Computers and Electronics in Agriculture 110 (2015) 1–8

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Identifying barley varieties by computer vision

Piotr M. Szczypiński [a,*], Artur Klepaczko [a], Piotr Zapotoczny [b]

[a] Institute of Electronics, Lodz University of Technology, Wolczanska 211/215, 90-924 Lodz, Poland
[b] Department of Agri-Food Process Engineering, University of Warmia and Mazury in Olsztyn, Heweliusza 14, 10-718 Olsztyn, Poland

ARTICLE INFO

Article history:
Received 8 May 2014
Received in revised form 12 August 2014
Accepted 27 September 2014

Keywords:
Barley classification
Computer vision
Machine learning

ABSTRACT

Visual discrimination between barley varieties is difficult, and it requires training and experience. The development of automatic methods based on computer vision could have positive implications for the food processing industry. In the brewing industry, varietal uniformity is crucial for the production of high quality malt. The varietal purity of thousands of tons of grain has to be inspected upon purchase in the malt house.

This paper evaluates the effectiveness of identification of barley varieties based on image-derived shape, color and texture attributes of individual kernels. Varieties can be determined by means of discriminant analysis, including reduction of feature space dimensionality, linear classifier ensembles and artificial neural networks, with high balanced accuracy ranging from 67% to 86%. The study demonstrated that classification results can be significantly improved by standardizing individual kernel images in terms of their anteroposterior and dorsoventral orientation and performing additional analyses of wrinkled regions.

1. Introduction

Barley is a major cereal grain used for both human consumption and animal feed. A total of 47 spring barley varieties (30 malting varieties and 17 fodder varieties) and 17 winter barley varieties are currently included in the Polish National List. Since certain varieties are suitable for specific applications, adequate selection is a crucial step in barley grain processing. For instance, barley grain used for malting should be characterized by low protein content, uniform size and high quality. Therefore, grain has to be controlled for varietal purity and technological quality at every stage of processing (Hulasare et al., 2003).

There are several methods of cereal grain testing, including immunological analysis, DNA analysis, high-performance liquid chromatography, protein electrophoresis and isoenzyme analysis. Most of those methods are labor intensive and expensive, and the analyses can only be performed by specialized laboratories. An alternative approach involves visual evaluation of grain for varietal classification in line with the International Rules for Seed Testing developed by the International Seed Testing Association (ISTA). The following physical parameters are generally evaluated during a visual inspection: kernel color, kernel shape, shape of lemma base, ventral crease hairs, rachis hairs, teeth on lateral dorsal nerves, wrinkling of the lemma and palea, shape and hairs of lodicules. This technique is much easier to apply than chemical methods, but the reliability of visual evaluation is largely dependent on the skills and experience of the evaluator.

Computer image analysis, a rapid and low-cost technique, is an alternative method that evaluates selected physical attributes of kernels. Varietal classification of cereal grains by computer vision has been widely discussed in literature (Mebatsion et al., 2013; Neuman et al., 1987; Visen et al., 2002; Zapotoczny, 2011a, 2011b). However, none of the cited studies involved correction of image orientation, and classification models relied on analyses of whole kernel regions only. By contrast, an algorithm (Szczypiński and Zapotoczny, 2012) was developed to analyze kernel images in terms of the dorsoventral (the side with or without the crease is visible) and anteroposterior (germ–brush direction along the semi-symmetric axis) orientation of kernels. Kernel images were segmented into specific areas that were inspected individually. The proposed procedure was motivated by the assumption that the shape and properties of a kernel's wrinkled region are genetically determined. A kernel's ventral and dorsal surfaces could have different properties that should be analyzed by distinct algorithms.

Image classification usually involves computation of image attributes (features). An image attribute is a numeric quantity that characterizes the image or its fragment. Numerous algorithms for feature computation focus on different aspects of image

* Corresponding author. Tel.: +48 426312642; fax: +48 426362238.
E-mail address: piotr.szczypinski@p.lodz.pl (P.M. Szczypiński).

http://dx.doi.org/10.1016/j.compag.2014.09.016
0168-1699/

appearance, such as brightness and color distribution, texture, shape and topology of a region. Those attributes can be examined in cereal analysis. Neuman et al. (1989) analyzed the color attributes of cultivars belonging to different wheat classes and demonstrated significant differences between varieties. Paliwal et al. (1999) relied on color attributes and shapes described by the Fourier transform of a radial function to distinguish between kernels of wheat, barley, oats and rye. Image histogram parameters computed from monochrome images of a wheat bulk sample proved to be useful for the estimation of moisture content (Manickavasagan et al., 2008). Zielinska et al. (2012) used image texture attributes and morphological (shape) properties to classify four red clover cultivars.

