Clustering Analysis Based on Segmented Images

Hongxu Zheng, Jianlun Wang(✉), and Can He

College of Information and Electrical Engineering,
China Agricultural University, Beijing 100083, China
496326832@qq.com, wangjianlun@cau.edu.cn,
348506582@qq.com

Abstract. Image segmentation plays an important role in the field of digital production management. Image resolution is an important factor affecting the size of its segmentation and segmentation efficiency, and the physical characteristics of the image capturing device is another important factor. With high-resolution segmentation algorithm in image segmentation, we often find that the edge contour image segmentation is difficult to accurately determine, more complex image arithmetic operation efficiency is not high and images taken with a different device in response to segmentation algorithms are very different. In this paper, the plant leaf image collected from different cameras was used as the object of study, and the feature quantity was extracted. The appropriate segmentation boundary was determined by cluster analysis. The leaf image was pretreated with the resolution adjustment, and the leaf image was in the appropriate segmentation feature range. After the clustering domain processing of the feature range in this paper, it solves the problem that the real edge of the leaf area information is too difficult to distinguish, and effectively solves the problem of complex image algorithm and ordinary pc machine in the process of complex image processing Efficiency issues. The appropriate segmentation feature range of the devices is established for different devices, which effectively solves the different response of different devices to the segmentation algorithm.

Keywords: Image segmentation · Resolution adjustment
Gray-level co-occurrence matrix · Clustering analysis

1 Introduction

Image segmentation is related to the resolution properties of the device and the image. Under the condition of natural light, the result is different from the plant leaves under different resolutions. In the field of image segmentation, the universality of a single algorithm is higher, such as Otsu [1], Canny algorithm [2], watershed algorithm [3], regional growth algorithm [4], mean drift clustering [5], Threshold [6], but when dealing with the same or different devices under different environmental conditions, the single segmentation operator is affected by factors such as physical condition, optical radiation environment, image noise, image complexity, threshold selection and so on. The image difference is mainly reflected in the complexity of the image, the resolution and the size of the image memory, and thus make its image segmentation efficiency and

X. Gu et al. (Eds.): MLICOM 2017, Part I, LNICST 226, pp. 57–75, 2018.
https://doi.org/10.1007/978-3-319-73564-1_7

effect are different. The complex composite algorithm effectively solves the influence of factors such as image noise, complexity and threshold selection when image segmentation is carried out. However, there is still no general segmentation algorithm for image acquisition devices under different environmental conditions to solve all the segmentation problems. And the existing segmentation algorithm is for different specific environmental conditions, specific equipment and the subject of the segmentation problem, so the division of the universal effect is low. With the improvement of various segmentation algorithms, the universal effect and efficiency of the composite segmentation algorithm for image segmentation of different specific subjects have been improved, such as Wang et al. [7] proposed an adaptive segmentation algorithm OTSU algorithm and Canny edge detection of plant leaf segmentation algorithm, you can get a better segmentation contour effect, the success rate to achieve a higher. The success rate is higher than the accuracy requirements, because in natural light conditions, the plant blade environment is complex, resulting in high image complexity. Not only requires a class of objects to accurately extract the edge, but also requires the algorithm. The segmented object has a certain degree of adaptability, that is, the same segmentation accuracy for objects with different environmental conditions. Dhalia Sweetlin et al. [8] proposed a method for segmenting CT images of lung disease based on patient-specific automated models, and also obtained a very high segmentation accuracy. However, this article only studies the accuracy of segmentation without involving success rates.

At present, all kinds of algorithm programs can solve the problem of incomplete and incomplete information when dealing with the complex images with high natural light and high resolution. At the same time, complex images with large amounts of information also slow down the image processing speed. So some scholars have made improvements. Such as Kim et al. [9], Frucci et al. [10], Liao et al. [11], Wang et al. [12], Zhu et al. [13]. In the high-resolution image segmentation, because of the large amount of image information, the high complexity of the factors is not conducive to the image of fast and accurate segmentation, but by adjusting the image resolution, Low resolution conditions, making the texture structure changes, to avoid excessive division, you can get the contours of the image, with better details to retain the characteristics and better anti-noise performance, thereby improving efficiency.

Some scholars have studied the problem of poor efficiency of segmentation image function algorithm in dealing with natural light image. Such as Moallem et al. [14], Delibasis et al. [15], Huang et al. [16], Saksa et al. [17], Zhang et al. [18]. The use of appropriate segmentation algorithm for a single device to obtain the image segmentation, can get a better segmentation effect. But for many devices to obtain the image, but rarely on the image segmentation effect is expressed. The study of multi-device image segmentation involves accuracy, but the efficiency of image segmentation is rarely mentioned. It can be seen that in the natural light state, the segmentation efficiency of the same segmentation algorithm for multi-device is still urgent to be solved, namely, the universality and efficiency of the algorithm. Natural light images of plant leaf images, not only by its complex background, shadow, light radiation and other factors, and different devices produce natural light conditions images, the same algorithm for them also have very different segmentation effect. The same type of camera and lens, because the impact of processing factors are not exactly the same, the imaging can not be exactly the same, but more than the different models to be closer. Therefore, the same algorithm in the different equipment image segmentation must effectively eliminate the impact of these factors, in order to make the image segmentation algorithm universality and efficiency is improved. In this paper, the plant leaf images obtained from three different cameras under natural light conditions are clustered according to the classification boundary formed by the sample base, and then the clustering is processed and then segmented to prove the versatility and goodness of the proposed method. The efficiency of image segmentation.

