J. Cotton Res. Dev. **37**(1), 16-24 (January, 2023) **ISSN No. 0972-8619**

Studies on the diversity of extant upland cotton (*Gossypium hirsutum* L.) genotypes based on seed image and fiber characteristics

K. RATHINAVEL*, H. KAVITHA AND C. PRIYADHARSHINI
Central Institute for Cotton Research, Regional Station, Coimbatore-641003
**Email : krathinavelus@gmail.com*

Abstract : An experiment conducted to explore diversity among the extant genotypes by analyzing the geometry of seeds using image analysis technique, fiber quality, and yield traits in cotton. The characteristics like area, perimeter, length, width, radial variance, CMRV, width X, height Y, fiber length, fiber strength, lint weight, and boll weight showed a significant positive association with seed index. Hence, these traits were used directly for the establishment of distinctiveness. The studies on PCA revealed seven components explained Eigenvalue >1, which contributes 82.10 per cent variability, and the remaining elements contribute only 17.90 per cent of the variability. The first component is attributed mainly by the traits like seed length, width, height, and seed index, which may be useful for the development of yield enhancement in cotton. The genotypes, namely AK 32, Kanchana, Sruthi, AK 23B, HLS 329, VICH5, AK 32B, RHC 003, RHC 004, and PKV Rajat shown distinct from other genotypes and are highly diversified. A set of 87 genotypes grouped into five clusters. Among the clusters, cluster III was the largest occupied by 33 genotypes. Thus the highly diversified genotypes are very much influenced by the above traits, and hence, these genotypes would easily be distinguishable from the rest. The analysis of seed data from the image capturing device was found useful for varietal identification, characterization, and classification of existing cotton genotypes.

Keywords : Clustering, correlation, Cotton seed, distinctiveness, image analysis, PCA

Cotton is the most important commercial crops grown worldwide. In India, varieties from all the four cultivated cotton species are grown. With the advent of the development of a large number of cotton varieties under public, private, and public private partnership model, the protection of intellectual rights as envisaged since 2001, and the system is in place. Intellectual property rights on varieties are issued based on the established Distinct, uniform, and stable morphological traits. The morphological descriptors were employed in the grow-out test to determine the genetic purity of these crop varieties for seed certification. These morphological descriptors have the advantage of traditional significance and are immediately accessible and do not require specialized equipment for measurement. However, this approach demands field assessment, which is laborious, time-consuming, and very much depends on the degree of experience of the experimenter. Most of the descriptors require subjective decisions on minor distinctness upon the interaction of cultivars with the environment. In order to minimize or eliminate the influence of the environment, seeds of these genotypes or a variety can be observed more effectively with the computerized image analysis system to capture and process morphological information for varietal identification.

Characterization of varieties through image analysis technique was successfully developed by studying the geometry of seeds *viz.*, area, perimeter, length, width, elongation, centroid, circularity, average radial, radial variance, CMRV and shape factor. In this process, image recognition, image based extraction of shape characteristics are significant. Shape characteristics can often provide critical information for seed identification. In particular, the graininess characteristics of seeds can be essential features for seed variety identification. Varietal identification and characterization using image analysis well documented for various crops like sorghum, lucerne, castor, sesame, phaseolus, mustard, and oat. The image analysis technique is one such system which offers the prospect that researches may be able to study the seed surface features more closely and also useful for classification of cotton genotypes. Image analysis enables to process large amounts of material for examination in a relatively short time (Barn and Baitley, 1987). The image analysis system has potential use in a wide range of tasks such as determination of the cultivar identity of seed lots and testing of the distinctness of new cultivars for the award of breeder's rights and cultivar registration (Keefe and Draper, 1986). Kapadia *et al.* 2017 reported that DUS characters play a crucial role for proper implementation of PVP and PBR programs in India, Image analysis of seed has been potentialy used for determination of the variety identification of seed lots and for testing distinctness new varieties for the grant of breeders' right and variety protection. Measurements based on Images are faster and easy to achieve the desired objectives and thus can provide data correlating with the genetic properties of germination and growth performance of seedling.

