Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 41 (2008) 1744–1755
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

Performance evaluation of soft color texture descriptors for surface grading using experimental design and logistic regression

Fernando López[a,*], José Miguel Valiente[a], José Manuel Prats[b], Alberto Ferrer[b]

[a]*Department of Computer Engineering, (DISCA), Technical University of Valencia (UPV), Camino de Vera s/n, 46022 Valencia, Spain*
[b]*Department of Applied Statistics, Operation Research and Quality, (DEIOAC), Technical University of Valencia (UPV), Camino de Vera s/n, 46022 Valencia, Spain*

Received 17 January 2006; received in revised form 10 September 2007; accepted 19 September 2007

Abstract

This paper presents a novel approach to the question of surface grading, the *soft color texture descriptors* method. This method is extracted from an extensive evaluation process of several factors based on the use of two well established statistical tools: *experimental design* and *logistic regression*. The utility of different combinations of factors is evaluated in regard to the problem of automatic classification of materials such as ceramic tiles that need to be grouped according to homogeneous visual appearance, that is, the surface grading application. The set of factors includes the number of neighbors in the *k*-NN classifier (several values of *k* parameter), color space representation schemes (CIE Lab, CIE Luv, RGB, and grayscale), and color texture features (mean, standard deviation, 2nd–5th histogram moments). A factorial *experimental design* is performed testing all combinations of the above factors on a large image database of ceramic tiles. Accuracy estimates are computed using *logistic regression* to determine the best combinations of factors. From the point of view of machine learning the overall process conforms a wrapper approach able to select significant design choices (*k* parameter in *k*-NN classifier and color space) and carry out a feature selection within the set of color texture features at the same time. Experiments were repeated with alternate color texture schemes from the literature: *color histograms* and *centile-LBP*. Comparisons of methods are presented describing both accuracy estimates and runtimes.

Keywords: Surface grading; Automatic inspection; Experimental design; Logistic regression; Color; Texture; Feature selection; Wrapper techniques

1. Introduction

There are many industries currently manufacturing flat surface materials that need to split their production into homogeneous series grouped by the global appearance of the final product. These kinds of products are used as wall and floor coverings. Some of them are natural products such as marble, granite or wooden boards, and others are artificial stuff such as ceramic tiles.

In these industries the quality control stage is crucial in remaining competitive, and one of the most important quality problems is the non-uniformity of the visual aspect of the product within the same lot of a specific model. As the final product is used to form areas which are supposed to be uniform in appearance, the presence of pieces which look different or even slightly different is considered a serious quality fault. Nowadays, industries rely on human operators to perform the task of surface grading. However, human grading is subjective and often inconsistent between different graders [1]. Thus, automatic and reliable systems are needed. Also, capacity to inspect overall production at on-line rates is important.

In this paper we approach the use of *experimental design* and *logistic regression* [2–4], two well established statistical tools, to carry out an extensive evaluation of different factors that will lead to define a new method for the issue of surface grading. These statistical tools in conjunction provide a methodology for finding the best combinations of factors in a set of experiments by seeking to maximize the classification accuracy. The studied factors include, *k* parameter in *k*-NN classifier, color space, and a set of color texture descriptors. All these factors are evaluated at the same time using the statistical methodology in a wrapper

* Corresponding author. Tel.: +34 963877007; fax: +34 963877579.
E-mail address: flopez@disca.upv.es (F. López).

0031-3203/$30.00
doi:10.1016/j.patcog.2007.09.011

manner [5] (see Section 5). After the evaluation two design choices are fixed (k parameter and color space), and a feature selection is performed within the color texture descriptors. The mentioned evaluation procedure was carried out successfully in previous works where multivariate projection models were used for the detection of defects in orange fruits and ceramics [6].

The new approach to surface grading, resulting from the previous evaluation study, is called the *soft color texture descriptors* method, and is also compared with two other methods coming from the literature, *color histograms* [7] and *centile-LBP* [8]. *Experimental design* and *logistic regression* are also used to tune design choices in these literature methods.

The presented work demonstrates that a simple set of global statistics softly describing color and texture properties, together with the well-known classifier k-NN and perceptually uniform color spaces, are powerful enough to meet stringent factory performance requirements, 95% surface grading accuracy.