Some scholars have studied the attribute adjustment of clustering data and the effect of different clustering methods on clustering results. Such as Meged and Gelbard [19], Dee Miller et al. [20], Ji et al. [21], Hong et al. [22]. For the clustering of attribute data, the general research is to adapt the attributes by adjusting the algorithm, but there is little research on the algorithm to adapt the attributes by adjusting the attributes. Our approach is to adjust the raw data attributes so that the adjusted data can meet the attribute requirements processed by the segmentation algorithm and generate new raw data clusters. We have two requirements: 1. Attribute adjustment is appropriate; 2. Method adjustment appropriate.

For attribute data clustering, its attribute value, clustering algorithm will affect the effect of clustering. Through the adjustment of the attribute value can make the effect of clustering better, to achieve higher classification accuracy. Different clustering algorithms often get different clustering results for the same data, and the clustering results of the data can be obtained by comparing the different clustering algorithms. However, there is little research on the relationship between image attributes and segmentation effects and efficiency. In this paper, we find that the method can easily and effectively judge the segmentation effect and efficiency, and adjust the image attributes to achieve the purpose of improving the segmentation effect and efficiency. The research of this paper focuses on the segmentation effect and efficiency of the different images obtained by different devices in the process of plant leaf segmentation. When the complex image is segmented by complex algorithm, the complex features of the image are complex background, the leaf overlap, the difference between the veins and the leaf brightness, and the leaf edge gradient change is not obvious, which causes the image foreground and background separation difficult.

In this paper, a leaf adaptive image segmentation algorithm [7] is used to segment the plant leaves collected by different devices under natural light, and the attribute feature data are obtained. Through analysis, we select the three-dimensional feature Quantity, memory ratio, unit pixel entropy ratio, energy than the characteristics of the parameters of the amount of three-dimensional structure, on the basis of clustering. By comparing the clustering boundary obtained by different clustering methods, the optimal three-dimensional feasible domain boundary is selected. The boundary provides the basis for rapid classification of newly acquired image data and, on this basis, performs resolution adjustment. So that different image sources and different resolution images can be separated from the complete foliage edge contours, for the subsequent plant leaf domain biological characteristics of identification, three-dimensional reconstruction and other biomass calculation work pave the way. We apply the clustering method to the image attribute data of different segmentation results, which proves the simplicity and efficiency of the method.

2 Method

2.1 Materials

As the different models and manufacturers of camera features may be different, so this paper in the completion of the experiment used in the shooting equipment selected three different manufacturers and different models of the camera, they are: 10 million pixel CCD Nikon color camera, 2 megapixel CCD Canon color camera and 8 million pixel CCD Sony color camera; use of these three cameras were in the same shooting environment conditions collected plant leaf image.

In this paper, the quality of the image picture has a strict request, the camera equipment to capture the image content to keep the natural, real and clear texture, the actual life applications, the impact of camera image imaging factors are many, such as outdoor light direct, Jitter, focal length adjustment, noise interference, overlapping blade morphology and large reflection surface interference, so when shooting plant leaves, to avoid the interference described above, select the appropriate natural light shooting environment, try to ensure that the image clarity.

2.2 General Natural Light Image Segmentation Results

In the natural light condition, the plant leaf image collected in farmland has large information volume, high resolution, complicated background, overlapping leaves, poor leaf blade and leaf brightness, and no obvious change of edge gradient. The background is difficult to separate. All kinds of segmentation algorithm to process the segmentation of the contour is difficult to accurately determine the split efficiency is poor. For the same object different devices to shoot the image of a large difference, and the precise division, the different equipment, the efficiency of the division is also different. So a single algorithm can not accurately segment the blade contour. Artificial light adopts the adaptive binarization of the subject to obtain the binary map with small amount of information, but it is not conducive to the analysis of biomass and other plant image texture features. A large number of studies are now effective in solving this problem. In this paper, a blade adaptive image segmentation algorithm is used, as shown in Fig. 1. However, regardless of which kind of segmentation algorithm, will face the natural light, plant blade image high resolution, segmentation efficiency is not high, the difference between different devices caused by the impact of such issues. Therefore, we analyze the feature quantity of the reaction image information to find a solution to these problems.

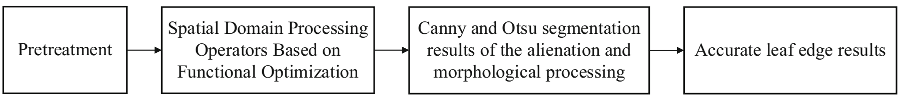

Fig. 1. The introduction of adaptive image segmentation algorithm

2.3 Analysis of Factors Affecting Image Segmentation

It is found that the image equipment and image information content have a great influence on the segmentation result, and it is necessary to carry out quantitative analysis.

2.3.1 Image Device on the Image Segmentation

The influence of image device and image information on image segmentation is as follows. First of all, different models and manufacturers of different physical characteristics of the camera, which CCD size, pixel size and processing accuracy, optical lens construction and processing accuracy and so on factors will cause imaging differences. Second, the same manufacturers of the same model of the physical characteristics of the camera can not be exactly the same, can also cause imaging differences. Again, the existing segmentation and other image algorithms and ordinary PC computing power is limited. In addition, the large amount of information on the leaf image is likely to cause the image processing function of the output is complicated, can not determine the real edge of the blade.

These problems are bound to result in the existing segmentation program can not be fully ideal for efficient and efficient segmentation. In this paper, the adaptive segmentation algorithm is based on the natural light conditions of the blade segmentation algorithm, the single blade of the success rate of about 70%. But the actual work of different camera equipment found different resolution of the blade image segmentation success rate is very different, but there are some centralized laws, as shown in Fig. 4.

