Journal of Agricultural Science (2006), **144**, 221–227. © 2006 *Cambridge University Press*
doi:10.1017/S0021859606006010 *Printed in the United Kingdom* First published online 6 March 2006

Image analysis of wheat grains developed in different environments and its implications for identification

J. K. SAINIS, S. P. SHOUCHE AND S. G. BHAGWAT*
Bhabha Atomic Research Centre, Mumbai-400 085, India
(*Revised MS received 11 November 2005*)

SUMMARY

Varietal identification is an important aspect of crop research and utilization. Identification using computer-based image analysis could be an alternative to visual identification. However, the effectiveness of image analysis systems needs to be established under various real conditions. Three wheat varieties were sown on three different dates. Variation in the grain size and shape of these varieties, brought about by changes in the environmental conditions, was measured using Comprehensive Image Processing Software (CIPS). Some parameters showed considerable grain-to-grain variation, which was either inherent or due to environmental changes during grain filling. Euclidean distances were calculated using either means of all the parameters (ED1), or using only those parameters that did not show a high coefficient of variation (ED2). For samples of the same variety sown at different times, Euclidean distances were smaller compared with samples of different varieties, indicating that grains of the same variety resembled one another. By using the criterion of minimum Euclidean distance it was possible to distinguish between varieties, in spite of variation in grain shape and size due to environmental conditions. It was possible to identify correctly an unknown sample, taken as a test case.

INTRODUCTION

Exact identification of varieties of crop plants is important in several areas such as research in plant breeding, seed certification, examination for breeders' rights, for commercial purposes and for appropriate end use. For instance, whether a given lot of wheat is of use for making bread, biscuit, *chapati* or pasta is decided by its quality characteristics, which are variety related, and hence correct identification of the variety is important.

Digital vision techniques are being developed for accurate quantitative estimation of morphological features of grains and other plant parts. Advances in digital image technology offer scope to estimate morphological features of grains objectively. Previously, attempts have been made to use digital image analysis for variety identification in wheat (Keefe & Draper 1986; Sapirstein 1995) and in other crops (Pietrzak & Fulcher 1995; Sakai *et al*. 1996; Keefe 1999). Visen *et al*. (2004) used neural networks to identify barley, oats, rye, wheat and durum wheat samples. Over 150 features related to grain colour and texture were studied and it was concluded that 20 features were adequate to classify them with complete accuracy, excepting oats.

Previously, in most image analysis systems, the images of grains were acquired using charge coupled device (CCD) or video cameras. In such systems the illumination conditions and focusing distances are critical to acquire accurate and comparative data. Shouche *et al*. (2001) used a scanner in transparency mode to acquire images of wheat grains and used these to measure 45 shape- and size-related parameters.

The Euclidean Distance (ED) calculated using these data enabled Shouche *et al*. (2001) to distinguish between 15 bread wheat varieties. Similarly, genetically related lines could also be distinguished (Bhagwat *et al*. 2003). Although the morphological features of grains are heritable, they are influenced by agronomic practices and climatic conditions. It is important, therefore, to investigate whether image analysis can be used for variety identification in practical situations, e.g. where there are possible variations in grain morphology caused by environmental conditions. Since variation in sowing time is

* To whom all correspondence should be addressed. Email: sbhagwat@apsara.barc.ernet.in

known to affect grain filling, three bread wheat varieties were sown on three different dates to simulate different environments. The present paper reports results on a study of the morphometry of wheat grains of the three varieties sown on different dates and their similarity or distinctness as indicated by Euclidean distance.

MATERIALS AND METHODS

At Trombay (0 m asl, 72·54°E and 18·55°N), the annual rainfall is 1800–2000 mm, occurring only during June to September. Wheat is cultivated during November to March, which is the dry period. Wheat cultivation is always irrigated. A few days' difference in sowing time results in different climatic conditions for the crop. A crop sown before the optimum date faces higher than optimum temperature in the early vegetative phase, whilst sowing after the optimum date results in higher than optimum temperature at grain-filling stage. Higher temperature during grain filling results in poor grain filling, seen as a higher proportion of smaller and shrivelled grains.