The paper is developed as follows. Section 2 presents an overview of the literature on the surface grading issue. The proposed method of *soft color texture descriptors* is described in Section 3. Literature methods chosen for comparison purposes are explained in Section 4. Section 5 presents the statistical procedure for the evaluation of factors based on *experimental design* and *logistic regression*. Section 6 deals with the experimental work and results. And finally, Section 7 summarizes the paper.

2. Surface grading

Surface grading is related with the automatic classification of flat pieces presenting random surface patterns. The aim of surface grading is to split production into different classes sorted by their global appearance which depends on color and texture properties.

In recent years many approaches to surface grading have been reported (see Table 1). Boukouvalas et al. [7] proposed color histograms and dissimilarity measures of these distributions to grade ceramic tiles.

Other works consider specific types of ceramic tiles; *polished porcelain* tiles, which imitate granite. These works include texture features. Baldrich et al. [9] proposed a perceptual approximation based on the use of discriminant features defined by human classifiers at factory. These features mainly concerned grain distribution and size. The method included grain segmentation and features measurement. Lumbreras et al. [10] joined color and texture through multiresolution decompositions on several color spaces. They tested combinations of multiresolution decomposition schemes (Mallat's, *àtrous* and wavelet packets), decomposition levels and color spaces (gray, RGB, Otha and Karhunen–Loève transform). Peñaranda et al. [11] used the first and second histogram moments of each channel in the RGB space.

Kauppinnen [1] developed a method for grading wood based on the percentile (or centile) features of histograms calculated for RGB channels. Kyllönen et al.'s approach [8] used color and texture features. They chose centiles for color, and LBP (Local Binary Pattern) occurrence histograms for texture description.

Lebrun and Macaire [12] described the surfaces of the Portuguese 'Rosa Aurora' marble using the mean color of the background and mean color, absolute density and contrast of marble veins. They achieved good results but their approach was very dependent on the visual properties of this marble. Finally, Kukkonen et al. [13] presented a system for grading ceramic tiles using spectral images. However, spectral images have the drawback of producing large amounts of data.

In the literature review, we found that many of these approaches specialized in a specific type of surface, others were not accurate enough or simply did not provide accuracy information, most of them performed a brief experimentation work and yet others did not take into account time restrictions in a real inspection process at factory. Thus, we think that surface grading is still an open field where more contributions are possible. In this sense, we attempt to fill these literature deficiencies by presenting an approach which is a generic method suitable for use in a wide range of random surfaces. Also we carry out extensive experimentation and achieve good accuracy with a representative data set of ceramic tiles. Our approach uses what we call *soft color texture descriptors*, which are simple and fast [to compute] global color and texture statistics. Thus, it is also appropriate for real-time compliance.

Table 1
Summary of surface grading literature

	Ground truth	Features	Time study	Accuracy (%)
Boukouvalas	Ceramic tiles	Color	No	–
Baldrich	Polished tiles	Color/texture	No	92.0
Lumbreras	Polished tiles	Color/texture	No	93.3
Peñaranda	Polished tiles	Color/texture	Yes	–
Kauppinen	Wood	Color	Yes	80.0
Kyllönen	Wood	Color/texture	No	–
Lebrun	Marble	Color/texture	No	98.0
Kukkonen	Ceramic tiles	Color	No	80.0

3. Soft color texture descriptors

The proposed method is simple, a set of statistical features describing color and texture properties are collected [14]. The features are computed in a perceptually uniform color space (CIE Lab or CIE Luv). These statistics form a feature vector used in the classification stage where the well-known k-NN approach was chosen as classifier [15]. We chose the k-NN classifier because it is simple and provides robust performance. In a recent work comparing several classification approaches, it still remains as one of the simplest providing very good performance [16].

CIE Lab and CIE Luv were designed to be perceptually uniform. The term 'perceptual' refers to the way that humans perceive colors, and 'uniform' implies that the perceptual difference between two coordinates (two colors) will be related to a measure of distance, which commonly is the Euclidean distance. Thus, color differences can be measured in a way close to the human perception of colors. These spaces were chosen