Expert Systems with Applications 36 (2009) 5822–5829

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Applying wavelets transform, rough set theory and support vector machine for copper clad laminate defects classification

Te-Sheng Li *

Department of Industrial Engineering and Management, Minghsin University of Science and Technology, Hsinchu, Taiwan

ARTICLE INFO

Keywords:
Copper clad laminate
Discrete wavelet transform
Inverse discrete wavelet transform
Feature extraction
Rough set theory
Support vector machine

ABSTRACT

In this paper, we present a multi-resolution approach for the inspection of local defects embedded in homogeneous copper clad laminate (CCL) surfaces. The proposed method does not just rely on the extraction of local textural features in a spatial basis. It is based mainly on reconstructed images using the wavelet transform and inverse wavelet transform on the smooth subimage and detail subimages by properly selecting the adequate wavelet bases as well as the number of decomposition levels. The restored image will remove regular, repetitive texture patterns and enhance only local anomalies. Based on these local anomalies, feature extraction methods can then be used to discriminate between the defective regions and homogeneous regions in the restored image. Rough set feature selection algorithms are employed to select the feature. Rough set theory can deal with vagueness and uncertainties in image analysis, and can efficiently reduce the dimensionality of the feature space. Real samples with four classes of defects have been classified using the novel multi-classifier, namely, support vector machine. Effects of different sampling approach, kernel functions, and parameter settings used for SVM classification are thoroughly evaluated and discussed. The experimental results were also compared with the error back-propagation neural network classifier to demonstrate the efficacy of the proposed method.

1. Introduction

Visual inspection plays a vital role of quality control in manufacturing systems. Manual inspection is subjective and highly dependent on the inspector's expertise. In this study, we use wavelet transform combining rough set feature selector and a novel classifier, support vector machine to detect and classify copper clad laminate (CCL) surface defects. Most CCL surface defects are tiny involving obvious faulty items such as pinholes, stains, scratches, strips and other ill-defined defects. These unanticipated defects are small in size, refractive in light and cannot be described using explicit measures, making automatic defect detection difficult.

Because CCL defects are similar to texture patterns, the inspection task in this study can be classified as texture detection and classification. In this environment, one must solve the problem of detecting small surface defects that break the local homogeneity of a texture pattern. Most defect detection methods for texture surfaces generally involve computing a set of textural features in a sliding window. The system searches for significant local deviations in the feature values. The most difficult task in this approach is feature extraction and feature selection which completely embody the texture information in the image. There is no arbitrary approach for selecting and judging the appropriate features to use. Therefore, selecting an adequate feature set for a new texture in the training process can be very time-consuming, requiring the help of expert knowledge. The more features extracted from the texture image, the more sophisticated the classifier, such as Bayes (Gonzalez & Woods, 1992), maximum likelihood (Cohen, 1992), and neural networks (Van Hulle & Tollenaere, 1993) needed to discriminate the texture classes.

Numerous methods have been proposed to extract textural features either directly from the spatial domain or from the spectral domain. In the spatial domain, the more reliable and commonly used features are second-order statistics derived from spatial gray-level co-occurrence matrices (Haralick, Shanmugam, & Dinstein, 1973). They have been applied to wood inspection (Conners, McMillin, Lin, & Vasquez-Espinosa, 1983), carpet wear assessment (Siew & Hogdson, 1988), and roughness measurements for machined surfaces (Ramana & Ramamoorthy, 1996).

In the past, Fourier transform and Gabor filter were two of the spectral domain approaches used to find the more adequate textural features that were less sensitive to noise and intensity variation. The Fourier transform is a global approach that characterizes only the spatial-frequency distribution, but it does not consider this information in the spatial domain. The Gabor filters are well regarded as a joint spectral representation for analyzing textured images with highly specific frequency and orientation characteristics. This technique extracts features by filtering the textured image with a set of Gabor filter banks characterized by the

* Tel.: +886 3 5593142x3234; fax: +886 3 5595142.
E-mail address: jeff@must.edu.tw

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.07.040

frequency, the sinusoid orientation, and the Gaussian function scale. However, obtaining the optimal Gabor filter design is difficult because human intervention is required to select the appropriate filter parameters for the texture under study. If the center frequency of a selected Gabor filter does not match any of the important harmonics in the textured image, only noisy information is produced (Pichler, Teuner, & Hosticka, 1996).

Recently, multi-resolution decomposition schemes based on discrete wavelet transforms (DWT) have received considerable attention as alternatives for textural feature extraction. The multi-resolution wavelet representation allows an image to be decomposed into a hierarchy of localized subimages at different spatial frequencies (Chen & Lee, 1997). It divides the 2-D frequency spectrum of an image into a lowpass (smooth) subimage and a set of highpass (detail) subimages. The textural features are then extracted from the decomposed subimages in different frequency channels and different resolution levels. Chen and Kuo (1993) proposed a tree-structured wavelet transform for texture classification. They used an energy criterion to select the subimage for decomposition. A set of textural features is derived from the energy values of the dominant channels, and distance measures are then employed to discriminate between the texture classes. Lambert and Bock (1997) proposed a feature extraction approach for texture defect detection. The textural features were derived from the wavelet packet decomposition coefficients. Neural network and Bayes classifiers were used to evaluate the feature vector. The wavelet-based feature extraction methods were applied to industrial material inspection such as LSI wafers (Lee, Choi, Choi, & Choi, 1997; Maruo, Shibata, Yamaguchi, Ichikawa, & Ohmi, 1999), cold rolled strips (Sari-Sarraf & Goddard, 1998), and woven fabrics (Tsai & Hsieh, 1999).