Bread wheat varieties Kalyansona (K) and Sonalika (S) and a genetic stock TW-1 (Tw) (for convenience, all are referred to hereafter as varieties) from the collection at the Bhabha Atomic Research Centre, India were sown in a field at Trombay in experimental plots of loamy soil. Three sowing dates were used in order to produce grains with different morphological characteristics within the same variety: 27 November 2002, 19 December 2002 and 3 January 2003. There were four replications for each variety at each sowing date, which were randomized. The sowing plan was identical for the three sowing dates. At maturity, which was judged by complete drying of plants and grains, each replicate was harvested separately. Grains for image analysis were taken at random, avoiding broken grains.

A flat bed scanner (HP Scanject 7400C automatic document feeder) with transparency adapter and HP Precision Scan Pro Software was used for image acquisition. A personal computer (Pentium IV 300 MHz, 256 MB RAM, with super VGA monitor having a 24-bit display card) was used for image analysis. The scanner resolution was set to 300 dpi. All the images were grabbed using identical illumination settings. The images were stored in .tif format for further analysis.

Fifty grains of each replicate were imaged and analysed using a Comprehensive Image Processing Software package (CIPS), and data on 45 morphometric parameters were collected for each grain, as described by Shouche *et al.* (2001) and Bhagwat *et al.* (2003). Grain area, perimeter, compactness, major axis length and minor axis length were measured in pixels. The axis ratio, shape factors 1 to 5, spread and slenderness were estimated. Moments such as standard moments (m00, m10, m01, m11, m20, m02, m30, m03, m12, m21), central moments (μ00, μ11, μ20, μ02, μ30, μ03, μ12, μ21), normalized central moments (η11, η20, η02, η30, η03, η12, η21) and invariant moments (ϕ1, ϕ2, ϕ3, ϕ4, ϕ5, ϕ6, ϕ7) were calculated. Standard formulae described in Gonzalez & Woods (1993) and Jain (1995) were used to calculate standard, central, normalized central and invariant moments. In the digital images of cereal grains, the grains had different orientations. Since the values of 2-D moments vary with position of kernels in the field of view, each grain in the image was rotated for normalization of orientation. After labelling the grain regions, the orientation of the grains was set to vertical by applying normalized rotation using the theory of moments, that provides raw and central moments as under

$$m_{p,q} = \iint f(x,y)x^p y^q \, dx\, dy, \quad p, q = 0, 1, 2 \ldots \quad (1)$$

$$\mu_{p,q} = \iint f(x, y)(x - x_{mean})^p (y - y_{mean})^q \, dx\, dy, \quad p, q = 0, 1, 2 \ldots \quad (2)$$

where $m_{p,q}$ represents standard moments, $\mu_{p,q}$ represents central moments and f(x, y) represents 2-dimensional grey level distribution in the image. The average values of the x, y (x, y) coordinates in the grain region are x_{mean}, y_{mean}. The angle of orientation of each grain is given by the angle of axis of least moment of inertia of each grain and is given by

$$\text{Angle } \theta = \tan^{-1}[2\mu_{1,1}/(\mu_{2,0} - \mu_{0,2})]. \quad (3)$$

All grains were rotated by an angle θ about the centre of mass. The rotated points were translated back to compensate for the original translation of the origin. Holes and clipping effects in the image were removed in the rotated image. The effect of jagged edges due to small angle of rotation was removed by applying bilinear interpolation. For computation of raw moments for each grain, origin of calculation was shifted to the lowest coordinate of the bounding rectangle of that grain, thus eliminating absolute displacement of x and y coordinates.

Geometric parameters such as perimeter, area, compactness etc. were calculated using the thresholded binary image of grains. All the images in these measurements were thresholded using a fixed value which resulted in better preservation of shape without any significant loss in information. Area of a region was defined as the number of pixels contained within its boundary. The length of its boundary represented the perimeter of a region. The compactness of a region was estimated as (Perimeter)$*$(Perimeter)/(4$*\pi*$Area). Using the values of axis length, perimeter and area, the rest of the geometric parameters were computed. Spread and slenderness of

https://doi.org/10.1017/S0021859606006010 Published online by Cambridge University Press

