Fast Image Segmentation Algorithm Based on Salient Features Model and Spatial-frequency Domain Adaptive Kernel

WU Fupei, LIANG Jiaye, LI Shengping

(*Department of Mechanical Engineering*, *College of Engineering*, *Shantou University*, *Shantou* 515063)

Abstract: A fast image segmentation algorithm based on salient features model and spatial-frequency domain adaptive kernel is proposed to solve the accurate discriminate objects problem of online visual detection in such scenes of variable sample morphological characteristics, low contrast and complex background texture. Firstly, by analyzing the spectral component distribution and spatial contour feature of the image, a salient feature model is established in spatial-frequency domain. Then, the salient object detection method based on Gaussian band-pass filter and the design criterion of adaptive convolution kernel are proposed to extract the salient contour feature of the target in spatial and frequency domain. Finally, the selection and growth rules of seed points are improved by integrating the gray level and contour features of the target, and the target is segmented by seeded region growing. Experiments have been performed on Berkeley Segmentation Data Set, as well as sample images of online detection, to verify the effectiveness of the algorithm. The experimental results show that the *Jaccard Similarity Coefficient* of the segmentation is more than 90%, which indicates that the proposed algorithm can availably extract the target feature information, suppress the background texture and resist noise interference. Besides, the *Hausdorff Distance* of the segmentation is less than 10, which infers that the proposed algorithm obtains a high evaluation on the target contour preservation. The experimental results also show that the proposed algorithm significantly improves the operation efficiency while obtaining comparable segmentation performance over other algorithms.

Keywords: Image Segmentation, Spatial-frequency Domain, Adaptive Convolution Kernel, Online Visual Detection

1 Introduction

Image segmentation is an important part of visual detection. The performance of image segmentation algorithm is crucial to the accuracy and reliability of target detection and recognition. Accurate and adaptive segmentation of target and background has been a hot topic in current image processing research. Traditional image segmentation algorithms mainly perform segmentation based on the similarity and discontinuity of pixel gray level. Segmentation algorithm based on pixel gray level value similarity can be divided into thresholding [1] and region-based segmentation [2]. Thresholding is based on a threshold, dividing image

pixels into target pixels and background pixels by designing appropriate thresholds. It is simple in form and clear in calculation target, but the segmentation is not ideal when facing complex images or requiring more objects to be distinguished. Therefore, a variety of statistical characteristics are introduced into threshold selection, such as maximum interclass variance [3] (OTSU), moment [4] (TSAI), entropy [5] (Kapur *et al.*) and the maximum likelihood [6] (Kurita *et al.*), etc. Region-based segmentation can be divided into two types: seeded region growing, region splitting and merging. Seeded region growing starts from a single pixel and gradually merges neighboring similar pixels to form the desired segmentation. In contrast, region splitting and merging are carried out gradually from the whole map until the required segmented regions are generated. Region-based segmentation considers location information such as region similarity, boundary and smoothness, and is robust to various images [2]. However, the selection of initial seeds has a great influence on the final segmentation. It is necessary to determine a set of seed pixels that can correctly represent the desired region according to the image characteristics and segmentation requirements. Kang *et al.* [8] presented a seeded region growing algorithm based on the fuzzy theory, which automatically selects the initial seeds combined with the fuzzy image pixel similarity, to avoid the initial seeds appearing in the details and complex background. Li *et al.* [8] introduced an optimized automatic seeded region growing (OASRG) algorithm, which uses the Affinity Propagation (AP) clustering algorithm to extract the seeds automatically. In the process of growing of seeds, the current pixel state can be merged with seeds is decided by the similarity measured by LAB color values and Canny edge, which improves the segmentation accuracy and robustness. The second method is to segment the image into several regions according to the mutation of pixel gray level at the boundary of each region, such as edge detection algorithm [9]. Traditional edge detection algorithms have high efficiency and accurate positioning, while they fail to ensure the continuity and closedness of the edges, and there are usually a large number of broken edges in the high-detail regions after segmentation, which causes confusion to form a large region. Arbeláez *et al.* [10] proposed a contour detection and hierarchical image segmentation algorithm. By designing a high-performance contour detector, the algorithm combines multiple local clues into a framework based on spectral clustering, and it converts all contour information into a hierarchical region tree while maintaining the contour quality, simplifies the image segmentation problem into a contour detection problem. Then it reconstructs each closed region by using discontinuous contour information through directional watershed transformation. Thanks to the development of digital image technology, image resolution has been continuously improved, and detailed texture information has been significantly increased. Traditional segmentation algorithms perform segmentation based on low-level semantics such as gray level, color, and edge texture, which usually causes confusion to segment high-definition images efficiently and accurately. As we know, clustering is an unsupervised classification method while unlabeled data sets can be classified by the similarity measure set in the cluster analysis to form multiple clusters. Cluster-based image segmentation algorithm [11] usually divides the pixels into different categories by quantitatively comparing the multi-dimensional features of pixels. Compared with traditional image segmentation methods, cluster-based image segmentation algorithms have higher accuracy and robustness [12]. For example, Felzenszwalb [13] developed a graph-based greedy clustering segmentation method, which uses the idea of minimum cut in graph theory to divide the image into several parts, and transforms the optimal segmentation into finding the minimum spanning tree (MST) of the sum of edge weights. Achanta *et al.* [14] proposed a simple linear iterative clustering (SLIC) algorithm based on the K-Means clustering segmentation. It converts the image into a five-dimensional feature vector in CIELAB space and XY coordinates, and it constructs a distance metric for the five-dimensional feature vector. It performs local clustering of image pixels to generate compact and approximately uniform

