Dimension fitting of wheat spikes in dense 3D point clouds based on the adaptive *k*-means algorithm with dynamic perspectives

Fuli Wang
School of Computer Science and Electronic Engineering
University of Essex
Colchester, United Kingdom
fuli.wang@essex.ac.uk

Vishwanathan Mohan
School of Computer Science and Electronic Engineering
University of Essex
Colchester, United Kingdom
vishwanathan.mohan@essex.ac.uk

Andrew Thompson
National Physical Laboratory
Teddington, United Kingdom
andrew.thompson@npl.co.uk

Richard Dudley
National Physical Laboratory
Teddington, United Kingdom
richard.dudley@npl.co.uk

Abstract—The use of dense 3D point clouds to obtain agricultural crop dimensions in the place of manual measurement is crucial for enabling high-throughput phenotyping. To achieve this goal, this paper proposes an adaptive *k*-means algorithm based on dynamic perspectives, which first performs segmentation in order to separate the wheat spikes. We also propose a method to fit the shape of each spike and measures the dimensions of each spike with the help of the Random Sample Consensus algorithm. The experimental results show that the proposed method can be applied in a complex environment where multiple wheat spikes are grown densely and that it can fit the size of most wheat spikes accurately.

Keywords—point clouds, wheat, clustering, segmentation, cuboid-fitting

I. INTRODUCTION

Image-based plant phenotyping is a rapidly emerging research area which has the potential to provide quantitative measurement of the structural and functional properties of plants for the development of new plant varieties. A variety of imaging methodologies have recently been explored for collecting data for quantitative studies of complex traits related to the growth, yield and adaptation to biotic or abiotic stress; see [1] and references therein. However, for wheat, the data collection of wheat spike size is still mainly carried out by means of manual sampling with a ruler [2]. Since the manual measurement is very time-consuming and labour-intensive, a switch to measuring the dimensions of wheat spikes using imaging methodologies is urgently needed to enable high-throughput phenotyping.

On the one hand, a few supervised neural networks have been proposed to fit some complex geometric primitives to 3D point clouds such as CAD model datasets of mechanical components [3, 4]. The maturity of wheat, on the other hand, can be judged only by fitting the spike size, which does not involve too many complex geometric primitives. Therefore, by using more classical clustering algorithms to segment the wheat, and then subsequently fitting to spikes, the process of training a supervised model can be omitted, and the fitting results can be obtained with less time cost. For example, [5] performs wheat spike segmentation using two different classical methods: voxel-based segmentation and mean shift segmentation. Besides, in [6], the density-based DBSCAN (Density Based Spatial Clustering of Applications with Noise) algorithm [7] has been developed for the task of segmentation, and then least-squares curve fitting used to obtain the size of the wheat spikes. Although the clustering algorithms such as DBSCAN, mean shift and *k*-means can be successful in some segmentation tasks, the segmentation task can be challenging for these algorithms in certain complex environments, such as when wheat crops are very dense.

To address the above concerns, the contributions of this paper can be summarized as follows:

1) A new *k*-means algorithm based on dynamic perspectives is proposed. The algorithm processes 3D point clouds according to different perspectives, thereby improving the segmentation performance.

2) The traditional *k*-means algorithm needs to determine the value of the parameter *k* before using it. The self-adaptive algorithm in this paper can update the appropriate *k* value iteratively according to the input point clouds.

3) The algorithm proposed in this paper also embeds the RANSAC (Random Sample Consensus algorithm) [8], which can fit the shape of the spike as a cuboid and output the dimensions after the segmentation.