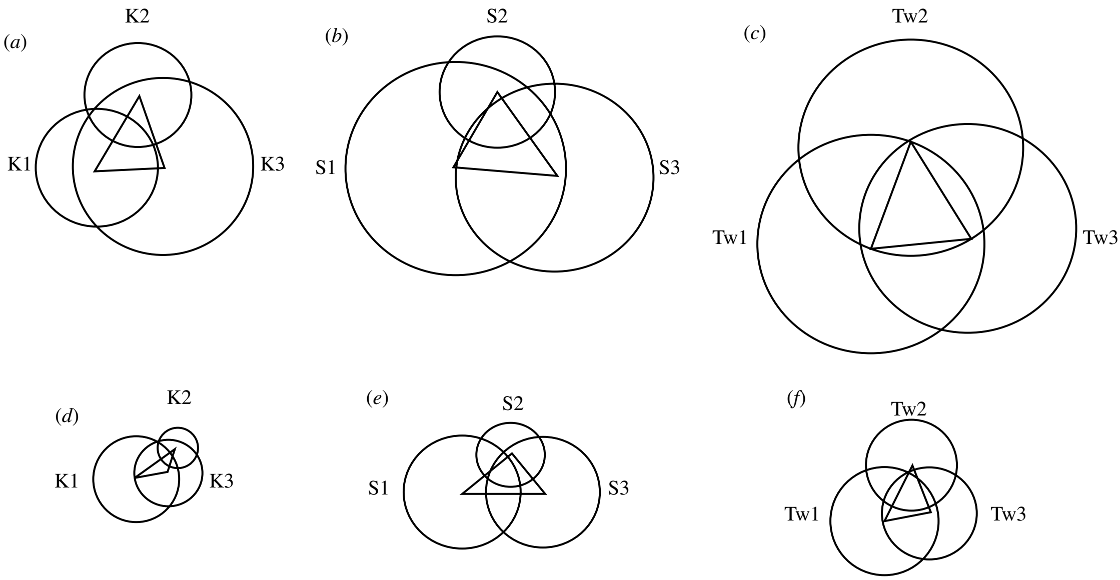

Fig. 1. Graphical representation of spread of ED1 and ED2 values: Taking average ED among the replicates at a given sowing time as radius in centimetres, a circle represents the spread of variation. For convenience, all the ED values were multiplied by a factor of 5 in all the graphical presentations. Kalyansona, Sonalika and TW-1 are represented as K, S and Tw, and the three sowing dates as 1, 2 and 3. Figs 1(*a*), (*b*) and (*c*) show the spread of ED1 within K, S and Tw sown at dates 1, 2 and 3 and Figs 1(*d*), (*e*) and (*f*) show ED2. The diagrams have been kept to the same scale.

individual grains was computed by moment analysis according to Hu (1962) using the binary image.

$$\text{SPREAD} = \mu_{20} + \mu_{02} \quad \text{......... using Eqn (2),}$$

$$\text{SLENDERNESS} = \sqrt{(\mu_{20} - \mu_{02})^2 + 4 * \mu_{11}{}^2}. \quad \text{......... using Eqn (2)}$$

The formulae described by Symons & Fulcher (1988) were used for calculations of various shapes. Binary images were used for calculations.

Shape factor 1: $(4\pi\text{Area})/\text{Perimeter}^2$ (i.e. 1/compactness)
Shape factor 2: Major axis length/Area
Shape factor 3: Area/(Major axis length)^3
Shape factor 4: Area/((Major axis length/2) (Major axis length/2) π)
Shape factor 5: Area/((Major axis length/2) (Minor axis length/2) π).

The average value of each parameter was calculated using 50 grains from each replicate plot, and Euclidean distance (ED) was calculated. Normalized ED was calculated using the following formula.

$$\text{Euclidean Distance (x, y)} = [\Sigma((x_i - y_i)/(x_i + y_i))^2]^{1/2}$$

where x_i is mean of sample variety and y_i is the mean of reference variety with which the comparison is being made for a given parameter (Bhagwat *et al.* 2003). All of the 45 parameters were used initially, to calculate ED1. The parameters showed different coefficients of variation (data not shown). The parameters with coefficients of variation greater than 10% (m01, m11, m03, m21, μ11, μ12, μ21,η11, η03, η12, η21) were omitted for calculation of another criterion, ED2.

A sample of the variety Kalyansona sown in December 2001 and harvested in March 2002 was used as an unknown sample for testing the system.