In order to enable the segmentation algorithm to efficiently segment different images and images of different devices under natural light conditions, it is necessary to process the images taken by different devices so that they can be in an ideal divisible Resolution domain and information domain, and establish the clustering and clustering threshold of the domain. On this basis, it is necessary to determine the three-dimensional boundary which is suitable for segmentation of different devices in order to solve the above four problems.

2.3.2 Image Resolution on the Segmentation of the Impact

The resolution of the image affects the segmentation effect of the image. High-resolution images are smooth edges, rich in detail and texture, giving a soft feel. However, when the high-resolution plant leaves are divided, there is a problem that the success rate of the equipment is high, the division time is too long, and the contour of the appropriate division is obtained. The edge of the low-resolution image will have obvious grain and jaggedness, the image's sharpness is poor, affecting the quality of the image. In the segmentation, + segmented contour is more smooth, loss of detail information, can not be a good response to plant leaf area and so on.

Resolution Adjust the experiment to be used in the steps: First, the resolution adjustment, access to different resolution images; Second, the image will be adjusted after the split experiment; Thirdly, the clustering threshold of the feature quantity of the image segmentation result of each experimental camera is analyzed.

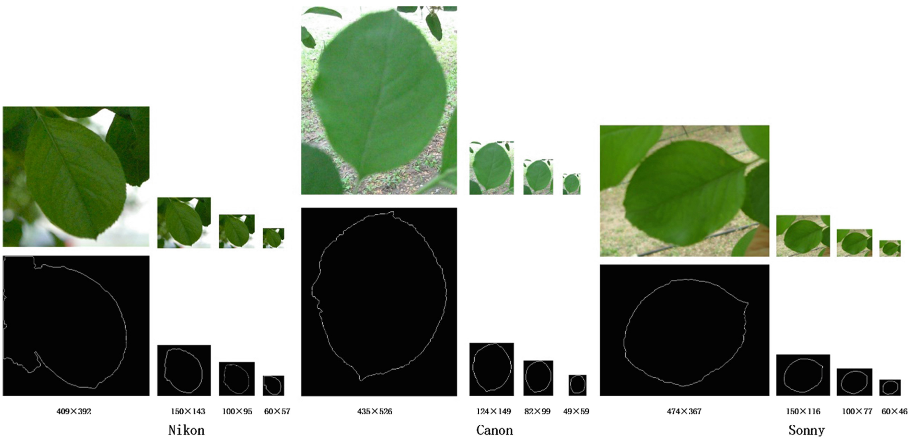

Fig. 2. The segmentation effect of the leaves taken by Nikon camera.

Table 1. The resolution range of single blade image which adapts segmentation algorithm.

Camera equipment	Image label	Suitable for segmentation of the resolution area	Suitable for segmentation of the pixel area
Nikon	Nikon (221)	80 * 77–210 * 201	6160–42210
Canon	Canon (15)	37 * 45–433 * 525	1685–227325
Sonny	Sonny (37)	35 * 26–525 * 404	945–212575

Through the experiment shown in Fig. 2, the resolution of the vane image affects the effect of dividing the edge contour.

Table 1 summarizes the range of suitable resolution and the number of suitable pixels for the monolithic images nikon (221), canon (15) and sonny (37) from the three camera devices.

Since each of the blade images has a different resolution domain that is suitable for segmentation, the resolution of the resolution range of the 90 monolithic leaf image samples of each group of camera devices is defined as the resolution spatial domain library of the device.

Through the observation and analysis, we found that the characteristics of the plant leaves obtained by different devices under natural light conditions also have important influence on the segmentation of the images. The success rate of the different resolution images obtained by different camera devices is very different. There are some centralized laws, through the analysis of the amount of features, you can effectively obtain the different camera equipment to capture the appropriate image of the plant blade area.

3 Image Feature Analysis

The energy of the image is the square sum of the elements of the gray covariance matrix. It can measure the stability of the gray scale of the texture. The larger the value is, the more stable the gray scale is, and the gray distribution of the image is more uniform. The entropy of the image belongs to the inherent attribute of the image. Entropy can not only reflect the density of the gray distribution in the image, but also reflect the spatial characteristics of the gray spatial distribution. The resolution of the image, that is, the resolution of the image, reflects the amount of information and detail stored in the digital image, usually expressed in terms of the number of pixels per unit inch.

3.1 Feature Selection

The resolution of the plant leaves with different complexity obtained under different natural conditions was analyzed. The resolution, pixel entropy, memory and gray covariance were selected as parameters to study the suitable segmentation of plant leaves.

3.1.1 Unit Pixel Entropy, Resolution and Memory

The higher the entropy of the image, the greater the detail in brightness and the conversion, the need for higher compression; otherwise, you need less compression. The unit pixel entropy reflects the size of the image on the unit pixel. Pixel is the abbreviation of the image element. In the computer operation analysis, if the image is magnified several times, the human eye can find that these continuous regions and texture are concatenated by a lot of squares with similar color or similar color, Square small area element is the smallest unit of pixels that make up the computer digital image, the pixel is also called the pixel or pixel element, the pixel is the resolution of the size unit. The larger the pixel, the higher the resolution, the clearer the picture, and the larger the output photo size. The size of the memory of the image file is proportional to the square of its image resolution, and the pixel of the image is proportional to the resolution of the image.

In this experiment, the unit pixel entropy, pixel ratio and memory ratio of 90 single leaf images of Nikon camera are used to obtain the ratio of the corresponding feature of the corresponding image by adjusting the resolution to select the characteristic parameter quantity suitable for clustering.

Fig. 3. Ratio of parameters before and after compression taken by Nikon camera.