II. CLUSTERING AND SEGMENTATION ALGORITHMS

A. Traditional k-means and DBSCAN

Although the *k*-means algorithm has become one of the top ten clustering algorithms because of its simplicity and efficiency [9], this algorithm needs to set the number of clusters, *k*, in advance. Similarly, DBSCAN is also non-adaptive and requires careful tuning of two parameters: the least number of neighbours and the neighbourhood radius.
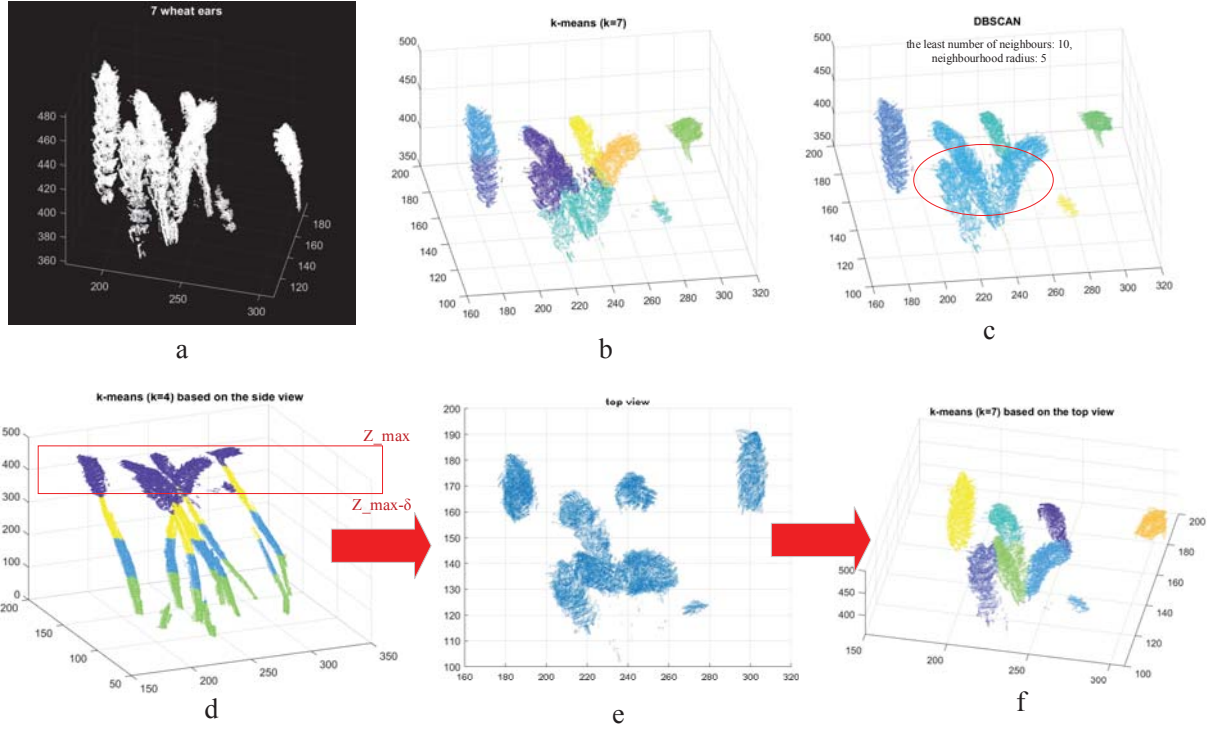
Fig. 1 (a) The initial point cloud of wheat spikes; (b) the segmentation result of *k*-means; (c) the segmentation result of DBSCAN; (d) a segmentation result of *k*-means based on side view; (e) the top view of spikes of wheat; (f) the segmentation result of *k*-means based on top view

These two algorithms are useful tools for image segmentation, but there are still some defects when dealing with some practical situations. As shown in a, b, c from Fig. 1, the 3D point clouds of several wheat spikes are input into the two algorithms respectively, and the running results are not ideal. In particular, in order to fit the shape of the spike, each wheat spike should be first segmented and then fitted. However, it can be seen from the results in Fig. 1 that *k*-means cannot distinguish the wheat spikes at all, while for DBSCAN it is impossible to segment dense wheat (the red circle part in Fig. 1c).

B. *The k-means algorithm based on dynamic perspectives*

As wheat is often observed to grow upwards, as shown in Fig. 1.a, when we observe these wheat crops from the side view, we can easily distinguish which part is the wheat spike and which part is the stem. Due to the overlap between the wheat, we cannot easily distinguish how many wheat crops there are. However, we can change our observation perspective: if we use the top view to observe again (Fig. 1.e), it is easy to distinguish how many wheat crops there are. This is the kind of deduction that would be common sense for human beings: by changing the perspective, we can get more information.