RESULTS

The results showed that in spite of growing in different environments, samples of the same variety showed the lowest ED values. There were exceptions, mainly due to the high degree of variation within TW-1. EDs among replicates at a particular sowing time for variety TW-1 were higher. TW-1 has short stature and duration with very little buffering capacity to maintain constant grain size under adverse conditions. Sonalika showed lower variation than TW-1 and Kalyansona showed the least variation. Hence, the results for the commercial cultivars were clearer (Fig. 1*a*, *b*, *d*, *e*); also, the third sowing date resulted in much poorer grain growth.

ED between replicates of each sowing date

Table 1 shows averages for ED1 and ED2 between replicates of each variety at each sowing date. The

https://doi.org/10.1017/S0021859606006010 Published online by Cambridge University Press

Table 1. *Average Euclidean distance between replicates*

Sowing Date	ED1 (average of 4 replications)	S.E.	ED2 (average of 4 replications)	S.E.
Kalyansona				
27 Nov 2002	0·45	0·046	0·32	0·057
19 Dec 2002	0·40	0·050	0·16	0·026
3 Jan 2003	0·67	0·112	0·25	0·028
Sonalika				
27 Nov 2002	0·81	0·124	0·43	0·117
19 Dec 2002	0·43	0·067	0·25	0·043
3 Jan 2003	0·73	0·157	0·42	0·108
TW-1				
27 Nov 2002	0·83	0·159	0·41	0·093
19 Dec 2002	0·82	0·133	0·34	0·083
3 Jan 2003	0·79	0·095	0·35	0·051

Euclidean distance ED1 and ED2 were calculated by pair-wise comparison of the replicates. Average ED values (n = 7) are represented.

Table 2. *Average Euclidean distance between replicates of different sowing dates*

	Sowing date x (4 replications)	Sowing date y (4 replications)	ED1	S.E.	ED2	S.E.
Kalyansona						
	27 Nov 2002	19 Dec 2002	0·58	0·048	0·23	0·024
	27 Nov 2002	3 Jan 2003	0·53	0·054	0·27	0·028
	19 Dec 2002	3 Jan 2003	0·56	0·053	0·20	0·018
Sonalika						
	27 Nov 2002	19 Dec 2002	0·59	0·076	0·33	0·053
	27 Nov 2002	3 Jan 2003	0·78	0·060	0·44	0·050
	19 Dec 2002	3 Jan 2003	0·63	0·051	0·33	0·037
TW-1						
	27 Nov 2002	19 Dec 2002	0·77	0·089	0·42	0·039
	27 Nov 2002	3 Jan 2003	0·75	0·074	0·36	0·046
	19 Dec 2002	3 Jan 2003	0·72	0·081	0·35	0·044

Euclidian distance ED1 and ED2 were calculated by pair-wise comparison of the replicates of the varieties sown at different dates. Average ED values (n = 16) are represented.

value of ED1 varied due to sowing date, although not uniformly for the varieties. ED1 for Kalyansona was lowest for the sowing on 19 December and highest for that on 3 January. For Sonalika ED1 was lowest for the sowing on 19 December. TW-1 showed relatively high ED1 values irrespective of the sowing date, suggesting inherent variation independent of environmental differences in this variety.

The ED2 values were lower compared with the corresponding ones for ED1, due to inclusion of only those parameters with lower coefficients of variation. Kalyansona and Sonalika showed lowest ED2 values among replicates in the 19 December sowing. TW-1 showed high ED2 values among replicates irrespective of sowing date. When the magnitudes of ED2 were compared, the ranking of sowing dates was the same for Sonalika but the rankings altered for Kalyansona and TW-1.

ED between different sowing dates

Table 2 shows average values of ED1 and ED2 between sowing dates for each variety, taking into account all replicates. For Kalyansona, ED1 was least between sowing dates 27 November and 3 January. The higher value of ED1 in comparisons involving the sowing on 19 December indicated larger differences in grain shape and size. The corresponding ED2 values were also lower. The lowest value of ED2

https://doi.org/10.1017/S0021859606006010 Published online by Cambridge University Press

Table 3. *Average Euclidean distance between replicates of varieties at different sowing dates*