super-pixels [15-16]. Kashif *et al.* [17] proposed a generalized kernel weighted fuzzy C-means clustering algorithm with local information (GKWFLICM) for m-dimensional input data sets. It adds neighborhood information to the m-dimensional data, which improves the shortcomings of the standard FCM clustering algorithm that is sensitive to noise and outliers, and it has poor performance for clusters of different sizes and different densities. This approach significantly improves the clustering performance. With the continuous improvement of industrial automation level, the visual online detection demand of production objects [18-19] is increasing, and image segmentation algorithms with both accuracy and efficiency are urgently needed. However, the segmentation performance of the currently developed image segmentation algorithm still needs to be improved in industrial applications where the morphological characteristics of samples are diverse and the detection environment is complex and vary. Therefore, traditional segmentation algorithms are difficult to suppress the background texture and noise information effectively when facing complex detection samples, and it is difficult to preserve the edge contour of the target regions accurately. On the other hand, although clustering-based image segmentation obtains better segmentation results, it is also difficult to be performed on online image segmentation and detection due to the high computational complexity. To solve the problems mentioned above, a fast image segmentation algorithm based on salient features model and spatial-frequency domain adaptive kernel is proposed. The organization of this paper is as follows. In Section 2, the salient feature model of the image is analyzed and established in spatial-frequency domain, and the design methods of the frequency domain bandpass filter and the spatial domain adaptive kernel are given. In Section 3, the structure design and flow of the algorithm are proposed. The results of comparative experiments with other algorithms are presented and discussed in section 4, and the conclusions are summarized in Section 5.

2 Establishing the Salient Feature Model in Spatial-frequency Domain

Currently, salient object detection (SOD) is a common feature extraction method [20] in image processing. It mainly aims to combine the visual characteristics of human beings, therefore the target region is highlighted from the cluttered background to attract visual attention. As we know, there are five basic requirements that should be met in the detection of salient objects [21]:

i. Emphasizing the largest salient object.

ii. Uniformly highlight whole salient regions.

iii. Establishing well-defined boundaries of salient objects.

iv. Disregard high frequencies arising from texture, noise and blocking artifacts.

v. Efficiently output full resolution saliency maps.

On this basis, the proposed algorithm combines the salient features of the image in the spatial frequency domain to improve the performance of salient object detection. By the filtering in frequency domain, the salient object is highlighted stably and efficiently, the background texture information is suppressed, and the high frequency interference is overcome. By the convolution with adapted kernel in spatial domain, the target edge and texture features are extracted completely, and the contour information of target is reconstructed.

2.1 Salient Region Enhancement in Frequency Domain

As shown in Fig.1, in the Fourier spectrum, the frequency component $F(0,0)$ at the origin of the spectrum is the DC component of the image, and its amplitude $\left|F\left(0,0\right)\right|$ reflects the average gray level of the image. The low-frequency components near the origin of the spectrum correspond to the components with slow gray level changes in the image, such as the large background with stable gray level and the internal area of glue. The high-frequency components are far from the origin of the spectrum corresponding to the components with sharp gray level varying in the image,

such as the step edge, background texture and high-frequency noise. The dominant feature of the Fourier spectrum of the glue dispensed sample is several distinctive components extending approximately ±45°, which corresponds to brightly colored glue outlines with a certain width.