Thousands of attributes can be used to characterize an individual image, depending on its specific application (Szczypinski et al., 2009). Thus, every analyzed image is described by attribute vectors with thousands of dimensions. A statistical analysis is then carried out to examine the distribution of attribute vectors in high-dimensional spaces that are not easy to handle. For instance, overfitting may take place in highly complex models where the number of parameters significantly exceeds the number of observations (Berthold and Hand, 2003). This leads to poor predictive performance as the classifier trained on a set of examples fails to correctly recognize other data. In cereal classification, the artificial neural network is often a method of choice (Goyal, 2013). Neural network classifiers such as multilayer perceptrons allow for nonlinear decision boundaries, and they easily fit exemplary data if the hidden layer contains a large number of neurons. Unfortunately, a higher number of neurons increases the risk of overtraining (Jain et al., 2000), which also leads to poor predictive performance.

The goal of this paper was to evaluate the feasibility of image-driven classification of 11 barley varieties and to identify subsets of attributes with the highest discriminatory ability. The study also set out to establish whether information related to a kernel's wrinkled subregion and its dorsoventral and germ–brush orientation improves classification performance.

2. Materials and methods

Grain samples used in this study were obtained from selected farms in the Region of Warmia and Mazury, NE Poland. The experimental material consisted of 11 varieties of two-rowed barley: Afrodita, Blask, Bordo, Conchita, Kormoran, Mercanda, Prymus, Serwal, Signora, STH and Victoriana. The differences between kernels belonging to the evaluated varieties are unlikely to be identified by an untrained assessor. Effective discrimination between varieties requires training and experience.

In accordance with ISTA recommendations, grain should be rinsed with water before a sensory evaluation to expose its texture and color. In this study, grain was not rinsed to improve its surface condition because in line with the applied methodology, unwashed grain is more difficult to classify. Therefore, the proposed methodology would be validated more effectively if tested on unwashed grain.

The images were acquired in color at the resolution of 400 dpi, 24 bits per pixel, in the Epson 4990 flatbed scanner. The scanner was placed in a compartment covered with black velvet and its top cover was removed. The kernels were positioned manually using a template with holes punctured specifically for this purpose. When all holes were filled with kernels, excess kernels and, consequently, the template were removed. This technique supported uniform spacing of 450 kernels per scan. However, individual kernels can be rotated, and they can produce images of the dorsal or the ventral side. The resulting images show relatively bright and disjoint kernel areas on a dark background (Fig. 1). In total, 33 scanned images of 11 barley varieties were acquired, with three

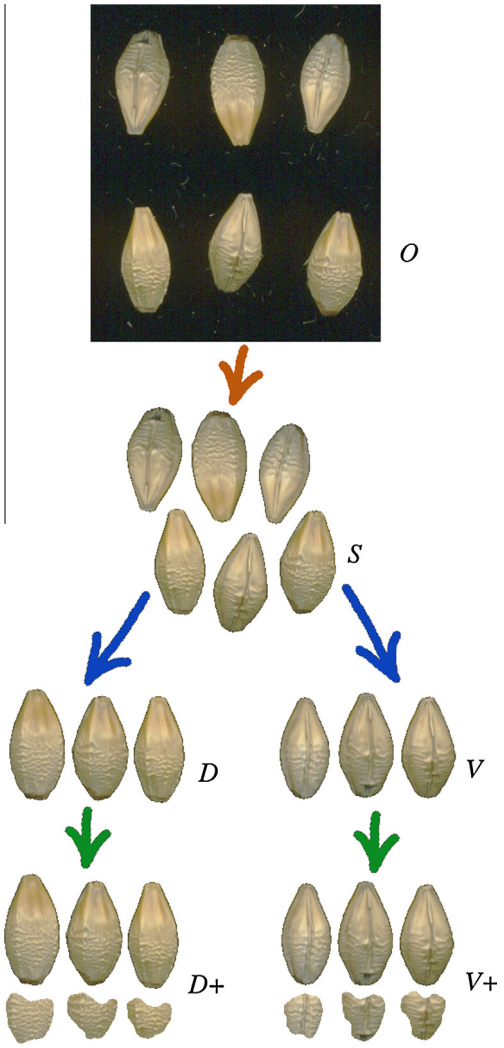

Fig. 1. Image preprocessing diagram and data sets for analysis: *O* – fragment of the original image, *S* – kernels identified after image segmentation, *D* – adjusted orientation of the dorsal view, *V* – adjusted orientation of the ventral view, *D+*– dorsal view and the corresponding wrinkled areas, and *V+*– ventral view and the corresponding wrinkled areas.

images per variety, which resulted in more than 13,000 individual objects for analysis.

One of the objectives was to determine whether identification of dorsoventral orientation, correction of the germ–brush direction and supplementary analysis of wrinkled subregions would improve classification results. The other goal was to identify attributes that best discriminate barley varieties. The experiment had the following design: kernels were identified in images and described individually in terms of their numerical properties – attributes. Images of individual kernels were rotated with respect to the germ–brush direction and split into two groups based on their dorsoventral orientation. The resulting images were repeatedly numerically characterized. Wrinkled kernel areas were determined and their attributes were computed to supplement the characteristics of rotated kernel images. The analysis produced data sets characterizing kernel images at three increasingly advanced preprocessing levels. A discriminant analysis was performed separately on the three data sets, and it involved attribute selection and classification. The analysis revealed kernel properties that were most suitable for varietal discrimination and demonstrated whether the proposed image processing method affected classification performance.