Variety (x) (4 replications)		Variety (y) (4 replications)					
Variety	Sowing Date	Variety	Sowing Date	ED1	S.E.	ED2	S.E.
Kalyansona	27 Nov 2002	Sonalika	27 Nov 2002	1·45	0·083	1·20	0·063
Kalyansona	27 Nov 2002	Sonalika	19 Dec 2002	1·25	0·052	1·07	0·047
Kalyansona	27 Nov 2002	Sonalika	3 Jan 2003	1·43	0·051	1·12	0·044
Kalyansona	19 Dec 2002	Sonalika	27 Nov 2002	1·45	0·062	1·25	0·052
Kalyansona	19 Dec 2002	Sonalika	19 Dec 2002	1·25	0·027	1·12	0·029
Kalyansona	19 Dec 2002	Sonalika	3 Jan 2003	1·35	0·020	1·17	0·026
Kalyansona	3 Jan 2003	Sonalika	27 Nov 2002	1·46	0·079	1·19	0·057
Kalyansona	3 Jan 2003	Sonalika	19 Dec 2002	1·26	0·049	1·07	0·038
Kalyansona	3 Jan 2003	Sonalika	3 Jan 2003	1·41	0·055	1·13	0·037
Kalyansona	27 Nov 2002	TW-1	27 Nov 2002	1·29	0·096	1·06	0·058
Kalyansona	27 Nov 2002	TW-1	19 Dec 2002	1·13	0·092	0·90	0·066
Kalyansona	27 Nov 2002	TW-1	3 Jan 2003	1·34	0·072	1·08	0·055
Kalyansona	19 Dec 2002	TW-1	27 Nov 2002	1·26	0·056	1·12	0·042
Kalyansona	19 Dec 2002	TW-1	19 Dec 2002	1·14	0·055	0·97	0·050
Kalyansona	19 Dec 2002	TW-1	3 Jan 2003	1·29	0·040	1·15	0·038
Kalyansona	3 Jan 2003	TW-1	27 Nov 2002	1·13	0·058	1·09	0·050
Kalyansona	3 Jan 2003	TW-1	19 Dec 2002	1·19	0·080	0·96	0·055
Kalyansona	3 Jan 2003	TW-1	3 Jan 2003	1·19	0·080	1·13	0·045
Sonalika	27 Nov 2002	TW-1	27 Nov 2002	1·00	0·087	0·59	0·059
Sonalika	27 Nov 2002	TW-1	19 Dec 2002	1·20	0·065	0·77	0·045
Sonalika	27 Nov 2002	TW-1	3 Jan 2003	1·19	0·057	0·73	0·049
Sonalika	19 Dec 2002	TW-1	27 Nov 2002	0·79	0·081	0·46	0·054
Sonalika	19 Dec 2002	TW-1	19 Dec 2002	0·97	0·042	0·62	0·024
Sonalika	19 Dec 2002	TW-1	3 Jan 2003	0·98	0·048	0·60	0·044
Sonalika	3 Jan 2003	TW-1	27 Nov 2002	0·82	0·085	0·45	0·058
Sonalika	3 Jan 2003	TW-1	19 Dec 2002	1·01	0·068	0·60	0·040
Sonalika	3 Jan 2003	TW-1	3 Jan 2003	0·93	0·088	0·55	0·061

Euclidean distance ED1 and ED2 were calculated by pair-wise comparison of the replicates of each variety with replicates of other variety at each of the sowing time mentioned. Average ED values (n = 16) are represented.

was observed between Kalyansona sown on 19 December and on 3 January, indicating that these samples most closely resembled one another. The ED1 values for variety TW-1 were generally higher than those of Sonalika and Kalyansona.

ED between different varieties at different sowing dates

To estimate similarities or differences between the varieties at different sowing dates, ED1 and ED2 were calculated using all combinations of replicates and sowing dates (Table 3). The ED1 values between varieties were higher than for replicates of the same variety at each sowing date (Table 1) or for replicates of the same variety in pair-wise comparisons of sowing dates (Table 2). The corresponding ED2 values were lower than ED1 (Table 3).

Comparison with unknown sample

Table 4 shows that the average ED1 of the 'unknown' variety was lowest in comparisons with Kalyansona sown on 19 December. The corresponding ED2 values were lower, with the least being between Kalyansona sown on 27 November. Kalyansona showed lower ED2 values at all sowing dates compared to Sonalika or TW-1, indicating that the 'unknown' variety had more in common with Kalyansona at all sowing dates than with the other two varieties. The 'unknown' sample was in fact a sample of the variety Kalyansona sown in December 2001 and harvested in March 2002, a year before the 'calibration' samples used.