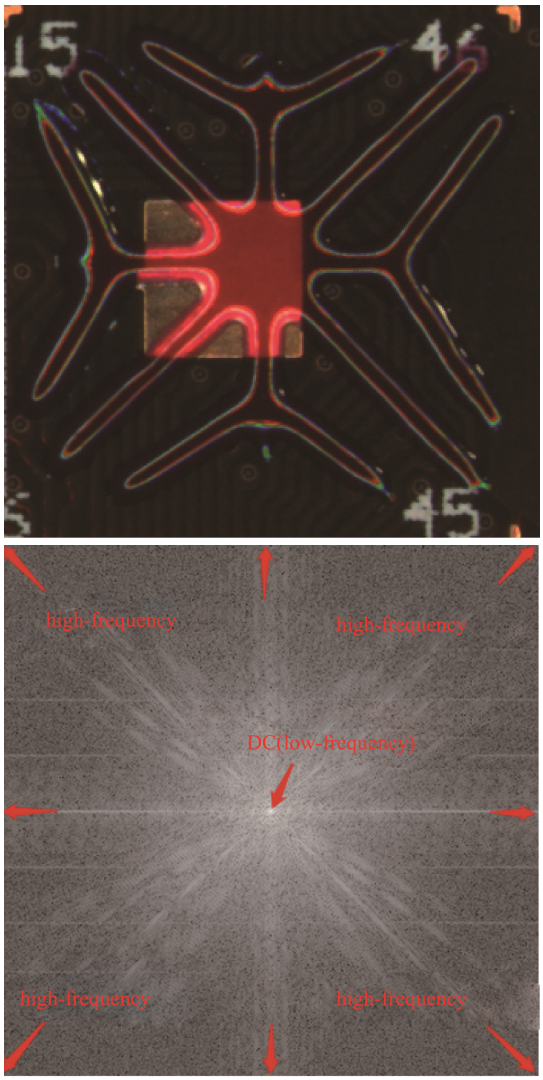


Fig.1 The Sample Image of Glue Dispensing and Its Fourier Spectrum

Based on the above analysis of salient objects detection, the designed band-pass filter should meet the following conditions:

1) The low cut-off frequency ω_{lc} should be as low as possible to highlight the relatively large salient region uniformly (salient object detection requirements i & ii).

2) There should be enough frequency margin between the low cut-off frequency ω_{lc} and the high cut-off frequency ω_{hc} to preserve the high-frequency components such as the edge contour in the image (salient object detection requirement iii).

3) The high cut-off frequency ω_{hc} should be less than the maximum frequency in the spectrum to filter the high-frequency noise and texture in the image (salient object detection requirement iv).

Fourier transform does not change the mathematical model of the Gaussian function, so the first-order Gaussian function is considered as the basic mathematical model of the filter to eliminate the ringing phenomenon of frequency filtering. Difference of Gaussians (DoG) is used to be the mathematical model of band-pass filter in the proposed algorithm. The DoG filter is the difference result of two Gaussian functions at different scales. Its spatial expression is defined as:

$$\begin{aligned} DOG(x,y) &= \frac{1}{2\pi\sigma_1^2} exp.(-\frac{x^2+y^2}{2\sigma_2^2}) - \frac{1}{2\pi\sigma_2^2} exp.(-\frac{x^2+y^2}{2\sigma_2^2}) \\ &= g_1(x,y,\sigma_1) - g_2(x,y,\sigma_2) \end{aligned} \tag{1}$$

where x and y are pixel coordinates of the image, and are the standard deviations of the Gaussian ($\sigma_1 > \sigma_2$).

From the properties of the Fourier transform, it can be seen that the Gaussian function with the larger spatial-domain standard deviation has a smaller frequency-domain cutoff frequency. Therefore, ω_{lc} and ω_{hc} of the band-pass filter depend on σ_1 and σ_2 , and the bandwidth depends on the ratio of σ_1 to σ_2 . In order to highlight the salient object in the image, σ_1 should be set to infinity, that is, ω_{lc} should be infinitesimal. Moreover, there should be a sufficiently large frequency width between ω_{lc} and ω_{hc} to preserve the frequency information of the salient objects. In this case, the DC component of the image ω = 0 will be filtered, and the average gray level of the image is zero, so the background gray level will be reduced to zero. In

addition, the value of ω_{hc} should be able to filter out high frequency noise and texture.

Considering the reality of filtering in frequency domain, the band-pass filtering with infinitesimal ω_{lc} cannot be realized and it is necessary to improve the Gaussian band-pass filter. The proposed algorithm sets the value of the component at the origin of the low-pass spectrum to zero, and filters the DC component of the image spectrum to realize the band-pass filtering with infinitesimal ω_{lc}, even if there is a slight ringing phenomenon. The expression of the improved Gaussian low-pass filter in frequency domain can be defined as:

$$G_{lp}(u,v)=\begin{cases}2\sigma^2 exp.\left[-2\pi^2\sigma^2\left(u^2+v^2\right)\right], & u,v\neq 0\\ 0 \quad , & u,v=0\end{cases} \tag{2}$$

where σ is the standard deviation of Gaussian low-pass function in spatial domain, u and v are coordinates in frequency domain. High cut-off frequency of the improved Gaussian low-pass filter is defined as:

$$\omega_{hc}=\frac{1}{2\pi\sigma} \tag{3}$$

In order to preserve the frequency information of the salient objects while filtering out high-frequency noise and texture, the high cut-off frequency ω_{hc} should be greater than the highest frequency m of the corresponding frequency band of the salient objects, and a certain frequency margin should be preserved. The improved Gaussian low-pass filter spectrum is shown in Fig.2.