As shown in Fig. 3, the abscissa order is sorted in ascending order according to the pixel ratio of the image before and after the resolution is adjusted. The memory ratio is proportional to the pixel ratio. The memory ratio and pixel ratio are independent of the unit pixel entropy ratio.

3.1.2 Gray Covariance Moment

The gray-level co-occurrence matrix [23] can describe the texture by studying the spatial correlation properties of the gray scale, and transform the spatial distribution information of the gray level into texture information. Haralick et al. [24] proposed 14 methods of texture quantization based on texture features such as uniformity, energy, variance, contrast, entropy, and inverse moment. In the boundary analysis, the energy parameters are used to construct the harmonic function boundary. Therefore, the energy ratio is chosen as the third dimension parameter.

3.2 Dimensional Space

Through the pixel ratio and the memory ratio of Fig. 3 is proportional to the ratio, with the associated nature, so these two parameters can only choose a participant in the data analysis; unit pixel entropy ratio is independent of the other two parameters. Therefore, this paper chooses the memory ratio, the unit pixel entropy ratio, and the energy ratio as the parameters to construct the clustering space.

Taking the Nikon camera as an example, the upper boundary vane image, which is suitable for segmentation, is suitable for segmentation of the blade image, and the resolution, unit pixel entropy of the image is extracted for each blade image and the appropriate segmented upper blade image after adjusting the resolution. Energy, adjust the resolution of these parameters and the appropriate division is the ratio of plant leaf parameters, constitute a three-dimensional space, as shown in Fig. 4.

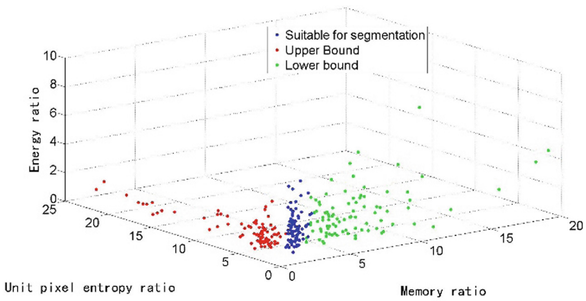

Fig. 4. Three dimensional data distribution map taken by Nikon camera.

4 Cluster Analysis

Cluster analysis itself is based on data to explore the data object and its relationship information, and the data grouping. The objects within each group are similar, and the objects between the groups are irrelevant. The higher the similarity in the group, the higher the heterogeneity between groups, the better the clustering. In general, the clustering method is divided into hierarchical clustering method, hierarchical clustering algorithm, density-based clustering algorithm, model-based clustering algorithm and grid-based clustering algorithm. The clustering algorithm based on the partitioning method requires the number of given clusters K, and the number of classes based on the user-defined hope in the hierarchical partitioning method is usually the end condition. The clustering algorithm based on density, the clustering algorithm based on the model and the grid-based clustering algorithm can not determine the number of clustering clusters in advance. The data studied in this paper are clustered by the upper and lower bounds of the appropriate segmentation of the plant leaf image, and the number of clustering clusters is known in advance. In order to obtain the most suitable segmentation region, K-Means [25], BIRCH and K-Medoids were used for comparative analysis.

4.1 Clustering Experiments

By analyzing the data in Fig. 4, it is found that the data space domain effect is ideal, but the threshold of the appropriate segmentation space domain is obtained. Through the calculation of clustering parameters, it is possible to obtain the appropriate spatial domain parameters of plant leaves, which can give the threshold of quantization adjustment resolution and other feature quantity.

The clustering analysis of the data points in the three-dimensional space is carried out. The Nikon camera uses the k-means algorithm as an example to cluster the data points of different colors respectively. Figure 5 is a Nikon camera three-dimensional data clustering map. According to this can be summarized Nikon camera image clustering data information, as shown in Table 2.

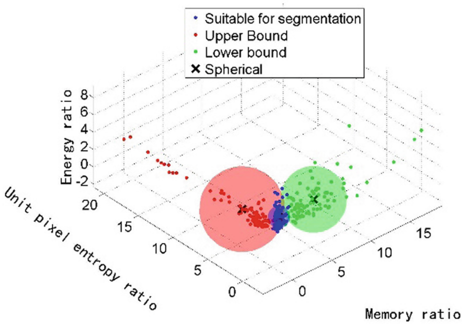

Fig. 5. Clustering space of three dimensional data taken by Nikon camera (ASM ratio). (Color figure online)

Table 2. Three dimensional data clustering information by Nikon camera.

Clustering algorithm	Spherical			
		Suitable for segmentation	Upper bound	Lower bound
K-Means	Coordinates	1.4073 0.6145 2.3564	0.3044 4.6680 1.7188	5.4744 0.1865 2.8003
	Radius	1.07	4.11	3.20
BIRCH	Coordinates	1.6388 0.6603 2.8889	0.4364 9.0446 2.1712	8.9659 0.23 4.3599
	Radius	3.55	9.12	6.32
K-Medoids	Coordinates	1.3769 0.6148 2.2725	0.3066 3.6632 1.6797	5.1053 0.1777 2.6835
	Radius	1.05	3.29	3.11

Similarly, Canon and Sony camera image clustering of data information, as shown in Tables 3 and 4.

Table 3. Three dimensional data clustering information by Canon camera.