Dehghan-Shoar *et al.* (1998) discriminated several varieties of Lucerne by capturing images of seed morphological characters and recommended his method for seed certification. Knowledge of the association of different traits is essential for cotton improvement that facilitates in discrimination of plants with desirable characteristics. The diversity study is vital for evolving the genotypes that meet the future demands under IPR. The rationale behind the present study undertaken was to study the geometry of seeds using an image analysis technique and to find out the diversity among the genotypes.

MATERIALS AND METHODS

In total, 87 extant upland cotton genotypes used in this study. Seeds were sown in randomized block design in the field for the measurements of morphological traits as per the guidelines. At harvest, 10 randomly selected bolls picked, weighed, and the average expressed in g/boll. After ginning, the lint was weighed and expressed in g/boll. Seed index (g) computed by taking the average weight of 100 seeds in eight replications. The fiber traits such as fiber length (2.5% span length mm), fiber strength (g/tex), fiber fineness (Mic.), fiber uniformity (%) measured through High Volume Instrument (Premier) under ICC mode. The seeds were defuzed, cleaned, and seeds with complete shapes used for capturing the images. For image characteristics, images of 30 seeds were taken in every replication of individual genotypes using Delta-T© (Delta- Instrument Device- Cambridge, UK) image analysis system by running custom-written software 'winDIAS' for measuring the geometry of seeds. (Nick Webb and Dick Jekins, 2000). In this process, thirty seeds placed above the lighting hood in such a way that the embryo of seed facing Image Analysis System and the longitudinal axis of the seed running parallel to the surface of the camera lens. Seeds viewed with a video camera (DSP surveillance color CCD camera (CVS 200/3300) using transmitted light so that the win DIAS software recorded a binary image of the silhouette of the seed.

Data measurement

Before going for actual measurement, calibration done by placing a transparent plastic ruler on the lighting hood illuminated from below. A ruler was aligned diagonally across the field of view, and the focus was adjusted to sharpen the image. Again aperture adjustment was made until optimum color and contrast achieved. Input length is given in centimeter.

Descriptors like area, perimeter, length, and width, from the menu object meter, were selected. After setting, the image was grabbed using image grabber, and the color threshold was done until the entire area was highlighted. For logging the data, the measurement button was pressed; the entire data was extracted each time by clicking entire objects. Data were viewed from the review module, and the data of each parameter were summed up for average value in the win DIAS. The seed parameters measured were as below area, perimeter, length: distance between two points marked on screen using the mouse (or) diameter of the smallest circumscribed circle that will fit around an object, width: Width is measured in horizontal X-axis, *Elongation*: It is the ratio of the length and width

$$E = \frac{W}{l}$$

Centroid: It is the most central point or centre of gravity of the object (measured from the top left-hand corner of the screen). Win DIAS calculates it using an algorithm based on the following equation.

$$\bar{x} = \frac{1/6 * \sum[(X2\,i+1 + X2\,i+1 * Xi + X2\,i)\,(Yi+1 - Yi)]}{\frac{1}{2}(Yi+1+Y)\,(Xi - X\,i+1)}$$

$$\bar{y} = \frac{\frac{1}{6} * \sum[(Y2\,i+1 + Y2\,i+1 * Yi + Y2\,i)\,(Xi+1 - Yi)]}{\frac{1}{2}(Xi+1+X)\,(Yi - Y\,i+1)}$$

Circularity: is the square root of the ratio of the actual area of the object to the area of a circle with the same circumscribed diameter.

$$C = \sqrt{\frac{A}{AP}}$$

Where, A is the actual area of the object and Ap is the area of a circle with a diameter equal to the circumscribed diameter of length of the object.

Average radial: It is an average of all distances measured from the centroid to each perimeter point.

$$\bar{R} = \frac{\sum Rj}{n}$$

Where, Rj is the j^{th} radius measured from the centroid to the jth perimeter point and n is the total number of perimeter points.

