Available online: www.notulaebiologicae.ro
Print ISSN 2067-3205; Electronic 2067-3264
Not Sci Biol, 2016, 8(1):93-97. DOI: 10.15835/nsb.8.1.9743
Notulae Scientia Biologicae

The Influence of Colour Features on Seed Identification Using Machine Vision

Sepideh ANVARKHAH[1*], Ali Davari Edalat PANAH[2], Alireza ANVARKHAH[3]

[1]Department of Crop Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran
[2]Research and Development Center of Astan-e-Qouds, Mashhad, Iran
[3]Azad University of Mashhad, Faculty of Agriculture, Department of Crop Science, Iran; sa_sah_sa25@yahoo.com (*corresponding author)

Abstract

Little studies have been done on morphology of medicinal plants seeds. This paper presents an automatic system for medicinal plant seed identification and evaluates the influence of colour features on seed identification. Six colour features (means of red, green and blue colours of the seed surface, as well as means of hue, intensity and saturation) were extracted by algorithm and applied as network input. Different combinations of colour features (one, two three, four, five and six colour features) were used to find out the most accurate combination for seed identification. Results showed that the six colour feature was the most accurate combination for seed identification (99.184% and 87.719% for training and test of neural network respectively). One colour feature had the lowest average accuracy values for seed identification (3.120% and 2.771%). In general, increasing the number of colour features increased the total average of accuracy values.

Keywords: automatic system, colour grading, hue, medicinal species, seed identification, seed morphology

Introduction

Although medicinal plants play an important role in the drug industry and health care, and thus draw much attention, few studies have been conducted on their seed identification. Knowledge of seed morphology is important in theoretical botany and could be useful within seed identification for seed testing, seed quarantine, seed dispersal and soil seed bank studies (Jensen, 1995).

Recently, machine vision has become a useful technology for quick seed control and identification (Ureña *et al*., 2001). Recent advances in hardware and software have enabled machine vision and imaging systems to identify, analyse and display finer details of objects from their digital images (Paliwal *et al*., 2003).

In laboratories, the most common method for cultivars' identification is to compare morphological characteristics of seeds with standard samples. Such characters include length, width, thickness, shape, weight, hilum colour and seed coat colour (Cope *et al*., 2012).

Studies showed that colour is a useful characteristic to divide different varieties based on seed coloration such as red-, amber- and white- coloured, but could not be useful for categorizing them into classes (Zhang *et al*., 2012; Lev-Yadun and Ne'eman, 2013).

Colour is one of the most important features in seeds classification and grading. Different seeds and their varieties are identified by their colours. Thomson and Pomeranz (1991) classified the Western Canadian wheat to six groups using a limited set of colour features (mean red (R), green (G) and blue (B) pixel reflectance features). In general, the red, white and amber coloured wheat types were well separated, while some confusion existed between certain red kernel types. Also, Luo *et al*. (1999) set an experiment for separation of healthy seeds of Western Canadian wheat from damaged ones using colour features.

Seed colour images might be used also to describe seed quality and hardness, fungal damages, viral diseases, as well as for separating immature seeds apart of mature ones (Ducournau *et al*., 2004; Liu *et al*., 2005). In addition, early identification of weed seeds for one crop might be a major interest in the agricultural industry. It can also be useful for chemical control of weed growth (Granitto *et al*., 2002).