Clustering algorithm	Spherical			
		Suitable for segmentation	Upper bound	Lower bound
K-Means	Coordinates	1.2723 0.7913 1.9610	0.2567 6.4207 1.4624	4.7749 0.2667 2.3417
	Radius	0.73	5.51	2.88
BIRCH	Coordinates	1.3633 0.8764 2.205	0.2722 10.9753 1.5037	7.7996 0.3356 2.4728
	Radius	2.7	10.49	5.65
K-Medoids	Coordinates	1.2642 0.7911 1.9121	0.2724 4.8999 1.4253	4.3015 0.2560 2.2711
	Radius	0.73	4.13	2.55

Table 4. Three dimensional data clustering information by Sonny camera.

Clustering algorithm	Spherical			
		Suitable for segmentation	Upper bound	Lower bound
K-Means	Coordinates	1.2548 0.8328 0.7834	0.2010 10.0731 0.6178	5.2534 0.2198 0.9291
	Radius	1.94	9.29	2.35
BIRCH	Coordinates	1.4507 0.8222 1.112	(0.2503 14.0611 0.7685	8.7385 0.2611 1.3843
	Radius	1.84	13.31	6.08
K-Medoids	Coordinates	1.2225 0.8390 0.7746	0.2089 7.3437 0.5997	4.9716 0.2151 0.8991
	Radius	1.94	6.84	2.30

Through the data from Tables 2, 3 and 4, we can see the clustering results, and provide the basis for improving the efficiency of blade segmentation.

(1) Through the analysis of the feature range, it is found that the leaf image has different feature quantity attributes with its resolution.
(2) Different camera equipment its suitable segmentation threshold range is different.
(3) Through the analysis of the appropriate threshold of different equipment, you can get the appropriate segmentation feature range of the device.

4.2 Clustering Results Analysis

Cluster analysis take Nikon camera as an example, select any single-leaf image, and adjust it to the three resolution fields shown in Fig. 5. Then, the image data points falling in the three clustering domains are selected, and their corresponding images are segmented edge contours to obtain three leaf edge contours as shown in Fig. 6.

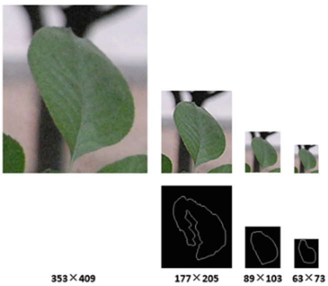

Fig. 6. The segmentation effect of new leave image by Nikon, Sonny and Canon camera.

From the edge of the camera segmentation effect, as shown in Figs. 5 and 6, the main process and the results are as follows.

(1) The three different sets of original resolution images of different cameras through the program automatically traverse cut to obtain a number of monolithic plant leaves of the image;
(2) According to the proportion of adjustment rules, the different equipment of the image group, adjust the resolution of the blade image and split its outline until the appropriate segmentation of the blade contour;
(3) Filtering the feature quantity of the adjusted image, and obtaining the appropriate resolution domain and the corresponding other feature range for each leaf image;
(4) For three-dimensional feature quantity clustering. The resolution field and the feature range of each segment are clustered according to their segmentation results. The image texture energy ratio of the gray level co-occurrence matrix based on the image is used as the third dimension data volume of the clustering space, and the three-dimensional data is clustered to obtain the appropriate segmentation feature. The range of gravity and the clustering radius of the upper bound of the feature quantity and the boundary of the appropriate segmentation feature;
(5) The clustering of each group is taken into the intersection to obtain the appropriate resolution of each device and the threshold of the eigenvalue clustering domain.

The resolution of the original image is adjusted by the clustering domain with the resolution ratio and the eigenvalue ratio, and the edge segmentation is carried out, which greatly improves the success rate and running efficiency of the existing segmentation algorithm. But because of its part of the data can not distinguish between its scope, therefore, need a reasonable boundary, in order to better define its appropriate segmentation of the region.

5 Three-Dimensional Harmonic Boundary

Aiming at the attribution problem of sample data points in overlapping regions of different camera clusters, this paper constructs the boundary function of overlapping clustering space, and divides the clustering space. And the three-dimensional clustering information of three camera devices is combined with Laplace equation and Gibbs free energy to reconcile the boundary function of clustering overlapping region.

5.1 Harmonic Boundary Function

The harmonic function equation is a second order continuous derivative function (which is an open subset) on the domain of the definition. The harmonic function satisfies the Laplace equation, that is, satisfies the Eq. (1):

$$\frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \ldots + \frac{\partial^2 f}{\partial x_{\mathrm{n}}^2} = 0 \tag{1}$$

Equations (1) can also be written or, where the symbol is called the Laplacian operator. The Laplace operator is defined as the divergence of the gradient and can represent the transport of matter due to the uneven distribution of matter. On the Riemannian manifold, the harmonic function has another definition. The Laplacian operator is called the Laplacian-Drumm operator. In this case, the harmonic function is satisfied.

Gibbs free energy is defined as:

$$\mathrm{G} = \mathrm{U} - \mathrm{TS} + \rho\mathrm{V} \tag{2}$$

Where G is Gibbs free energy, U is the internal energy of the whole system, T is the temperature of the system, S is the entropy of the system, ρ is the pressure, V is the volume, and $\mathrm{H} = \mathrm{U} + \rho\mathrm{V}$ is the enthalpy of the system. At room temperature and pressure system, Gibbs free energy can be completely determined by the system's internal energy.