Radial variance: It is square of the standard deviation of all distances measured from the centroid to each perimeter point.

$$RV = \frac{\sum Rj - \bar{R})2}{n-1}$$

CMRV: It is the correlation of the Average Radial and Radial Variance.

$$CMRV = \frac{RV * 100}{\bar{R}}$$

Shape factor (S factor): Shape factor is the ratio of the actual perimeter to that of a circle with the same area.

$$S = \frac{P}{PC}$$

Where P is the perimeter of the object and Pc is the perimeter of a circle with the same area as the object. Pc is calculated as follows.

$$(Pc = 2\,(\pi * A)^{0.5}$$

Where, A is the actual area of the object.

Pearson correlation coefficient was worked out for 22 traits and correlation matrix was prepared for comparing different traits. Principal component analyses (PCA) was executed to find out the relative importance of different traits in capturing the genetic variation and also clustering using PAST 3 software (Hammer *et al.* 2001).

RESULTS AND DISCUSSION

The variation among the genotypes recorded through descriptive statistics is presented in Table 1. The cancroid Y showed the most significant variation among the genotypes, which ranged from 22.43 to 197.96, followed by centroid X from 29.21 to 230.05. The trait-like fiber maturity showed the least variation and ranged from 0.65 to 0.89.

The traits like area, perimeter, length, width, radial variance, CMRV, width X, height Y,