To improve the segmentation performance, this common sense is introduced into the *k*-means algorithm. Firstly, in the 3D coordinate system, all the points in the 3D point cloud are projected onto the XZ plane, and the 3D view is converted into a 2D side view. Similarly, all 3D points are projected onto the XY plane to obtain a two-dimensional top view. Specifically, given a cluster consisting of points $N_{n\times3}$, where n is the number of points, and 3 is the number of dimensions, we denote by $\{x_i\}$, $\{y_i\}$ and $\{z_i\}$ the x, y and z coordinates of the point i ($i \in n$). For

the side view, we transfer the $N_{n\times3}$ array into an $N_{n\times2}$ array, which only contains the two dimensions of $\{x_i\}$ and $\{z_i\}$. Then, we input the 2D points from the side view into the *k*-means, which outputs the labels of all points. Fig. 1.d can be obtained by using theses labels to mark all 3D points. Then the result of this procedure is the extraction of all 3D points that belong to spikes of wheat. Similarly, transferring the 3D points of spikes into the top view $N'_{n\times2}$, which only contains two dimensions of $\{x_i\}$ and $\{y_i\}$, we obtain the final result in Fig.1.f, which is better than the previous results.

More specifically, two algorithms are proposed in this paper. Firstly, Alg. 1 projects the 3D point cloud image of wheat onto the 2D side view, and then uses *k*-means to segment wheat spikes and stems. Therefore, in Alg.1, we can obtain all points that belong to the wheat spikes. Then we input the point clouds of spikes into Alg.2, which transfers the 3D points within spikes into a 2D top view image, and then uses the self-adaptive *k*-means to separate each spike and fit the shape of each spike one by one.

Therefore, this is a two-stage method. Alg.1 separates spikes from the wheat, and Alg.2 identifies the individual spikes and realizes the cuboid fitting of each spike.

Alg. 1 is described in detail as follows. This algorithm uses *k*-means to segment the whole point cloud based on the side view. Then, it defines a value space according to the highest point which has the max value of $\{z\}$. By extracting all the segments in the value space, the points belonging to spikes are obtained. Note that, since each segment has been determined by *k*-means, the space value does not have to be set too accurately; it is sufficient to ensure that δ is a small value.

145

Algorithm 1: Wheat spike segmentation based on *k*-means
Require: 3D_points // *wheats*
Initialize parameter of *k*-means (3 or 4) and δ;
Convert the 3D points into side view 2D points;
Use *k*-means for segmentation based on 2D_points;
Obtain the point with the highest Z coordinate value: Z_max;
Calculating a value space of Z coordinates:
[Z_max- δ, Z_max];
Extract all segments in the value space;
return all points within wheat spikes.

C. Cuboid fitting based on self-adaptive k-means

In order to obtain good results from Alg.1, it is sufficient to set the *k*-means parameters to be 3 or 4. However, if we need to perform further shape fitting for each spike one by one, we have to set the number of clusters in advance, that is, the number of spikes.

In Fig. 1.f, the result is pretty good by setting a correct number of clusters. However, when dealing with different scenarios, it is impossible to know in advance how many wheat spikes there are. Therefore, this paper proposes an adaptive algorithm to self-update the appropriate parameter values. First, it needs to calculate the initial parameter *k* according to the input number of points:

$$k = floor(\frac{n}{a})$$

where, *n* is the number of input points and *a* is a constant. In the images collected in this paper, a single wheat spike contains between 18000 ~ 30000 points, the parameter *a* can be set to 31000~35000, so a small initial value of *k* is calculated. $floor(.)$ is a function that guarantees that the output is rounded down.