By constructing the surface satisfying the harmonic function Eq. (1), we can determine the boundary between the "complete segmentation" and the "lower bound" medium clustering domain. The Laplace pressure between the two clusters is used to obtain the curve equation. Together, the formula for the Laplace pressure can be deduced from the Young-Laplacian formula:

$$\Delta\mathrm{P} \equiv \mathrm{P}_{\mathrm{in}} - P_{out} = \gamma\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{3}$$

Where R_1 and R_2 are the radius of curvature and γ is the tension coefficient of the surface. In general, the convex surface can be represented by a positive curvature radius, and the concave surface is represented by a negative curvature radius. By the definition of the surface tension coefficient, it can be seen that the surface tension coefficient γ can be obtained by the partial derivative of the area A of Gibbs free energy G in the case where the temperature T and the pressure P remain constant:

$$\gamma = \left(\frac{\partial \mathrm{G}}{\partial \mathrm{A}}\right)_{T,P} \tag{4}$$

The Gaussian free energy G of the sphere can be substituted into the Eq. (4) by calculating the sum of the sum of the squares of the elements of the normalized image and the second order moments in the angular direction. The curve equation fits the desired set of points.

5.2 The Construction of 3D Boundary Function and the Correctness

The attribution of sample data points in overlapping regions of different camera clusters requires the use of Gibbs free energy as the basis for constructing the harmonic boundary.

The two spherical domains are modeled into two different media, and the data fitting points of the three cameras are obtained by calculating their free energy, and then the harmonic boundary equation is found.

The steps to obtain three fitting points include:

First, calculate the Laplace pressure difference:

Gibbs free energy expression for the $G = U - TS + \rho V$, due to the system at room temperature and pressure, so the free energy changes only by the system's own energy. The total value G_1 of the image energy value after the compression resolution of 90 data points in the "complete segmentation" sphere, the surface area S_1 of the "complete segmentation" sphere, and the compression resolution of 90 data points in the "lower bound" sphere The total value of the image energy after the value of G_2, the "lower bound" sphere surface area S_2. The surface tension γ_1 and γ_2 of "complete segmentation" and "lower bound" are obtained by using the surface tension formula $\gamma = G/S$, then the difference between the energy per unit area of the "suitable segmentation" and the "lower bound" cluster sphere is $\Delta\gamma = \gamma_1 - \gamma_2$. Into the formula (2) can be obtained between the two spheres of the Laplace pressure difference ΔP;

Second, the boundary fixed:

The pressure difference is projected onto the two core wires, and the boundary vertices of the harmonic function are obtained.

Third, calculate the two fitting points:

With the energy of the ball on the distance from the two spherical distance from the farthest point as a fitting point.

Fourth, calculate the boundary parameters:

According to the vertex and two fitting points can be determined to determine the center of the boundary and radius.

According to the above steps, the boundary functions are determined as shown in Table 5.

The results show that the efficiency of the segmentation of the plant leaves in the boundary is shown in Table 6 by analyzing the results of 90 suitable segmentation resolution domains.

Through the above comparison, the boundary obtained by K-Means clustering algorithm is the best, because the boundary obtained by this algorithm can adjust the blade image segmentation.

5.3 Segmentation Success Rate of Three-Dimensional Limit Function

Under natural light, the obtained leaf image of the plant was selected and 30 leaflets were selected for validation. The experimental flow is shown in Fig. 7. The efficiency of the plant leaves before and after the adjustment of the equipment is shown in Table 7. The results show that there is no adjustment of the resolution directly on the split edge of the edge of the edge of the success rate is relatively low, the average running time is longer.

Table 5. Three dimensional harmonic boundary function.

Clustering algorithm	Camera equipment	Dimensional harmonic boundary function	Domain
K-Means	Nikon	$(x-0.25)^2+(y-4.87)^2+(z-1.69)^2=4.19^2$ $(x-5.6)^2+(y-0.17)^2+(z-2.81)^2=3.29^2$	$(4.19*x-0.44*y+0.46*z-24.66<0)$ $(1.16*x-4.26*y+0.67*z-19.3>0)$
	Canon	$(x-0.14)^2+(y-7.05)^2+(z-1.41)^2=6.11^2$ $(x-7.85)^2+(y+0.19)^2+(z-2.68)^2=6^2$	$(1.13*x-6.25*y+0.55*z+43.12>0)$ $(6.58*x-0.98*y+0.71*z-53.73<0)$
	Sonny	$(x-0.18)^2+(y-10.28)^2+(z-0.61)^2=9.3^2$ $(x-5.37)^2+(y-0.2)^2+(z-0.93)^2=2.36^2$	$(1.08*x-9.45*y+0.17*z+96.82>0)$ $(4.11*x-0.63*y+0.15*z-22.08<0)$
BIRCH	Nikon	$(x-0.33)^2+(y-9.79)^2+(z-2.11)^2=9.15^2$ $(x-9.75)^2+(y-0.18)^2+(z-4.52)^2=6.38^2$	$(1.31*x-9.13*y+0.78*z+87.22>0)$ $(8.11*x-0.48*y+1.63*z-86.42<0)$
	Canon	$(x-0.23)^2+(y-11.33)^2+(z-1.48)^2=10.49^2$ $(x-8.34)^2+(y-0.3)^2+(z-2.5)^2=5.68^2$	$(1.13*x-10.52*y-0.73*z+117.82>0)$ $(6.98*x-0.52*y+0.29*z-58.74<0)$
	Sonny	$(x-0.24)^2+(y-14.19)^2+(z-0.77)^2=13.31^2$ $(x-8.87)^2+(y-0.25)^2+(z-1.39)^2=6.09^2$	$(1.21*x-13.37*y+0.35*z+189.15>0)$ $(7.42*x-0.57*y+0.28*z-66.09<0)$
K-Medoids	Nikon	$(x-0.2)^2+(y-3.96)^2+(z-1.62)^2=3.44^2$ $(x-5.33)^2+(y-0.15)^2+(z-2.71)^2=3.25^2$	$(1.17*x-3.34*y+0.65*z+11.95>0)$ $(3.96*x-0.46*y+0.44*z-22.2<0)$
	Canon	$(x-0.01)^2+(y-6)^2+(z-1.29)^2=5.21^2$ $(x-10.97)^2+(y-0.92)^2+(z-3.06)^2=9.33^2$	$(1.26*x-5.21*y+0.62*z+30.48>0)$ $(9.7*x-1.71*y+1.15*z-111.53<0)$
	Sonny	$(x-0.16)^2+(y-7.64)^2+(z-0.59)^2=6.85^2$ $(x-5.19)^2+(y-0.18)^2+(z-0.91)^2=2.32^2$	$(1.06*x-6.8*y+0.18*z+51.61>0)$ $(3.96*x-0.66*y+0.13*z-20.56<0)$

Table 6. The correctness of the constructor.