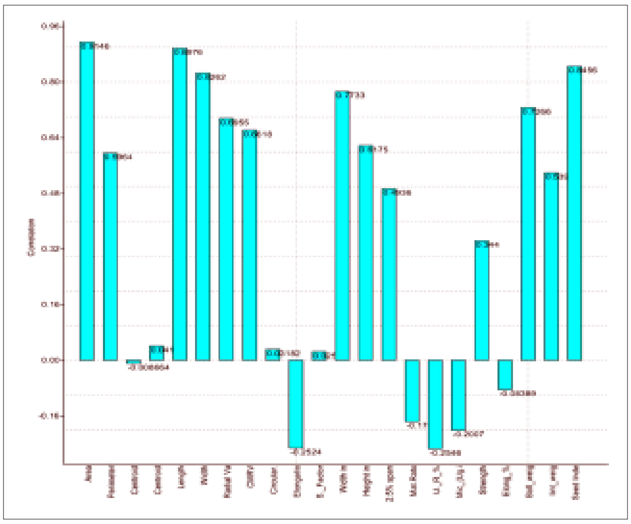

Fig. 1. PCA factor loadings for extant cotton genotypes

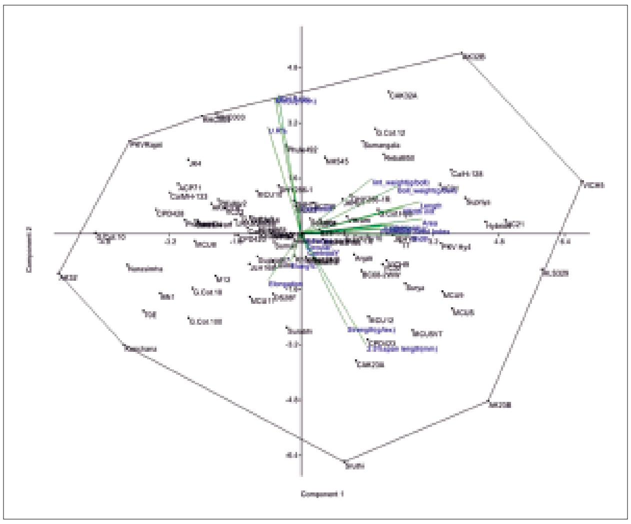

Fig. 2. Scatter plot for extant cotton genotypes using PC components 1 and 2

fiber length, fiber strength, lint weight, and boll weight showed a significant positive association with seed index (Table 2). Therefore, these traits may be directly concentrated while the selection of genotypes for cotton improvement, and they may establish distinctiveness among genotypes. In relation to inter correlation, the trait like area had a significant positive inter correlation with perimeter, length, width, radial variance, CMRV, width X, height Y, fiber length, fiber strength, boll weight, and lint weight. The perimeter had a significant inter correlation with length, width, radial variance, CMRV, S factor, width X, boll weight, and lint weight. The seed length had a significant positive inter correlation with width, radial variance, CMRV, width X, height Y, fiber length, boll weight, and lint weight. The seed width had a positive inter correlation with radial variance, CMRV, elongation, width X, height Y, fiber length, fiber strength, boll weight, and lint weight. The radial variance with CMRV, S factor, width X, height Y, fiber length, boll weight, and lint weight. CMRV with width X, height Y, fiber length, boll weight, and lint weight; circularity with elongation; width X and height Y with boll weight and lint weight; Height Y with fiber length; fiber length with fiber strength and boll weight; fiber maturity with fiber uniformity, fiber fineness and lint weight; fiber uniformity with fiber fineness and boll weight with lint weight. The trait, fiber fineness showed significant positive inter-correlation with lint weight; this may help for further improvement of yield and quality traits.

Li *et al.* (2012) extracted fourteen shape characteristic parameters of cottonseeds *viz.*, the area, perimeter, NCI ratio, circular degree, center of gravity *X*, center of gravity *Y*, primary diameter, short diameter, second moment *X*(*Mx*2), second-moment *Y*(*My*2), second moment *XY* (*Mxy*), major axis of oval, short axis of oval, and shape coefficient of oval using image analysis technique for the varietal identification

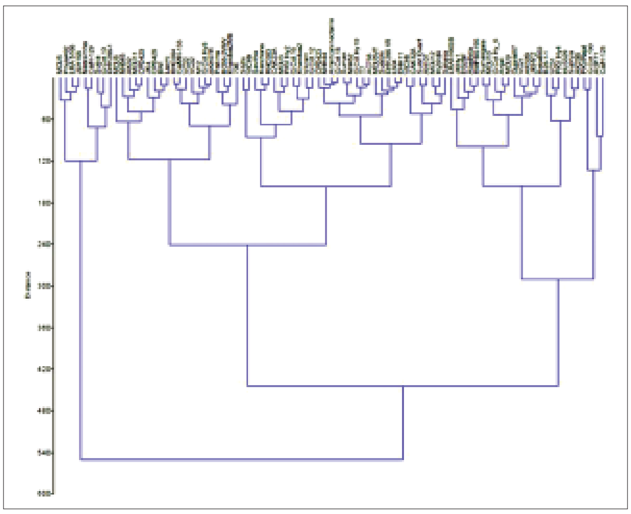

Fig. 3. Dendrogram of extant cotton genotypes by Ward's Minimum Variance

Table 1. Variability parameters of seed characters, fiber quality and yield component traits among extant cotton genotypes

Characters	Mean	Minimum	Maximum	Range	Standard Deviation
Area	**54.34**	41.18	68.75	27.57	5.83
Perimeter	**44.39**	37.03	60.04	23.01	3.83
Cent. X	**65.54**	29.21	230.05	200.84	32.04
Cent. Y	**76.92**	22.43	197.96	175.53	67.40
Length	**11.74**	9.72	13.52	3.80	0.70
Width	**6.72**	6.07	7.60	1.53	0.34
Rad. Var.	**26.54**	17.20	37.67	20.47	4.55
CMRV	**157.01**	92.93	241.40	148.47	32.68
Circularity	**0.72**	0.66	1.06	0.40	0.04
Elongation	**0.58**	0.50	0.65	0.15	0.03
S factor	**1.71**	1.42	2.18	0.76	0.14
Width X	**10.09**	8.58	11.77	3.19	0.73
Height Y	**8.44**	7.02	9.67	2.65	0.57
Fibre length (2.5% span length)	**27.96**	23.00	35.73	12.73	2.55
Maturity (%)	**0.78**	0.65	0.89	0.24	0.04
Uniformity (%)	**47.47**	41.50	51.17	9.67	1.86
Fibre fineness (micronaire)	**4.18**	2.97	5.43	2.46	0.42
Fibre strength (g/tex)	**21.86**	18.83	28.33	9.50	1.84
Elongation (%)	**6.03**	4.90	6.67	1.77	0.30
Boll weight (g/boll)	**2.75**	2.10	3.74	1.64	0.29
Lint weight (g/boll)	**1.69**	1.10	2.39	1.29	0.24
Seed Index	**10.55**	8.53	13.57	5.04	0.99

of delinted cotton seeds using the color and shape parameters based on BP neural network.