Studies showed that there is a correlation between seed colour and seed quality. For example, it has been reported that the seeds of naturally occurring yellow seeded genotypes *B. rapa*, *B. juncea* and *B. carinata* contained greater oil, higher protein and lower fibre contents than the seeds of black/brown seeded genotypes of these species (Rahman and McVetty, 2011). The yellow seeded *Brassica* genotypes of these species had a thinner and more translucent seed coat, lower hull proportion with a bigger embryo and consequently greater oil and protein percentage (Rahman and McVetty, 2011). Proanthocyanidins and tannins are the major compounds involved in seed coat pigmentation. These are deposited in the seed coat of black/brown seeded *Brassica* genotypes and reduce the digestibility of seed meal for livestock. However, the seeds'

Received: 01 Jan 2016. Received in revised form: 09 Feb 2016. Accepted: 07 Mar 2016. Published online: 16 Mar 2016.

coat of black/ brown seeded *Brassica* genotypes contained more fiber and less protein than those of yellow seeded genotypes. Therefore, *B. napus* lines have been developed from interspecific crosses with related species, namely *B. rapa*, *B. oleracea* spp. *alboglabra*, *B. juncea* and *B. carinata* (Rahman and McVetty, 2011).

One of the most important attributes for introducing sesame grains in the market was seed colour (Pandey *et al*., 2013; Zhang *et al*., 2013). Although most are light coloured, there is a wide variability in sesame seed coat colour, which varies from white to black. Due to the importance of this trait for the export market, seed colour is a central target in sesame breeding programs; however, there are few studies on the inheritance of this essential seed attribute, and determination of genetic factors affecting any trait is necessary to establish useful breeding programs (Laurentin and Benítez, 2014).

Seeds acquire primary dormancy during their development to enhance adaptation, as the capacity of wild species, to diverse environments, by distributing germination over time and space. Domestication tends to reduce dormancy by selection for rapid and uniform germination. Differentiation in seed dormancy between cereal crops and wild relatives has been associated with some factors such as seed morphologies (Guo *et al*., 2000). For example, the most persistent type of weedy rice is red rice, which is characterized by a red pericarp colour. Red rice has strong seed dormancy (Lim and Ha, 2013).x

In some cases, there is a correlated relationship between seed coat colour and seed quality. Some studies showed that seed lots of red clover visually inspected in terms of seed colour were separated based on a larger colour spectrum thereafter, by digital colour measurement equipment, as seed colour yellow, purple, brown and mixed. Results revealed that seed coat colour of red clover could be preferred as an indicator of seed quality and seedling growth ability. Yellow coloured seeds lots of red clover had higher vigour and seed quality than other colours. Mean germination time (MGT) and electrical conductivity (EC 4 h) test showed significant differences among the seed coat colour. Meanwhile, tests also showed a highly significant correlation in emergence and seedling percentage in salt stress conditions (Atis *et al*., 2011).

Some seeds are valued according to their appearance, and thus colour is the most important factor for grading (Copeland and McDonald, 2012). The purpose of the current research was to determine the influence of colour features on seed automatic identification.

Materials and Methods

Grain samples

Seeds from 75 species of medicinal plants (Table 1) were used for this study.

Seeds were photographed using a Dinolite Digital Microscope model 4050 with 640 × 480 to 1024 × 768 pixel resolution at 30- to 80-times magnification, depending on their original size. A database containing 1,800 images of the 75 species was constructed.

Algorithm development

For algorithm development MATLAB 7.9 (Version 2009b) software and windows Vista (Service Pack 1) were used. Employed hardware was an IBM compatible laptop (model Vostro 1500 from DELL Company). In the algorithm, the seed image was segmented from the background image, and its features were extracted and used for the neural network training (Anvarkhah *et al*., 2012).

Six colour features were extracted by algorithm and applied as network input:

- Mean of red colour of seed surface (R)
- Mean of green colour of seed surface (G)
- Mean of blue colour of seed surface (B)
- Hue means (H)
- Intensity (I)
- Saturation means (Sa)

Different combinations of colour features were used to find out the most accurate combination for seed identification.

Results and Discussion

One colour feature

Table 2 shows the total average values of training and test parts of neural network, when using each colour feature individually. The use of hue had the highest accuracy values of training and test with values of 9.239% and 8.771% respectively. However, employing one colour feature led to a low rate of accuracy values. For example, no accurate identification was shown when using red, blue and saturation features separately (0%). The rest of the colour features tested hereby had low accuracy percents, below 8%.