Tsai and Hsieh (1999), Tsai and Hsiao (2001), and Tsai and Huang (2003) proposed a global approach based on a Fourier image reconstruction scheme for inspecting surface defects in textures. Their method does not depend on local texture features but detecting defects via the high-energy frequency components in the Fourier spectrum. In the restored image of a textured surface, the regular region with periodic lines in the original image will have an approximately uniform gray level, whereas the defective region will be distinctly preserved.

Rough set theory (RST), which was introduced by Pawlak in the early 1980s, (Pawlak, 1982) is a new mathematical tool that can be employed to handle uncertainty and vagueness. It focuses on the discovery of pattern in inconsistent data (Pawlak, 1996; Slowinski & Stefanowski, 1989) and can be used as the basis to perform formal reasoning under uncertainty, machine learning, and rule discovery (Yao, Wong, & Lin, 1997). The basic philosophy behind rough set theory is based on equivalence relations or indiscernibility in the classification of objects (Walczak & Massart, 1999). It does not need any a priori knowledge, such as probability in statistics, or the basic probability assignment in the Dempster–Shafer theory of evidence or membership grade in the fuzzy set theory. Using rough set theory as a data mining tool, it can explore the hidden patterns in the data set and obtain the decision rules. That is, rough set theory can be used for (a) reduction of feature spaces; (b) finding hidden patterns; and (c) generation of decision rules (Kusiak, 2001; Pawlak, 1982). In the past decades, the rough set theory has been successfully applied to many real-world problems in medicine, pharmacology, engineering, banking, financial and market analysis (Ahna, Cho, & Kim, 2000; Kusiak, 2001; the related methodology refers Refs. Pawlak (1982), Pawlak (1996)).

From Bayes classifiers to neural networks, there are many possible choices for classifier selection. Among these, support vector machines (SVM) would appear to be a good candidate because of their ability to generalize in high-dimensional spaces, such as spaces spanned by texture patterns. The appeal of SVM is based on its strong connection to the underlying statistical learning theory. That is, an SVM is an approximate implementation of the structural risk minimization (SRM) method (Vapnik, 1995; Vapnik, 1998). For several pattern classification applications, SVM have already been shown to provide better generalization performance than traditional techniques, such as neural networks (Schokopf et al., 1997).

The aim of this paper is to illustrate the potential of DWT, RST and SVM in CCL defects detection and classification. The proposed method incorporates the DWT, inverse DWT, and RST with the SVM classifier, using the selected features fed directly into the SVM classifier. As a result, the features obtained from the restored defects image are nonlinearly mapped into the SVM architecture. Since SVM was originally developed for two-class problems, its basic scheme is extended to multi-texture classification by adopting the *one-against-all* and *one-against-one* decomposition methods. This works by applying SVM to separate one class from all the other classes, or separate each pair among all classes. Thereafter, feature detection and classification are both performed in accordance with a unique criterion referred to as the classification rate.

This paper is organized as follows. Section 2 describes the DWT and inverse DWT transform in image, the rough set theory for defects feature selection and classification approach used to determine the multi-class defects by SVM. Section 3 presents experimental results on four categories of CCL surface defects classification using SVM classifier. Section 4 discusses the effects of the proposed method. The paper is concluded in Section 5.

2. Methodology

2.1. Wavelet transform and reconstruction

Wavelets are functions generated from one single function by dilations and translations. The basic idea of the wavelet transform is to represent any arbitrary function as a superposition of wavelets. Any such superposition decomposes the given function into different scale levels where each level is further decomposed with a resolution adapted to that level (Mallat, 1989).

This process continues until some final scale is reached. The values or transformed coefficients in approximation and detail images (sub-band images) are the essential features, which are as useful for texture discrimination and segmentation. Since textures, either micro or macro, have non-uniform gray-level variations, they are statistically characterized by the values in the DWT transformed sub-band images or the features derived from these sub-band images or their combinations. In other words, the features derived from these approximation and detail sub-band images uniquely characterize a texture.

The 2-D wavelet analysis operation consists of filtering and down-sampling horizontally using 1-D lowpass filter L and highpass H to each row in the image and produces the coefficient matrices $f_L(x, y)$ and $f_H(x, y)$. Vertically filtering and down-sampling follows, using the lowpass and highpass filters L and H to each column in $f(L)$ and $f(H)$, and produces four subimages $f(LL)$, $f(LH)$, $f(HL)$, and $f(HH)$ for one level decomposition. The 2-D pyramid algorithm can iterate on the smooth subimage $f(LL)$ to obtain four coefficient matrices in the next decomposition level.

For images, there exist an algorithm similar to the 1-D case for 2-D wavelets and scaling functions obtained from 1-D ones by tensorial product. This kind of 2-D DWT leads to a decomposition of approximation coefficients at level j in four components: the approximation at level $j+1$, and the details in three orientations (horizontal, vertical, and diagonal). The following chart describes the basic decomposition steps for images.