



Original papers

Wheat grain classification by using dense SIFT features with SVM classifier



Murat Olgun^{a,1}, Ahmet Okan Onarcan^{b,2}, Kemal Özkan^{b,2}, Şahin Işık^{b,2}, Okan Sezer^c, Kurtuluş Özgişi^c, Nazife Gözde Ayter^{a,1}, Zekiye Budak Başçiftçi^{a,1}, Murat Ardiç^c, Onur Koyuncu^{c,*}

^a Eskişehir Osmangazi University, Faculty of Agriculture, Department of Field Crops, Eskişehir, Turkey

^b Eskişehir Osmangazi University, Faculty of Engineering and Architecture, Department of Computer Engineering, Eskişehir, Turkey

^c Eskişehir Osmangazi University, Faculty of Science and Letters, Department of Biology, 26480 Eskişehir, Turkey

ARTICLE INFO

Article history:

Received 19 February 2015

Received in revised form 23 December 2015

Accepted 29 January 2016

Available online 15 February 2016

Keywords:

Wheat identification

Grain classification

Dense SIFT features

k-means clustering

Support Vector Machines

Bag of Word Model

ABSTRACT

The demand for identification of cereal products with computer vision based applications has grown significantly over the last decade due to economic developments and reducing the labor force. With this regard, we have proposed an automated system that is capable to classify the wheat grains with the high accuracy rate. For this purpose, the performance of Dense Scale Invariant Features (DSIFT) is evaluated by concentrating on Support Vector Machine (SVM) classifier. First of all, the concept of k-means clustering is operated on DSIFT features and then images are represented with histograms of features by constituting the Bag of Words (BoW) of the visual words. By conducting an experimental study on a special dataset, we can make a commitment that the proposed method provides the satisfactory results by achieving an overall 88.33% accuracy rate.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Wheat (*Triticum* spp.) is a major food source in the world and it is commonly grown in most of the countries. It has wide adaptability to various environments including irrigated and dry land conditions; this explains that why it prevails in food production of the world. Healthy wheat production mainly requires certified pure grain use and in production process grains shouldn't be mixed with different genotypes. The identification of varieties requires some knowledge of the appearance of grains and is assisted by information regarding its grain appearance. Confusion in grain appearance is a frequently encountered problem, especially in Turkey, where the number of wheat genotypes is very large. Commercially, wheat can be categorized with two groups as grain hardness (soft, medium-hard and hard) and appearance (color, degree of damage by insects or fungi, shriveling, shape of embryo). This separation can be branch out by considering growing habit (spring or winter). Also each subclasses can be ordered by their grades depending upon the price of a wheat stock as applying premiums or penalties by taking such properties (rain, heat, frost, insect and mould dam-

age) and the cleanliness (dockage and foreign material) of the wheat lot into account. Grading factors can also be varied with respect to quality of grain such as protein content and sedimentation test weight in wheat trading as emphasized in Peña (2002).

Classification of wheat grains can be made in two ways as manually or automatically. In manual way, the type and the quality of wheat grain is specified based on an expert judgement. However, the judgement of the expert is inaccurate for some cases when the difference between the variety and the quality of wheat species is very close to each other. Thereby, the decision of the expert could result in financial loss, bankrupt or loss of confidence on behalf of manufacturer. In another way, using image processing and pattern recognition algorithms, called expert systems, with a purpose for wheat classification are slightly more accurate than manual way. In this way, the shape, the texture, the length and the color features are considered and combined to construct the feature vector, which represents the image with reduced dimension by discriminative characteristics. Later, the obtained feature set is put forward as input into a machine learning algorithm, i.e., K-Nearest Neighbor (K-NN), Decision Tree (DT) or Artificial Neural Network (ANN), to obtain a concise decision about its label. It should be noted that using expert systems depending upon the pattern recognition methodologies is more effective, fairer, cheaper and faster when compared with expert judgement.

* Corresponding author.

E-mail address: omurkoyuncu@gmail.com (O. Koyuncu).

¹ Tel.: +90 222 3242991/4862.

² Tel.: +90 222 2393750/3275.

Going through the previous studies on wheat classification, we found a few studies as counted something on the fingers of two hands. In Zayas et al. (1989) developed a structural prototype for discrimination of wheat and non-wheat components in a grain sample, by using a multivariate discriminant analysis technique. In referred paper, the execution time for discrimination between wheat and non-wheat was elapsed as about 10 s. Also, priorly, the same authors carried out another two experiments as discrimination between Arthur and Arkan wheats (Zayas et al., 1985) and discrimination between wheat classes (hard red winter, soft red winter and hard red spring) and varieties by benefiting from image processing methods (Zayas et al., 1986) with maximum 99.52 (1 out of 209 grain was not identified correctly) and 83% accuracy rates in the test stage, respectively. In another work, Majumdar and Jayas had reported four different approaches for classification commercial cereal grains by employing morphology models (Majumdar and Jayas, 2000a), color models (Majumdar and Jayas, 2000b), texture models (Majumdar and Jayas, 2000c) and hybrid one as combination of morphology, color and texture models (Majumdar and Jayas, 2000d). Moreover, the discrimination of wheat class and variety by digital image analysis of whole grain samples (Neuman et al., 1987), the performance of pattern recognition methods for the separation of cereal grains (corn, soybeans, rice, sorghum grain, barley and wheat) (Lai et al., 1986), the identification of Australian wheat varieties (Myers and Edsall, 1989), identification of Canada Western Red Spring (CWRS) wheat (Sapirstein and Kohler, 1999) and analyzing the shape Indian wheat varieties (Shouche et al., 2001) by image analyses techniques, particularly with morphological operations (Zapotoczny et al., 2008) had been presented and realized by concentrating on some computer programs.