Graphical representation

The relationships within and between the samples can be represented graphically. The value of the ED between replicates for a variety can be used as the radius of a circle to denote the variability inherent in the sample. The distance between different samples, whether different sowing dates or different varieties, can be represented as lines joining the centres of the circles for each sample. In the present study these lines form a triangle.

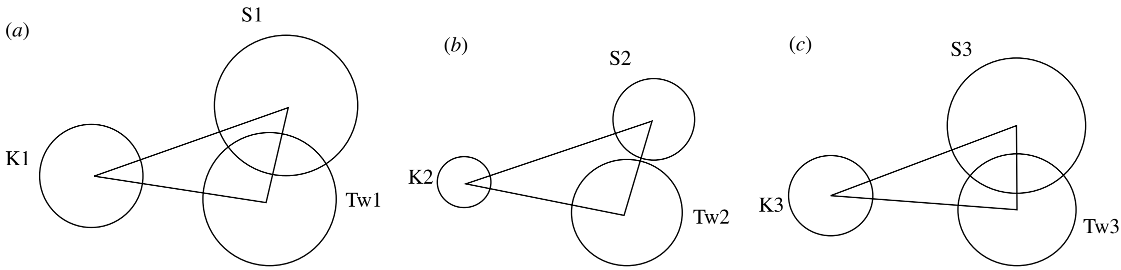

Fig. 2. Graphical representation of the spread of ED2 values. (*a*) K1, S1, Tw1; (*b*) K2, S2, Tw2; (*c*) K3, S3, Tw3.

Table 4. *Average Euclidean distance (ED1, ED2) of the unknown with all the varieties*

Variety	Sowing date	Average ED1	S.E.
Kalyansona	19 Dec 2002	0·66	0·043
Kalyansona	3 Jan 2003	0·83	0·122
Sonalika	19 Dec 2002	0·85	0·038
Kalyansona	27 Nov 2002	0·88	0·095
TW-1	19 Dec 2002	0·89	0·096
TW-1	27 Nov 2002	0·91	0·105
Sonalika	3 Jan 2003	0·93	0·022
TW-1	3 Jan 2003	0·98	0·099
Sonalika	27 Nov 2002	1·07	0·121
Variety	Sowing date	Average ED2	S.E.
Kalyansona	27 Nov 2002	0·48	0·097
Kalyansona	3 Jan 2003	0·49	0·054
Kalyansona	19 Dec 2002	0·53	0·021
TW-1	19 Dec 2002	0·56	0·126
Sonalika	19 Dec 2002	0·69	0·068
TW-1	27 Nov 2002	0·70	0·121
TW-1	3 Jan 2003	0·75	0·104
Sonalika	3 Jan 2003	0·76	0·049
Sonalika	27 Nov 2002	0·84	0·131

Euclidean distance ED1 and ED2 were calculated by pair-wise comparison of the unknown with four replicates of the varieties sown at different dates. Average ED values (n = 4) are represented.

Fig. 1*a* illustrates the spread within Kalyansona at the three sowing dates, when estimated by ED1. Large overlaps between the circles indicate the extent of similarity of the Kalyansona samples between the different sowing dates. Figures 1*b* and 1*c* show similar diagrams for Sonalika and TW-1. Clearly ED1 is not good for distinguishing seeds produced at different sowing dates. ED2 values can also be used to draw similar diagrams (Fig. 1*d, e, f*). Since ED2 was based on those parameters with lower coefficients of variation, the circles are much smaller.

ED2 values were also used to draw such graphical representation of the three varieties sown at three different sowing dates (Fig. 2). Kalyansona could be distinguished from the other two varieties at all three sowing dates, while overlap was observed for Sonalika and TW-1 sown on the 27 November and 3 January. There were no overlaps between the three varieties sown on 19 December, indicating that they were all distinguishable from each other.

DISCUSSION

Computer-based image analysis has the potential to fulfil a long-standing requirement for a rapid, inexpensive and objective method for varietal identification. Research worldwide is using a variety of hardware and software. A major difficulty for correct identification is the variation in seed size and shape caused by environmental conditions. Samples of the same variety coming from different environments have to be correctly identified, in spite of the variation in size and shape.

Euclidean distances (ED1 and ED2) were calculated using 45 shape- and size-related parameters. In theory, the ED for two samples of the same variety would be zero. However, sampling and experimental errors result in values of ED >0. Even so, related samples should show lower ED values compared with unrelated samples. If a database were to be generated for all varieties with all possible variations for each variety, an unknown sample could be confidently identified as that with which the ED would be lowest.