Shouche *et al.*, (2001) extracted geometric features and shapes related feature analysis of the cereal grains. Geometry related features including area, perimeter, major and minor axis lengths, compactness, axis ratio, shape factor 1 to 5, and spread and slenderness were measured from the binary images. Standard or raw, central, normalized central, and invariant moments were computed from the digital images of each grain and mean, standard deviation (S.D.), and standard error (S.E.) calculated.

Thangavel *et al.* (2003) studied the varietal characterization of sorghum seeds through the image analysis technique. The accurate, meaningful data and the rapidity of measurement were done using this technique.

Shahin and Symons, 2003 studied the lentil type identification using the change of color, uniformity, and discoloration due to developmental variation using machine vision.

In the principal component analysis, the most commonly used criterion for resolving the number of components is the Eigenvalue one criterion. Seven principal components depicted the Eigenvalue more than 1, which contributes 82.10 per cent variability, and the remaining components contribute only 17.90 per cent of the variability. In the first component, area (0.915), length (0.898), seed index (0.846), width (0.826), width X (0.773), boll weight (0.727), radial variance (0.696), CMRV (0.662) and height Y (0.618) contributes totally 30.398 per cent of variability. For the second component, maturity (0.911), fibre fineness (0.894) and fibre uniformity (0.702) contributes 15.96 per cent of variability (Table 3). The component one is mainly attributed to the seed length, width, height, and seed index, which may be useful for the development of yield improvement in cotton.

The traits like area, perimeter, centroid Y, length, width, radial variance, CMRV, circularity, S factor, width in X, height in Y, 2.5 per cent span length, fiber strength, boll weight, lint weight, and seed index were the positive PCA loadings.

Table 2. Correlation coefficient between seed parameters, fibre quality and yield contributing traits in extant cotton genotypes

Characters	Area	Perimeter	CentroidX	CentroidY	Length	Width	Radial	CMRV	Circularity	Elongat.	S factor	WidthX	HeightY
Area	1	0.356**	0.038	0.012	0.915**	0.879**	0.511**	0.615**	0.146	-0.209	-0.276**	0.820**	0.612**
Perimeter		1	0.045	0.042	0.479**	0.464**	0.647**	0.272*	-0.275**	-0.106	0.798**	0.503**	0.175
CentroidX			1	0.056	0.088	-0.035	0.049	0.061	0.003	-0.127	0.024	0.085	-0.018
CentroidY				1	0.028	0.023	0.118	-0.077	-0.074	-0.012	0.044	0.037	-0.019
Length					1	0.691**	0.607**	0.616**	-0.086	-0.541**	-0.096	0.801**	0.666**
Width						1	0.449**	0.489**	0.163	0.226*	-0.088	0.742**	0.435**
Radial							1	0.480**	-0.196	-0.310**	0.333**	0.487**	0.402**
CMRV								1	0.080	-0.243*	-0.119	0.519**	0.418**
Circularity									1	0.340**	-0.369**	0.059	-0.116
Elongat.										1	0.027	-0.210	-0.399**
S factor											1	-0.003	-0.221*
WidthX												1	0.131
HeightY													1

** Significant at 1% * Significant at 5%

Elongat. – Elongation; S factor – Shape factor; FL – Fibre length; FM – Fibre maturity; UR – Uniformity ratio; FF – Fibre fineness; FS- Fibre strength; FE – Fibre elongation; BW – Boll weight; LW- Lint weight; SI- Seed Index

Contd..,

Table 2. Correlation coefficient between seed parameters, fibre quality and yield contributing traits in extant cotton genotypes