Recently, a crowded set of machine learning techniques have been experimented to determine some wheat types. In a work (Guevara-Hernandez and Gomez-Gil, 2011), a machine vision system was developed for classification of wheat and barley grains based on the 21 morphological features, 6 of them were color features, 72 of them were texture features, totally 99 features. To reduce high dimension of feature set and eliminate the ones which gain the less contribution, the sequential forward feature selection has utilized after zero mean normalization, weighting and sorting processes with respect to larger Fisher discriminant ratio. The experimented approach achieves an accuracy rate that is higher than 99% when conducting on two classes. In another work (Ronge and Sardeshmukh, 2014), the widely known four Indian wheat seed varieties (Lokvan Gujrat, Lokvan MP, MP sure, Khapali) are classified by extracting 131 textural features as 32 gray level textural features, 31 Local Binary Pattern (LBP) features, 31 Local Similarity Patterns (LSP) features, 15 Local Similarity Number (LSN) features, 10 Gray Level Co-occurrence matrix (GLCM) features and 12 Gray Level Run-length Matrix (GLRM) features. By observing the results, the average accuracy values are 66.68% and 39% in case of intra class classification with ANN and K-NN classifiers, respectively. Again, the Multi-Layer Perceptron (MLP) Neural Network based classification system (Pazoki and Pazoki, 2013) has been developed to distinguish the six classes of rain fed wheat grain cultivars with 21 statistical features. The average accuracies returned from the system have been reported as 86.48% and 87.22% as before and after applying the utility additive (UTA) algorithm to ignore the less promised features. Moreover, six varieties (Demir, Gün, İkizce, Mızrak, Seval, Tosunbey) of bread wheat are classified by using the common vector approach (CVA), which is a subspace based feature extraction method (Gulmezoglu and Gulmezoglu, 2015). By using the CVA, firstly, a common vector which represents common or invariant properties of each class is computed and then a given test image is assigned to its label based

on minimum distance criteria. The average accuracy for 500 test images has been reported as 36.7. Also, the impact of four machine learning algorithms (One-R, J48, IBK and Apriori) (Romero et al., 2013) for the prediction of wheat yield from several phenotypic plant traits has been examined by using the machine learning software WEKA (Hall Mark, 2003). Authors emphasized that among the aforementioned algorithms, the best overall accuracy obtained from Apriori, which is noted as 90% when executed to predict durum wheats for three cities. In given study, the measured yield components have considered as features in case of prediction.

Despite the good performance of proposed methods, the limitations of them become apparent when considering big datasets with high sample size and dimension, which are not carried out in related works. In another word, training a pattern recognition system with the low number of samples does not give stable and precise results in terms of accuracy and effectiveness in reality. We consider this detail and present a solution to close the gap with a different wheat classification approach. With this aim, we have proposed an algorithm based on the dense SIFT features, which is a popular feature extraction method utilized in tasks of object recognition (Loncomilla and Ruiz-del-Solar, 2005), object tracking (Zhou et al., 2009) and image retrieval (Ledwich and Williams, 2004). The reason for choosing DSIFT instead of SIFT, is attributed to its good results by obtaining descriptors from every locations, when compared with performance generated from specific locations as performed in SIFT algorithm. Moreover, the study on comparison of feature detectors and descriptors for object class matching (Hietanen et al., 2015) emphasizes that the performance DSIFT is superior to some feature selection methods. Also, SVM is selected as an optimized classifier that generates an optimal decision boundary between classes. The obtained features are concentrated on SVM classifier. The experimental results show that the proposed method gives satisfactory, realistic and convincing accuracy rates when making trials on 40 classes as consists of 160 samples per each class.

The rest of paper is organized as follows. Section 2 introduces the materials and procedure for wheat classification. The related work is presented in Section 3. In Section 4, the performance of proposed method on special dataset is discussed with objective evaluation measures. Finally, a conclusion is touched and future work is discussed.

2. Material and methods

The classification procedure of wheat objects is conducted on most of the state of art feature extraction methods. The related work is summarized with Fig. 1. By inspecting Fig. 1, some stages that are required for a particular classification problem are considered as dense SIFT features that are extracted in (i), (ii) the k -means clustering is operated on DSIFT features, (iii) the BoW model from the histogram clustering features are acquired and finally SVM classifier is utilized on BoW model in order to identify wheat objects.

2.1. Dense scale invariant feature transform

Scale-invariant feature transform (or SIFT) (Lowe, 1999, 2004) is a computer vision algorithm to represent and identify objects with some local and distinctive features, proposed by David Lowe in 1999. The main objective of SIFT algorithm obtains the descriptors that are invariant to rotation, scale, variation in illumination and robust to geometric transformations which are isometry, similarity, affine, projective and inversion transformations.

Download English Version:

<https://daneshyari.com/en/article/84050>

Download Persian Version:

<https://daneshyari.com/article/84050>

[Daneshyari.com](https://daneshyari.com)