Clustering algorithm	Camera equipment	Correct rate %
K-Means	Nikon	92.22
	Canon	97.78
	Sonny	94.44
BIRCH	Nikon	78.89
	Canon	80
	Sonny	77.5
K-Medoids	Nikon	80
	Canon	84.44
	Sonny	58.89

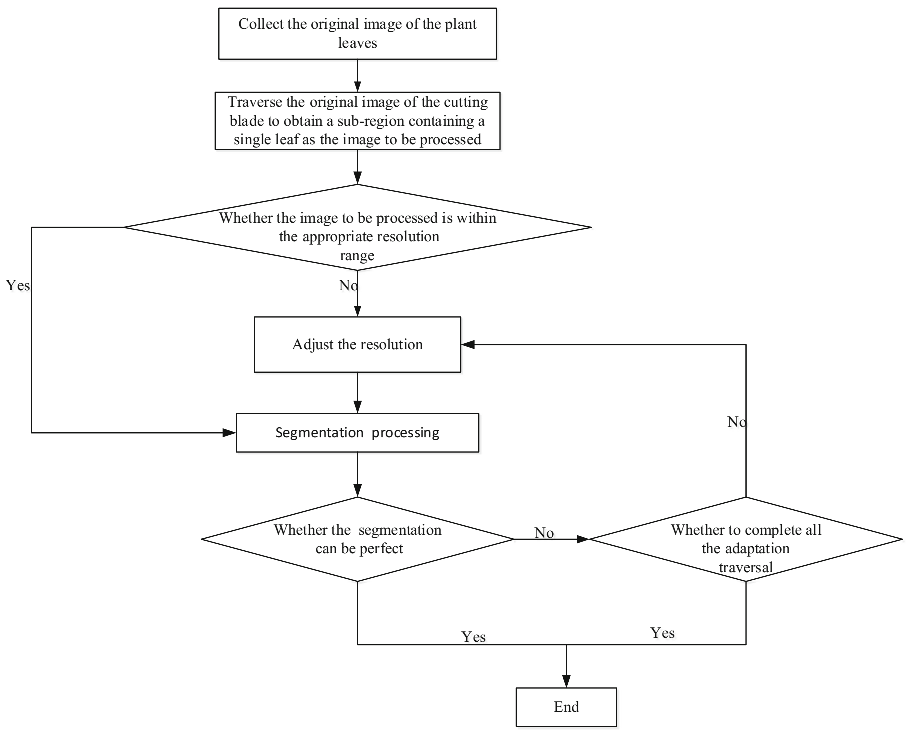

Fig. 7. Flow chart of image resolution adjustment.

Table 7. The segmentation efficiency changes between the leave images before and after the characteristic parameters.

Camera equipment	Before adjustment		After adjustment	
	The rate of segmentation/%	Time/m	The rate of segmentation/%	Time/m
Nikon	26.67	>20	73.33	1.14
Sonny	23.33	>20	66.67	1.26
Canon	33.33	>20	70	1.05

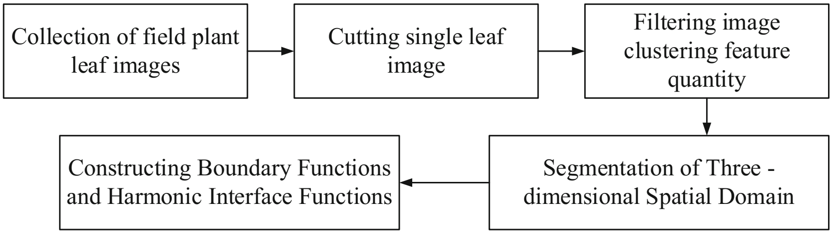

Fig. 8. The process of obtaining the spatial domain of image features and spatial resolution

6 Conclusion

In this paper, the plant leaf images collected from different cameras (Nikon, Canon and Sony, respectively) were used as the object of study. After a large number of feature information were extracted and sorted, the clustering analysis was carried out, and the leaf image was pretreated by resolution adjustment. The image is in a reasonable feature range based on clustering analysis. Which effectively prevents the computer from consuming too much load operation. At the same time, it effectively solves the problem that the real edge of the leaf area information is too large is difficult to distinguish. After analyzing the feature area of different devices by clustering processing of feature area, the appropriate segmentation feature area of each device is established. And the image based on the domain is adjusted according to the feature size of the image for efficient segmentation.

According to the above experimental and experimental results, the process of leaf image feature and resolution adjustment is shown in Fig. 8.