Discrimination between two samples depends upon how pure and uniformly well developed each one is. If there is contamination or developmental variation, the ED for replicates would be larger, which may lead to overlapping and difficulty in distinguishing one from the other. Although these varieties are unlikely to be cultivated in an environment as adverse as at sowing date 3 January, this was included in the analysis to cover maximum possibilities. In practice, commercial cultivars grown in optimum conditions would be expected to show lower intra-varietal ED, thus facilitating identification.

There is considerable value in characters that remain invariable and can be used to distinguish varieties reliably. These are potentially useful in Distinctness, Uniformity and Stability (DUS) testing. However, such characters, which show uniformity and stability, often may not be useful for distinctness. In computing the Euclidean distance, many parameters are included and they contribute to identification of a given variety. Some of the parameters used in the present paper were more stable, and the Euclidean distance (ED2) based on these yielded better results in classification of an unknown sample (Table 4). These characters may also be more useful for DUS purposes.

The technique collects data on a grain-to-grain basis. Hence, it will be possible to identify individual grains with deviant properties. The potential of the method can be improved by increasing the number of parameters, which will be helpful in using it in more complex situations.

In conclusion, though the environmental variation increased ED within varieties, in most cases it was lower than ED between varieties. Thus, using the criterion of lowest ED, it may be possible to identify a variety. Some parameters were more stable than others across the sowing dates. An ED based on this limited set of parameters (ED2) showed improved discrimination between varieties. Computer-based image analysis thus has the potential to correctly identify samples from different environments in the normal range of variation.

We are thankful to Ms Radhika Vishwanathan for collecting images and Ms R. Rastogi for her support and encouragement during the course of the work.

REFERENCES

BHAGWAT, S. G., SAINIS, J. K., SHOUCHE, S. P. & RASTOGI, R. (2003). Grain morphometric characterization of genetically related wheat selections using image analysis. *Cereal Research Communications* **31**, 205–212.

GONZALEZ, R. C. & WOODS, R. E. (1993). Representation and description. In *Digital Image Processing* (Eds R. C. Gonzalez, & R. E. Woods), pp. 643–675. New York: Addison-Wesley.

HU, M. K. (1962). Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory* **IT 8**, 179–187.

JAIN, A. K. (1995). *Fundamentals of Digital Image Processing* (Ed. A. K. Jain), pp. 377–394. New Delhi: Prentice Hall of India Private Limited.

KEEFE, P. D. (1999). Measurement of linseed (*Linum usitatissimum* L.) seed characters for distinctness, uniformity and stability testing using image analysis. *Plant Varieties and Seeds* **12**, 79–90.

KEEFE, P. D. & DRAPER, S. R. (1986). The measurement of new characters for cultivar identification in wheat using machine vision. *Seed Science and Technology* **14**, 715–724.

PIETRZAK, L. N. & FULCHER, R. G. (1995). Polymorphism of oat kernel size and shape in several Canadian oat cultivars. *Canadian Journal of Plant Science* **75**, 105–109.

SAKAI, N., YONEKAWA, S., MATSUZAKI, A. & MORISHIMA, H. (1996). Two-dimensional image analysis of the shape of rice and its application to separating varieties. *Journal of Food Engineering* **27**, 397–407.

SAPIRSTEIN, H. D. (1995). Varietal identification by digital image analysis. In *Identification of Food-Grain Varieties* (Ed. C. W. Wrigley), pp. 91–130. St. Paul. MN: American Association of Cereal Chemists Inc.

SHOUCHE, S. P., RASTOGI, R., BHAGWAT, S. G. & SAINIS, J. K. (2001). Shape analysis of grain of Indian wheat varieties. *Computers and Electronics in Agriculture* **33**, 55–76.

SYMONS, S. J. & FULCHER, R. G. (1988). Determination of wheat kernel morphological variation by digital image analysis. I Variation in eastern Canadian milling quality wheats. *Journal of Cereal Science* **8**, 211–218.

VISEN, N. S., PALIWALL, J., JAYAS, D. S. & WHITE, N. D. G. (2004). Image analysis of bulk grain samples using neural networks. *Canadian Biosystems Engineering* **46**, 7.11–7.15.