pham T.D. "Multiple Features Based Two-stage Hybrid Classifier Ensembles for Subcellular Phenotype Images Classification" International Journal of Biometrics and Bioinformatics, 4(5), 2011.

[7] Haralick R.M., Shanmugam K, Dinstein I. "Textural features for image classification" IEEE Transactions on Systems, Man, and Cybernetics, 3(6), 1973, pp. 610-621.

[8] Roques E.J. & Murphy R.F. "Objective Evaluation of Differences in Protein Subcellular Distribution" 3, 2002, pp. 61 – 65.

[9] Boland M.V. & Murphy R.F. "A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells" Bioinformatics, 17(12), 2001, pp. 1213-1223.

[10] Zernike F. "Beugungstheorie des Schneidenverfahrens und Seiner Verbesserten Form, der Phasenkontrastmethode" Physica 1, (8), 1934, pp. 689–704.

[11] Khotanzad A. & Hong Y. H. "Invariant Image Recognition by Zernike Moments" IEEE Transactions on Pattern Analysis and MachineIntelligence, 12(5), 1990, pp. 489-497

[12] Chu S. R., Shoureshi, R., and Tenorio, M. "Neural networks for system identification" IEEE Control Systems Magazine, 1990, pp. 31-34.

[13] Jovanović O. "Identification of Dynamic System using Neural Network" 1998.

[14] Rumelhart D.E., Hinton G.E., Williams, Ronald J. "Learning representations by back-propagating errors" 323(6088), 1986, pp. 533–536.

[15] Brodatz P. "Textures: A Photographic Album For Artists And Designers" Dover Publications, New York, 1966.