According to the study of this paper, the resolution size is a factor that affects its segmentation effect and segmentation time. Another factor is the physical characteristics of the image capture device. In order to improve the segmentation success rate of the program and improve the efficiency of the algorithm, this paper obtains the resolution range of each device adapting to the algorithm by comparing the resolution of different cameras and improves the success rate and efficiency of the algorithm. (Image energy, entropy, moment of inertia, correlation), and select the appropriate parameters to do the clustering analysis, to find out the resolution of the image (the number of pixels), the memory, the unit pixel entropy and the image grayscale covariance matrix texture feature vector. The center of gravity of the clustering domain and the radius and feasible domain, and the boundary of the resolution adjustment feature corresponding to the segmentation algorithm for the captured image of the device is determined.

Acknowledgments. This work was supported by the Beijing 'The agricultural technology promotion information project in Hebei Province' program and was undertaken by China Agricultural University.

References

1. Zhou, Y., Gao, M., Fang, D., Zhang, B.: Unsupervised background-constrained tank segmentation of infrared images in complex background based on the Otsu method. SpringerPlus. **5**(1), 1–15 (2016)
2. Jiang, W., Zhou, H., Shen, Y., Liu, B., Fu, Z.: Image segmentation with pulse-coupled neural network and Canny operators. Comput. Electr. Eng. **46**, 528–538 (2015)
3. Wang, M., Wang, Y., Chai, L.: Application of an improved watershed algorithm in welding image segmentation. Russ. J. Nondestr. Test. **47**(5), 352–357 (2011)
4. Frulla, L., Lira, J.: An automated region growing algorithm for segmentation of texture regions in SAR images. Int. J. Remote Sens. **19**(18), 3595–3606 (1998)
5. Wu, G., Zhao, X., Luo, S., Shi, H.: Histological image segmentation using fast mean shift clustering method. Biomed. Eng. Online **14**(1), 1–12 (2015)
6. Chen, M.F., Zhu, H.S., Zhu, H.J.: Segmentation of liver in ultrasonic images applying local optimal threshold method. Imaging Sci. J. **61**(7), 579–591 (2013)
7. Wang, J., He, J., Han, Y.: An adaptive thresholding algorithm of field leaf image. Comput. Electron. Agric. **2013**(96), 23–39 (2013)
8. Dhalia Sweetlin, J., Khanna Nehemiah, H., Kannan, A.: Patient-specific model based segmentation of lung computed tomographic images. J. Inf. Sci. Eng. **32**(5), 1373–1394 (2016)
9. Kim, B.-G., Shim, J.-I., Park, D.-J.: Fast image segmentation based on multi-resolution analysis and wavelets. Pattern Recogn. Lett. **24**(16), 2995–3006 (2002)
10. Frucci, M., Ramella, G., Sanniti di Baja, G.: Using resolution pyramids for watershed image segmentation. Image Vis. Comput. **25**(6), 1021–1031 (2007)
11. Liao, C.-C., Xiao, F., Wong, J.-M., Chiang, I.-J.: A multiresolution binary level set method and its application to intracranial hematoma segmentation. Comput. Med. Imaging Graph. **33**(6), 423–430 (2009)
12. Wang, W., Qin, J., Chui, Y.P., Heng, P.A.: A multiresolution framework for ultrasound image segmentation by combinative active contours. IEEE Eng. Med. Biol. Soc. **2013**, 1144–1147 (2013)
13. Zhu, H., Cai, L., Liu, H., Huang, W.: Information extraction of high resolution remote sensing images based on the calculation of optimal segmentation parameters. PLoS One **11**(6), 1–15 (2016)
14. Moallem, P., Razmjooy, N., Mousavi, B.S.: Robust potato color image segmentation using adaptive fuzzy inference system. Iran. J. Fuzzy Syst. **11**(6), 47–65 (2014)
15. Delibasis, K.K., Goudas, T., Maglogiannis, I.: A novel robust approach for handling illumination changes in video segmentation. Eng. Appl. Artif. Intell. **49**, 43–60 (2016)

16. Huang, Z., Zhang, J., Li, X., Zhang, H.: Remote sensing image segmentation based on dynamic statistical region merging. Optik - Int. J. Light Electron Optics **125**(2), 870–875 (2014)
17. Saksa, T., Uuttera, J., Kolström, T., Lehikoinen, M., Pekkarinen, A., Sarvi, V.: Clear-cut detection in boreal forest aided by remote sensing. Scand. J. For. Res. **18**(6), 537–546 (2003)
18. Zhang, X., Xiao, P., Song, X., She, J.: Boundary-constrained multi-scale segmentation method for remote sensing images. ISPRS J. Photogramm. Remote Sens. **78**, 15–25 (2013)
19. Meged, A., Gelbard, R.: Adjusting fuzzy similarity functions for use with standard data mining tools. J. Syst. Softw. **84**(12), 2374–2383 (2011)
20. Dee Miller, L., Soh, L.-K., Scott, S.: Genetic algorithm classifier system for semi-supervised learning. Comput. Intell. **31**(2), 201–232 (2015)
21. Ji, J., Pang, W., Zhou, C., Han, X., Wang, Z.: A fuzzy k-prototype clustering algorithm for mixed numeric and categorical data. Knowl.-Based Syst. **30**, 129–135 (2012)
22. Hong, X., Wang, J., Qi, G.: Comparison of semi-supervised and supervised approaches for classification of e-nose datasets: case studies of tomato juices. Chemom. Intell. Lab. Syst. **146**, 457–463 (2015)
23. MacQueen, J.: Soma methods for classification and analysis of multivariate observations. In: Proceedings of the Symposium on Mathematical Statistics and Probability, pp. 281–297. University of California Press, Berkeley (1967)
24. Haralick, R.M., Shanmugam, K., Dinstein, I.H.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)
25. Kim, H.A., Karp, B.: Autograph: toward automated distributed worm signature detection. In: Proceedings of the 13th Usenix Security Symposium (2004)