Two Colour Features

It was noted that two colour features led to a more accurate identification. However, combinations of red, green, blue and saturation with hue caused 0% of accuracy values, while by using other features paired two by two the results had higher values. The most accurate identification was shown within the combination of hue and intensity, which led to 24.184% and 19.298% for training and test parts of neural network respectively (Table 3).

Three colour features

Table 4 shows the training and test accuracy values obtained by using three colour features within the different colour combinations. Except of two combinations ([red + hue + intensity], [blue + hue + saturation], both with 0%), all others had accuracy values of 30-99% for training and 20-85% for test parts of neural network.

Four colour features

Except two combinations of [hue + saturation + red + green] and [hue + saturation + red + blue], all other combinations of four colour features caused above 60% and 50% for training and test parts of neural network respectively (Table 5). It seems that using triple effect of

Table 1. Scientific name and family of the 75 species of medicinal plant seeds used in the experiment

No.	Scientific name	Family	No.	Scientific name	Family
1	*Allium sativum*	Alliaceae	41	*Astragalus siliquosus*	Leguminoseae
2	*Amaranthus albus*	Amaranthaceae	42	*Astragalus squarosus*	Leguminoseae
3	*Amaranthus retroflexus*	Amaranthaceae	43	*Hedysarum scoparium*	Leguminoseae
4	*Anethum graveolens*	Apiaceae	44	*Onobrychis radiata*	Leguminoseae
5	*Foeniculum vulgare*	Apiaceae	45	*Onobrychis sp.*	Leguminoseae
6	*Dorema ammoniacum*	Apiaceae	46	*Securigera securidaca*	Leguminoseae
7	*Prangos ferulaceae*	Apiaceae	47	*Allium ampeloprasum persicum*	Liliaceae
8	*Achillea millefolium*	Asteraceae	48	*Allium cepa L*	Liliaceae
9	*Anthemis tinctoria*	Asteraceae	49	*Allium schoenoprasum L.*	Liliaceae
10	*Calendula officinalis*	Asteraceae	50	*Linum usitatissimum*	Linaceae
11	*Centaurea cyanus*	Asteraceae	51	*Althaea officinalis*	Malvaceae
12	*Chrysanthemum superbum*	Asteraceae	52	*Malva dendromorpha*	Malvaceae
13	*Cynara scolymus*	Asteraceae	53	*Malva sylvestris*	Malvaceae
14	*Pseudohandelia umbellifera*	Asteraceae	54	*Caryophyllus aromaticus*	Myrtaceae
15	*Silybum marianum*	Asteraceae	55	*Oenothera biennis*	Onagraceae
16	*Echinacea purpurea*	Asteraceae	56	*Sesamum indicum*	Pedaliacae
17	*Rudbeckia hirta* 'Marmalade'	Asteraceae	57	*Digitalis purpurea*	Plantaginaceae
18	*Silybum marianum*	Asteraceae	58	*Plantago major*	Plantaginaceae
19	*Taraxacum sect*	Asteraceae	59	*Plantago ovata*	Plantaginaceae
20	*Borago officinalis*	Boraginaceae	60	*Plantago purshii*	Plantaginaceae
21	*Lepidium perfoliatum*	Brassicaceae	61	*Plantago psyllium*	Plantaginaceae
22	*Lepidium sativum*	Brassicaceae	62	*Agropyron pectiniforme*	Poaceae
23	*Cannabis sativa*	Cannabaceae	63	*Avena sativa*	Poaceae
24	*Saponaria officinalis*	Caryophyllaceae	64	*Bromus kopetdaghensis*	Poaceae
25	*Erotia ceratoides*	Chenopodiaceae	65	*Hordeum bulbosum*	Poaceae
26	*Kochia prostrata*	Chenopodiaceae	66	*Melica persica*	Poaceae
27	*Cucurbita pepo*	Cucurbitaceae	67	*Pennisetum orientale*	Poaceae
28	*Ricinus communis*	Euphorbiaceae	68	*Portulaca oleracea*	Portulacaceae
29	*Fumaria parviflora*	Fumariaceae	69	*Aquilegia vulgaris*	Ranunculaceae
30	*Hyssopus officinalis*	Lamiaceae	70	*Nigella sativa*	Ranunculaceae
31	*Marrubium vulgare*	Lamiaceae	71	*Ruta graveolens*	Rutaceae
32	*Melissa officinalis* L.	Lamiaceae	72	*Physalis alkekengi*	Solanaceae
33	*Melissa axillaris* (Benth.) Bakh.f.	Lamiaceae	73	*Hyoscyamus niger*	Solanaceae
34	*Ocimum album*	Lamiaceae	74	*Hyoscyamus pusillus*	Solanaceae
35	*Ocimum basilicum*	Lamiaceae	75	*Zygophyllum eurypterum*	Zygophyllaceae
36	*Origanum majorana*	Lamiaceae			
37	*Salvia dorrii*	Lamiaceae			
38	*Salvia sclarea*	Lamiaceae			
39	*Satureja hortensis*	Lamiaceae			
40	*Ziziphora clinopodioides*	Lamiaceae			