Next, after obtaining the initial parameter *k*, the algorithm uses *k*-means to segment the 2D points of the top view and then calls the RANSAC to fit a cuboid to each segmentation. Since the initial *k* value is small, the fitting result is not accurate. As shown in Fig. 2, when the *k* value is small, some abnormal wheat size will be output (the fitting size of the purple part is significantly larger than that of normal wheat). Therefore, once the algorithm detects unreasonable results, the *k* value will be superimposed until a reasonable final result is outputted. The procedure is described in detail in Alg. 2.
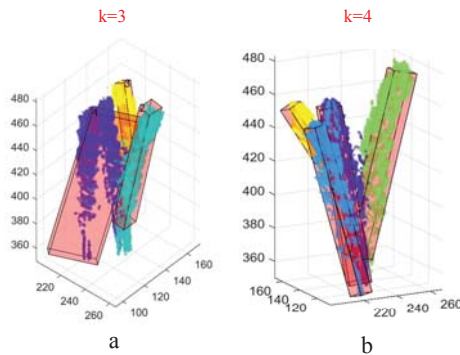

Fig. 2 (a) The shape fitting result with a small value of *k*; (b) the final shape fitting result

Note that the algorithm does not make any intrinsic change to the *k*-means algorithm, and instead, it requires several iterations of any existing implementation of *k*-means. Therefore, the algorithm can call any version of the *k*-means algorithm. In this paper, the lite *k*-means [10] is selected to carry out experiments on MATLAB 2019a.

Algorithm 2: Shape fitting based on the self-adaptive *k*-means
Require: 3D_points // *spikes*
Initialize parameter: *a*
Calculate the initial parameter according to equation (1)
if (*k*<1)
k=1;
end if
Convert the 3D point into top view 2D points
repeat
Use the *k*-means for segmentation based on top view;
Use RANSAC to fit each segment;
Evaluate the size of each segment;
if (there is an abnormal size)
k++;
break;
end if
until there is no abnormal size
return the updated shape model.

III. EXPERIMENTAL RESULTS

The method proposed in this paper can be applied to the scenario of multiple dense wheat crops. Therefore, this section describes experiments in which the method is applied to scenes involving multiple wheat crops.

Scenes featuring three, seven, and twelve wheat crops were tested with this method, and the results are shown in Fig. 3. In order to investigate how well the proposed method performs cuboid fitting using RANSAC [11], some shape fitting results are recorded in Table 1. The iterations of RANSAC was set to 3000.

In Fig. 3, the images in the left column are the original images, and the right column is the final output results. As described in Section II, the original image is processed by Alg. 1 after which only the points within wheat spikes are retained, and then Alg. 2 performs cuboid fitting for each spike one by one.

Besides, we also used the 12 wheat spikes point cloud to test the DBSCAN and traditional *k*-means, and the results are shown in Fig. 4. As is shown, the DBSCAN algorithm can only roughly divide the 12 spikes into 4 segments, while our method can output 11 segments. In other words, DBSCAN identifies that there are only 4 spikes, which is quite different from the actual situation. Meanwhile, in the traditional *k*-means algorithm, even if we have set the number of clusters to 12 (*k*=12), the output result is still not satisfactory. This is because the lack of perspective-changing makes *k*-means unable to output what we want. Compared with these results, it is obvious to see that the feasibility of our method and the shape fitting results are reasonably good.

Admittedly, it also can be seen that the fitting results of 12 wheat spikes are not all accurate in Fig. 3. Sometimes the method can get very good results, sometimes it cannot. This is

because *k*-means is a random algorithm and it cannot guarantee that the result of each run is always accurate. Due to the uncertainty of *k*-means, the method cannot guarantee that the dimensions of each spike are always accurate, but it can output the estimated dimensions of most of wheat spikes.

