SUPPLEMENTARY MATERIAL

Zanthoxylum infructescence detection based on adaptive density clustering

Diwei Wu,[1,2] Shaohua Zeng,[1,2] Shuai Wang,[3] Yanan Chen,[4] Yidan Xu[5]

[1]College of Computer & Information Science, Chongqing Normal University; [2]Chongqing Center of Engineering Technology Research on Digital Agricultural & Service; [3]Chongqing Master Station of Agricultural Technology Promotion; [4]Chongqing Wanzhou District Station of Soil Fertilizer and Agricultural Ecological Protection; [5]Chongqing Beibei District Agriculture and Rural Committee, Chongqing, China

Correspondence: Shaohua Zeng, College of Computer & Information Science, Chongqing Normal University, Chongqing, 401331, China.
E-mail: zsh_cqu@126.com

Key words: adaptive density clustering; infructescence detection; yield estimation; Zanthoxylum.

Acknowledgments: this research was supported by Chongqing University Innovation Research Group Funding [Grant No. CXQT20015] and Research Project of Chongqing Normal University [Grant No. YKC21060].

Conflict of interest: the authors declare no potential conflict of interest.

Availability of data and material: the raw data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study.

Table S1. Framework of the density-based rough clustering.

Input:	Fruit regions.
Output:	Result of density-based rough clustering.
1.	Calculate the density of each point according to Eq. 3, and get the density map *DM;*
2.	Obtain feature weight map *FW* according to Eq. 7;
3.	Calculate the Hadamard product of *DM* and *FW* to get the feature density map *FDM*. The non-zero elements in *FDM* are extracted in the order of row-first to form $X = \{x_1, x_2, \ldots, x_N\}$;
4.	Find the maximum density points of *FDM* as the initial cluster centers;
5.	Assign labels to centers and form a series of cluster sets called *cluster_list* which only contains corresponding centers;
6.	Allocate unmarked points to the possible cluster sets according to Table 2;
7.	Merge clusters with duplicate elements into one cluster to update *cluster_list* by Table 3;
8.	Return *cluster_list.*

DM, density map; FW, feature weight; FDM, feature density map.

Table S2. Points allocation (step 6 of Table 1).

<table>
<tr><td>Input:</td><td>Points $X = \{x_1, x_2, \dots, x_N\}$, cluster sets cluster_list.</td></tr>
<tr><td>Output:</td><td>The updated cluster set cluster_list.</td></tr>
<tr><td>1.</td><td>Initializing: is_changed=0</td></tr>
<tr><td>2.</td><td>while True do</td></tr>
<tr><td>3.</td><td>| is_changed=0</td></tr>
<tr><td>4.</td><td>| for i from 1 to $len(X)$ do</td></tr>
<tr><td>5.</td><td>| | if x_i not in any cluster_list then</td></tr>
<tr><td>6.</td><td>| | | for each x_j in 4-neighbors of x_i do</td></tr>
<tr><td>7.</td><td>| | | | if x_j in cluster_list then</td></tr>
<tr><td>8.</td><td>| | | | | cluster_list[index of cluster set x_j belongs].append(x_i)</td></tr>
<tr><td>9.</td><td>| | | | | is_changed = 1</td></tr>
<tr><td>10.</td><td>| | | | end if</td></tr>
<tr><td>11.</td><td>| | | end for</td></tr>
<tr><td>12.</td><td>| | end if</td></tr>
<tr><td>13.</td><td>| end for</td></tr>
<tr><td>14.</td><td>| if is_changed == 0 then</td></tr>
<tr><td>15.</td><td>| | break</td></tr>
<tr><td>16.</td><td>| end if</td></tr>
<tr><td>17.</td><td>end while</td></tr>
</table>

Table S3. Cluster sets merging (step 7 of Table 1).

Input:	Cluster set *cluster_list*.
Output:	Density-based rough clustering results.

```
for i from 1 to len(cluster_list)-1 do
|    for j from (i+1) to len(cluster_list) do
|    |    if cluster_list [i]∩cluster_list [j] is not empty then
|    |    |    cluster_list [i] = cluster_list [i]∪cluster_list [j]
|    |    |    cluster_list [j] = empty
|    |    end if
|    end for
end for
```

Table S4. Local Calinski-Harabasz-based cluster merger.

Input:	Result of density-based rough clustering.
Output:	The optimal clusters.

```
while True do
|    for v_j in centers do
|    |    v_i = argmin_{v_x} d(v_x, v_j), where x ≠ j
|    |    Calculate the LCH between v_i and v_j by Eq. 13
|    end for
|    Find the optimal merging with Eq. 15 and record the p and q
|    Calculate the silhouette coefficient of the current merging, and remember it
|    if number of clusters==2 then
|    |    break
|    end if
end while
Find the optimal merge result by Eq. 16 and return the corresponding merger.
```

LCH, local Calinski-Harabasz.

Table S5. Framework of adaptive-density-clustering-based *Zanthoxylum* infructescence detection.

Input:	*Zanthoxylum* infructescence image *img*.
Output:	Number of infructescence and the location of infructescence.
1.	Extract the HOG feature map I_{HOG} from original image *img*.
2.	Obtain the thresholded HOG feature map I_{HOG}'.
3.	Extract the joint-direction-intensity feature map I.
4.	Calculate the ExGR index, and segment the plant and non-plant area to get *PM*.
5.	Fuse I and *PM* to obtain the fruit regions F'.
6.	Perform density-based rough clustering (Table 1).
7.	Iteratively merge the rough clustering results of Step 6, obtain the optimal clustering by Table 4.
8.	Mark the infructescence positions in the original image *img* and calculate the number of infructescence of infructescence based on the clustering results.

HOG, histogram of oriented; ExGR, excess green minus excess red.