hue, saturation and red with features of green and blue may cause training errors.

Five colour features

Table 6 shows identification accuracy using five colour features. These combinations led to training and test accuracy values higher than 90% and 75% respectively, for all features' combinations.

Six colour features

The most accurate identification was shown using combination of six colour features (99.18% and 87.71% for training and test of neural network respectively) (Table 7). In general, increasing the number of colour features increased the total average of accuracy (Anvarkhah *et al.*, 2013).

Comparison among colour features combination

Fig. 1 shows the average accuracy values of different colour combinations. Increasing the number of colour features led to higher accuracy values. Combination of six colour features was the most accurate combination with

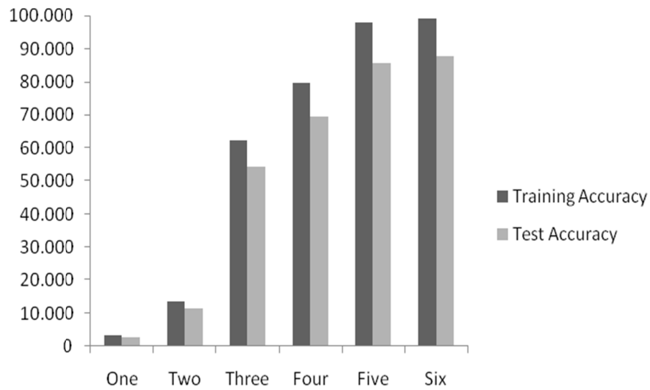

Fig. 1. Average accuracy values using different colour features combinations

Table 2. Accuracy values using one colour feature

Feature	Training accuracy (%)	Test accuracy (%)
Green	3.508	4.347
Red	0	0
Blue	0	0
Hue	9.239	8.771
Intensity	5.978	3.508
Saturation	0	0
Average	3.120	2.771

Table 3. Accuracy values using two colour features

Feature	Training accuracy (%)	Test accuracy (%)
Red +Green	43.070	43.859
Green+Blue	14.402	9.649
Red +Blue	18.750	14.035
Red+Hue	0	0
Red+Intensity	10.597	9.649
Red+Saturation	19.021	16.666
Green+Hue	0	0
Green+Intensity	5.163	4.386
Green+Saturation	20.380	16.666
Blue+Hue	0	0
Blue+Intensity	8.695	7.017
Blue+Saturation	18.206	14.035
Hue+Intensity	24.184	19.298
Hue+Saturation	0	0
Intensity+Saturation	18.750	15.789
Average	13.414	11.403