Characters	FL	FM	UR	FF	FS	FE	BW	LW	SI
Area	0.341**	-0.074	-0.113	-0.095	0.242*	-0.030	0.581**	0.382**	0.796**
Perimeter	0.201	-0.099	-0.173	-0.117	0.107	-0.109	0.430**	0.351**	0.420**
Centroid X	-0.113	0.067	0.095	0.072	-0.212*	0.044	-0.023	0.048	-0.179
Centroid Y	0.095	-0.063	-0.045	-0.075	0.055	0.007	-0.021	-0.049	0.051
Length	0.247*	-0.005	-0.070	-0.023	0.150	0.013	0.555**	0.377**	0.730**
Width	0.348**	-0.119	-0.159	-0.148	0.245*	-0.119	0.557**	0.387**	0.712**
Radial	0.254*	-0.114	-0.106	-0.113	0.180	-0.015	0.448**	0.364**	0.448**
CMRV	0.234*	-0.101	-0.147	-0.110	0.093	-0.014	0.375**	0.271*	0.457**
Circularity	0.037	-0.033	0.010	-0.060	0.090	-0.078	0.145	0.114	0.147
Elongat.	0.056	-0.156	-0.087	-0.169	0.076	-0.151	-0.114	-0.074	-0.162
S factor	-0.010	-0.045	-0.107	-0.050	-0.044	-0.098	0.067	0.111	-0.072
Width X	0.156	-0.019	-0.041	-0.049	0.118	-0.072	0.480**	0.328**	0.620**
Height Y	0.312**	-0.034	-0.133	-0.032	0.171	0.073	0.394**	0.260*	0.547**
FL	1	-0.630**	-0.728**	-0.617**	0.683**	0.049	0.217*	0.084	0.444**
FM		1	0.602**	0.978**	-0.549**	-0.187	0.166	0.215*	-0.032
UR			1	0.596**	-0.200	-0.417**	-0.064	-0.029	-0.116
FF				1	-0.541**	-0.147	0.124	0.177	-0.062
FS					1	-0.364**	0.070	-0.071	0.388**
FE						1	-0.167	-0.116	-0.217*
BW							1	0.950**	0.659**
LW								1	0.392**
SI									1

** Significant at 1% * Significant at 5%

Among these traits, area (0.9146) recorded the highest factor loadings followed by seed length (0.898), seed index (0.846), and seed width (0.826). So these traits might be useful for the crop improvement program. The traits like centroid X, elongation, fiber maturity, fiber uniformity ratio, fiber fineness, and fiber elongation were the negative PCA loadings (Fig.1). It was reported that Draper and Travis (1985) separated seeds of 49 different crops and weed species based on their shape in combination with length factor using an image processing system. Myers and Edsall (1989) reported that the possibility of discriminating Australian wheat varieties using machine vision procedures.

The scatter plot between the principal

Table 3. Principal components of seed, fibre characters and yield component in extant cotton genotypes

Characters	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7
Area	0.915	0.097	-0.282	-0.116	-0.012	0.174	-0.033
Perimeter	0.596	0.037	0.704	0.260	0.027	0.110	-0.1186
Centroid X	-0.009	0.189	0.123	-0.215	0.216	0.493	0.2648
Centroid Y	0.042	-0.098	0.137	-0.025	-0.146	0.205	0.8791
Length	0.898	0.211	-0.016	-0.308	-0.112	0.098	-0.036
Width	0.826	-0.006	-0.204	0.227	0.104	0.241	-0.064
Radial Variance	0.696	0.060	0.425	-0.060	-0.076	0.038	0.027
CMRV	0.662	0.058	-0.107	-0.211	-0.053	0.146	-0.200
Circularity	0.032	-0.062	-0.648	0.333	0.307	0.149	0.047
Elongation	-0.252	-0.302	-0.226	0.674	0.286	0.190	-0.033
S factor	0.025	-0.021	0.902	0.348	0.032	0.007	-0.094
Width X	0.773	0.173	-0.016	0.033	0.028	0.450	-0.107
Height Y	0.618	0.059	-0.137	-0.411	-0.153	-0.379	0.051
Fibre length	0.494	-0.727	-0.036	0.016	-0.011	-0.225	0.124
Maturity (%)	-0.176	0.911	-0.064	0.070	-0.027	-0.117	0.047
Uniformity (%)	-0.255	0.702	-0.141	0.165	-0.436	0.184	-0.026
Fibre fineness (micronaire)	-0.201	0.894	-0.053	0.030	-0.036	-0.125	0.035
Fibre strength (g/tex)	0.344	-0.601	-0.147	0.258	-0.523	-0.102	0.032
Elongation (%)	-0.084	-0.186	0.092	-0.627	0.590	-0.027	-0.016
Boll weight (g/boll)	0.727	0.311	-0.024	0.307	0.281	-0.351	0.163
Lint weight (g/boll)	0.539	0.363	0.049	0.315	0.413	-0.381	0.186
Seed Index	0.846	0.041	-0.190	0.139	-0.171	-0.123	0.032
Eigen values	6.687	3.511	2.244	1.883	1.394	1.231	1.013
Variability (%)	30.398	15.961	10.200	8.560	6.338	5.597	4.606
Cumulative variability (%)	30.397	46.329	56.618	65.277	71.772	77.492	82.098