IV. CONCLUSIONS

We have proposed an adaptive *k*-means algorithm based on dynamic perspectives, which can deal with the complex environment where multiple wheat spikes are grown densely. The results demonstrate clearly a certain level of robustness of our algorithm to identify and measure the dimensions of wheat crops in the scenario of multiple spikes.
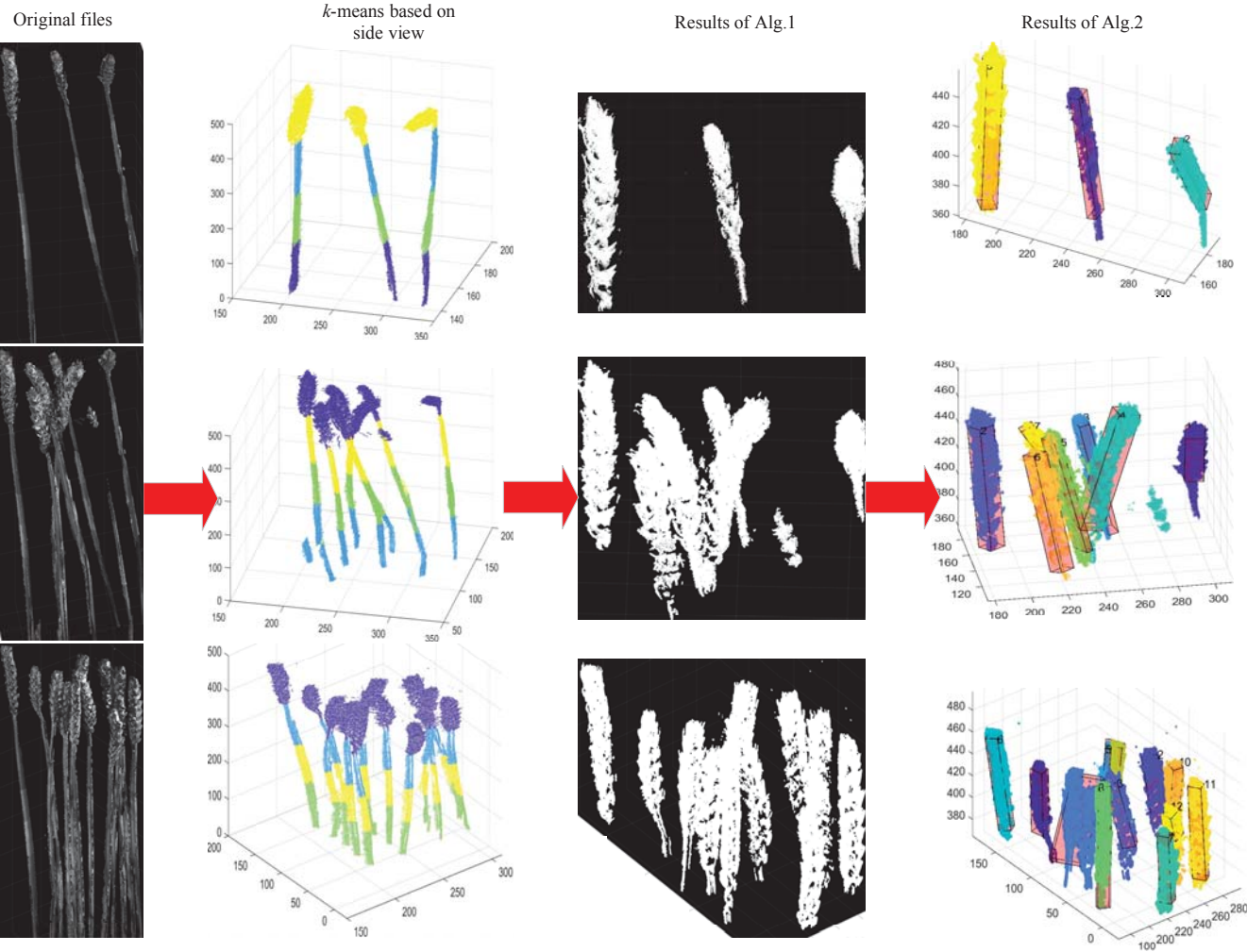

Fig. 3 The results of the proposed method

Table. 1 Estimation for crop height, spike length and width (height*length*width)/ cm