Table 4. Accuracy values using three colour features

Feature	Training accuracy (%)	Test accuracy (%)
Red+Green+Blue	52.445	49.122
Red+Green+Hue	92.119	76.315
Red+Green +Intensity	32.065	28.947
Red +Green Saturation	91.032	80.701
Red+Blue+Hue	93.750	78.070
Red+Blue+Intensity	33.695	32.456
Red+Blue+Saturation	65.217	61.403
Green+Blue+Hue	78.532	68.421
Green+Blue+Intensity	40.489	40.350
Green+Blue+Saturation	59.782	52.631
Red+Hue+Intensity	0	0
Red+Hue+Saturation	98.641	84.210
Red+Intensity+Saturation	72.282	66.666
Green+Hue+Intensity	55.434	44.736
Green+Hue+Saturation	97.826	84.210
Green+Intensity+Saturation	66.847	59.649
Blue+Hue+Intensity	83.967	69.298
Blue+Hue+Saturation	0	0
Blue+Intensity+Saturation	32.608	25.438
Hue+Intensity+Saturation	97.826	84.210
Average	62.227	54.341

accuracy values of 99.184% and 87.719% for training and test of neural network. The lowest training and test accuracy values belonged to one colour feature with 3.121% and 2.771%.

In general, increasing the number of colour features increased the total average of accuracy (Anvarkhah *et al.*, 2013).

Colour is one of the most important features in seeds classification and grading. Different seeds and their varieties are identified by their colours. Thomson and Pomeranz (1991) classified the Western Canadian wheat to six groups using a limited set of colour features (mean Red (R), Green (G) and Blue (B) pixel reflectance features). In general, the red, white and amber coloured wheat types were well separated, while some confusion existed between certain red kernel types. Also, Luo *et al.* (1999) set an experiment for separation of healthy seeds of Western Canadian wheat from damaged ones using colour features.

Table 5. Accuracy values using four colour features

Feature	Training accuracy (%)	Test accuracy (%)
Red+Green+Blue+Hue	98.641	85.964
Red+Green+Blue+Intensity	64.673	57.894
Red+Green+Blue+Saturation	93.206	82.456
Hue+Intensity+Saturation+Red	99.184	86.842
Hue+Intensity+Saturation+Green	98.369	82.456
Hue+Intensity+Saturation+Blue	98.641	84.210
Hue+Intensity+Red+Green	97.282	83.333
Hue+Intensity+Red+Blue	98.097	85.964
Hue+Intensity+Green+Blue	88.858	78.947
Hue+Saturation+Red+Green	0	0
Hue+Saturation+Red+Blue	0	0
Hue+Saturation+Green+Blue	98.641	83.333
Intensity+Saturation+Red+Green	93.478	82.456
Intensity+Saturation+Red+Blue	80.706	72.807
Intensity+Saturation+Green+Blue	85.869	75.438
Average	79.709	69.473

Table 6. Accuracy values using five colour features

Feature	Training accuracy (%)	Test accuracy (%)
Red+Green+Blue+Hue+Intensity	98.641	86.842
Red+Green+Blue+Hue+Saturation	99.184	87.719
Red+Green+Blue+Intensity+Saturation	93.206	78.947
Hue+Intensity+Saturation+Red+Green	99.184	86.842
Hue+Intensity+Saturation+Red+Blue	99.184	86.842
Hue+Intensity+Saturation+Green+Blue	98.913	85.964
Average	98.052	85.526

Table 7. Accuracy values using six colour features

Feature	Training accuracy (%)	Test accuracy (%)
Red+Green+Blue+Hue+Intensity+ Saturation	99.184	87.719
Average	99.184	87.719

Conclusions

Different combinations of colour features (one, two three, four, five and six colour features) were used to find out the most accurate combination for seed identification by machine vision and algorithm determinations. Results showed that the six colour feature was the most accurate combination for seed identification for training and test of neural network respectively, while employing one colour feature led to a low rate of accuracy values, with a lack of accurate identification when using red, blue and saturation features separately.