Table 4. Clustering of extant cotton genotypes for seed, fibre traits and yield components

Clusters	Groups	No. of genotypes	Name of the genotypes
I	1	4	MCU5, Phule 492, LRA 5166, JLH 168
	2	5	Narasimha, Ca-H-129, Surya, G.cot. 10, DHY 286-1
II	1	9	MCU8, NH 545, AK32,MCU11, CPD 420, JK4, CPD 428, BN1, Laxmi
	2	11	RHC 004, Ca/MH-133, VC 22, VC 32, M12, G.cot. Hy. 8, AKH07R, Pratima, BC 68-2WW, Khandwa 2 MB, 70E
III	1	12	Sruthi, VICH 9, Savitha, Kanchana, RHC 003, CAK 23A, MCU 3, PKV Hy 2, G.Cot. 16, Khandwa 2, Vikram, G.Cot.12
	2	14	CPD 423, NHH 44, American Nectariless, G. Cot. 18, Suman, NH 452, G.Cot. Hy. 10, T7, AC 738, MCU5VT, Surabhi, DHY 286-1R, L 604, DHH 11
	3	7	VICH 5, CAK 32A, Swart Dwarf, AH107, MCU12, Hybrid 4, Sahana
IV	1	5	LRA 5166 SB, JKHy1, DS 28F, Abaditha, RMPBS 155
	2	10	Sumangala, MCU 10, G. Cot. Hy6, Anjali, MCU 9, Gujarat 67, VC 21, AK 32B, JK Hy2, Reba B50
	3	6	B 28-1-1, VC 31, PKV Hy4, HLS 329, Supriya, AK 23B
V	1	2	PKV Rajat, G.Cot. 100
	2	2	ACP 71, Ca/H-128

components 1 and 2, which explained the arrangement of grouped genotypes in a factor plane (Fig.2). The distance of the genotypes from the base of the spread-out plot to the outer side of the polygon showed the longest distance from others, which generates greater genetic diversity. The genotypes like AK 32, Kanchana, Sruthi, AK 23B, HLS 329, VICH5, AK 32B, RHC 003, RHC 004, and PKV Rajat were found at the vortex of the polygon and showed distinct from other genotypes and are highly diversified. Malik et al., 2013 also reported that the genotypes placed at the vortex of the polygon showing the longest distance from the base depicted the greatest genetic variability, and closer genotypes displayed lower genetic variation due to narrow genetic base.

Five clusters were found with respect to the measurement of the geometry of seed by image analysis technique, fiber quality, and yield traits in cotton (Fig. 3). Among the clusters, cluster III is the largest having 33 genotypes, clustered into three groups consists of 14, 12, and 7 genotypes followed by cluster IV, which consists of 21 genotypes in three groups, each containing 10, 6, and 5 genotypes respectively. The cluster V is the smallest containing four genotypes, two genotypes in each group (Group I - PKV Rajat, G. Cot. 100 and Group II - ACP 71, Ca/H-128) (Table 4). The divergent genotypes can be used for the crop improvement program.