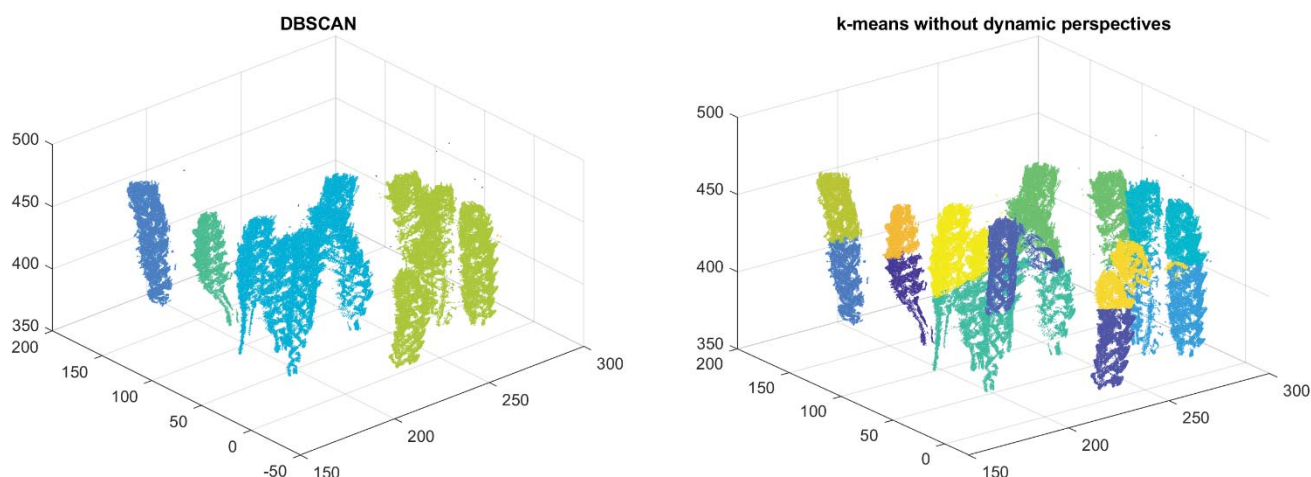| Filename	Ear number	RANSAC results
3 ear straight	1	85.35*12.03*5.61
	2	50.05*9.13*8.94
	3	95.06*11.58*9.69

147

Fig. 4 The results of DBSCAN and traditional *k*-means

REFERENCES

[1] L. Li, Q. Zhang, and D. Huang, "A Review of Imaging Techniques for Plant Phenotyping," *Sensors*, vol. 14, pp. 20078-20111, October 2014.

[2] A. Torres and P. Julian, *Crop morphological traits*, CIMMYT, 2012, ch. 19.

[3] H. Su et al., "SPLATNet: Sparse Lattice Networks for Point Cloud Processing," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, 2018, pp. 2530-2539.

[4] L. Li, M. Sung, A. Dubrovina, L. Yi and L. J. Guibas, "Supervised Fitting of Geometric Primitives to 3D Point Clouds," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019, pp. 2647-2655.

[5] K. Velumani, S.O. Elberink, M.Y. Yang M Y, "Wheat ear detection in plots by segmenting mobile laser scanner data," in *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, Wuhan, China: 2017, pp. 18–22.

[6] A. Thompson, V. Livina, P. Harris, I. Mohamed and R. Dudley, "Model-based algorithms for phenotyping from 3D imaging of dense wheat crops," in *2019 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor)*, Portici, Italy, 2019, pp. 95-99.

[7] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Knowledge Discovery in Databases (KDD)*, vol. 96, no. 34, 1996, pp. 226–231.

[8] R. Schnabel, R. Wahl and R. Klein, "Efficient RANSAC for point‐cloud shape detection," *Computer graphics forum*. Vol. 26, pp. 214-226, Oxford, UK: Blackwell Publishing Ltd, 2007.

[9] X. Wu, V. Kumar, J.R. Quinlan, et al, "Top 10 algorithms in data mining," *Knowledge and information systems*, Vol. 14, pp. 1-37, 2008.

[10] D. Cai, "Litekmeans: the fastest matlab implementation of kmeans," http://www.zjucadcg.cn/dengcai/Data/Clustering.html, 2011.

[11] U. Mehmood, "Cuboid fit (RANSAC)," https://github.com/usamamehmood3/Cuboid-Fitting-RanSAC, GitHub.