References

Atiş I, Atak M, Can E, Mavi K (2011). Seed coat color effects on seed quality and salt tolerance of red clover (*Trifolium pratense*). International Journal of Agriculture and Biology 13(3):363-368.

Anvarkhah S, Hajeh-Hosseini M, Davari-Edalat-Panah A, Rashedmohassel MH (2013). Medicinal plant seed identification using machine vision. Seed Science and Technology 41(1):107-120.

Anvarkhah S, Khajeh-Hosseini M, Davari-Edalat-Panah A (2012). Seed identification of ten rangeland species based on machine learning using combination of RBF and Feed Forward neural networks. International Journal of Agriculture and Crop Sciences 4(14):993-1004.

Cope JS, Corney D, Clark JY, Remagnino P, Wilkin P (2012). Plant species identification using digital morphometrics: A review. Expert Systems with Applications 39(8):7562-7573.

Copeland LO, McDonald M (2012). Principles of seed science and technology. Springer Science & Business Media.

Ducournau S, Feutry A, Plainchault P, Revollon P, Vigouroux B, Wagner MH (2004). An image acquisition system for automated monitoring of germination rate of sunflower seeds. Computer and Electronics in Agriculture 44:189-202.

Granitto PM, Navone HD, Verdes PF, Ceccatto HA (2002). Weed seeds identification by machine vision. Computer and Electronics in Agriculture 33:91-103.

Guo Q, Brown JH, Valone TJ, Kachman SD (2000). Constraints of seed size on plant distribution and abundance. Ecology 81(8):2149-2155.

Jansen PI (1995). Seed production quality in *Trifolium balansae* and *T. resupinatum*: The role of seed colour. Seed Science and Technology 23:353-364.

Laurentin H, Benítez T (2014). Inheritance of seed coat color in sesame. Pesquisa Agropecuria Brasiliera 49(4):290-295.

Lev-Yadun S, Ne'eman G (2013). Bimodal colour pattern of individual *Pinus halepensis* Mill. seeds: a new type of crypsis. Biological Journal of the Linnean Society 109(2):271-278.

Lim SH, Ha SH (2013). Marker development for the identification of rice seed color. Plant Biotechnology Reports 7(3):391-398.

Liu ZY, Cheng F, Ying YB, Rao XQ (2005). Identification of rice seed varieties using neural network. Journal of Zhejiang University-Science B 6(11):1095-1100.

Luo X, Jayas DS, Symons SJ (1999). Identification of damaged kernels in wheat using a colour machine vision system. Journal of Cereal Science 30:49-59.

Paliwal J, Visen NS, Jayas DS, White NDG (2003). Cereal grain and dockage identification using machine vision. Biosystems Engineering 85(1):51-57.

Pandey SK, Das A, Dasgupta T (2013). Genetics of seed coat color in sesame (*Sesamum indicum* L.). African Journal of Biotechnology 12(42):6061.

Rahman M, McVetty PBE (2011). A review of Brassica seed color. Canadian Journal of Plant Sciences 91(3):437-446.

Thomson WH, Pomeranz Y (1991). Classification of wheat kernels using three dimensional image analysis. Cereal Chemistry 68:357-361.

Ureña R, Rodríguez F, Berenguel M (2001). A machine vision system for seeds quality evaluation using fuzzy logic. Computer and Electronics in Agriculture 32(1):1-20.

Zhang H, Miao H, Wei L, Li C, Zhao R, Wang C (2013). Genetic analysis and QTL mapping of seed coat color in sesame (*Sesamum indicum* L.). PLoS One 8(5):p.e63898.

Zhang X, Liu F, He Y, Li X (2012). Application of hyperspectral imaging and chemometric calibrations for variety discrimination of maize seeds. Sensors 12(12):17234-17246.