CONCLUSION

Machine vision has been utilized for cultivar description, characterization, and identification of varieties using seeds and plant parts (Draper and Travis, 1984). Finch-Savage (1986) measured the seedling vigor of some vegetables using an image analysis system and reported a positive relationship of the results with seed vigor and field performance. The above study concludes that the traits like area, perimeter, length, width, radial variance, CMRV, width X, height Y, fiber length, fiber strength, lint weight, and boll weight had a significant positive association with seed index. The seven principal component analysis showed Eigenvalue >1, which contributes 82.10 per cent variability. The first component is attributed mainly by the traits like seed length, width, height, and seed index, which will be useful for the development of yield improvement in cotton. In the scatter diagram, the genotypes like AK 32, Kanchana, Sruthi, AK 23B, HLS 329, VICH 5, AK 32B, RHC 003, RHC 004, and PKV Rajat shown distinct from other genotypes and are highly diversified. The cluster analysis grouped the genotypes into five clusters with cluster III is the largest containing 33 genotypes. So these genotypes are very much useful for the crop improvement program. The image analysis system will be useful for varietal identification in cotton and also for identifying the admixture seeds in the seed lot and testing of the distinctness of new cultivars.

REFERENCES

Barn, B.R. and Baitley, L.G. 1987. A survey of endosperm starch granules in the genus *Hordeum*: a study using image analysis and numerical techniques. *Canadian Journal of Botany*, **65**: 1563-69.

Dehghan-Shoar, M., Hampton, J.G. and Haslett, S.J. 1998. Identification and discrimination among Lucerne (Medicago sativa L.) varieties using seed image analysis. Pl. *Varieties Seeds*, **11**: 107-27.

Draper, S.R. and Travis, T.J. 1984. Preliminary observations with a computer based system for analysis of the shape of seeds and vegetable structures. *J. Nati. Instit. Agricult.* Botany, **16**: 387-95.

Finch-Savage, W.E. 1986. A study of the relationship between seedling characters and rate of germination within a seed lot. *Ann. Appl. Biol.*, **198**: 441-44.

Hammer, O., Harper, D.A.T. and Ryan, P.D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electr.* **4**: 1–9

Kapadia, V., Sasidharan, N., Kalyanrao, P. 2017. Seed Image Analysis and Its Application in Seed Science Research. *Adv Biotech Micro.*, **7**: 1-3

Kneefe, P.D. and Draper, S.R. 1986. The measurement of new characters for cultivar identification in wheat using machine vision. *Seed Sci. Tech.* **14**: 715-24.

Li, J., Chen, B., Shao, L., Tian, X. and Kan, Z. 2012. "Variety identification of delinted cotton seeds based on BP neural network," *Trans. Chinese Soci. Agri. Engin.* **28** : 265-69.

Malik, W., Khan, A. and Sadia, B. 2013. In situ characterization of coloured cotton genotypes. *Australian J. Crop Sci.*, **7**: 299- 304.

Myers, D.G. and Edsall, K.J. 1989. The application of image processing techniques to the identification of Australian wheat varieties. *Pl. Varieties Seeds*, **2**: 109-16.

Nick, Webb and Jekins, Dick 2000. winDIAS user manual, Delta T. Devices Ltd., Cambridge, UK.

Shahin, M.A. and Symons, S.J. 2003. Lentil type identification using machine vision. *Canadian Bio-systems Engineering*, **45**: 3.5-3.11.

Shouche, S. P., Rastogi, R., Bhagwat, S. G. and Sainis, J.K. 2001. "Shape analysis of grains of Indian wheat varieties," *Computers Electronics* Agriculture, **33**: 55-76.

Thangavel, P. 2003. Characterisation of sorghum (*Sorghum bicolor* L.) Moench. Cultivars by traditional and modern approaches. *M.Sc.*(*Ag.*) *Thesis*, Tamil Nadu Agricultural University, Coimbatore-3.

Received for publication : October 14, 2022
Accepted for publication : November 20, 2022