In order to reduce the complexity of the algorithm, the specific filtering operation is usually implemented in the spatial domain after the frequency domain filter being determined, therefore the inverse discrete Fourier transform (IDFT) is eliminated. By calculating the difference between the image after Gaussian low-pass filtering and the arithmetic mean image, the high-frequency noise, texture and DC components are filtered, and the frequency band information of the salient regions are preserved, so that approximately replaces the above band-pass filtering, which can be expressed as:

$$s(x,y)=\left|f_{\omega_{hc}}(x,y)-f_{\mu}(x,y)\right| \tag{4}$$

where $s(x,y)$ is the output of salient object detection, $f_{\omega_{hc}}(x,y)$ is the image after Gaussian low-pass filtering, $f_{\mu}(x,y)$ is the arithmetic mean image, of which the size is the same as the original image, and the gray level of the pixels are the arithmetic mean of the gray level of the whole image.

Then, the standard deviation of the spatial domain Gaussian low-pass filter should be:

$$\sigma=\frac{1}{2\pi\omega_{hc}}\leqslant\frac{1}{2\pi\omega_{m}} \tag{5}$$

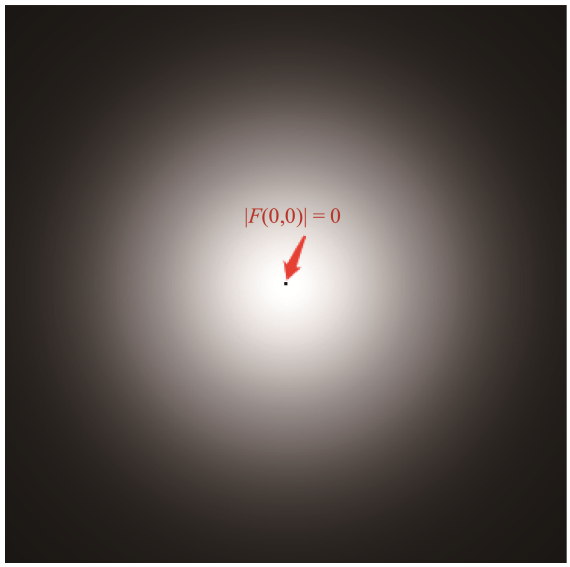
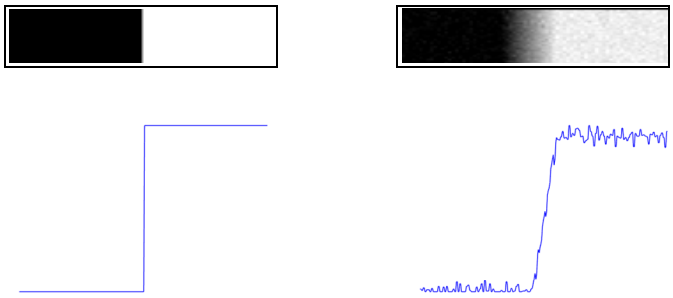


Fig.2 The Fourier Spectrum of Improved Gaussian Low-pass Filter

2.2 Salient Feature Extraction in Spatial Domain

Now, the saliency of the target regions is enhanced after filtering in frequency domain. Based on this salient enhanced image, edge detection in spatial domain will be used to extract the salient edge features in this section.

In actual digital images, the real edge model is not an ideal "step", but a fuzzy and noisy "slope" (as shown in Fig.3). The ambiguity of the edge depends on the limitation of the optical imaging mechanism, and the noise level is mainly affected by the illumination level and the electronic components of the imaging system.



(a) The Ideal Edge Model (b) The Actual Edge Model

Fig.3 The ideal Edge Model and the Actual Edge Model

Traditional edge detectors, such as Sobel and Laplacian, are simple and efficient. However, when the edge of the image is not obvious, the texture is intricate and the noise pattern cannot be measured, it is difficult for traditional edge detectors to achieve a high comprehensive evaluation in terms of edge detection and anti-noise interference [22]. In response to the accuracy and rapidity requirements of online detection for edge detection, an edge detection method based on adaptive convolution kernel will be proposed in this manuscript. The eight-neighborhood square kernel is set as an example.

$f(i-1,j-1)^2$	$f(i,j-1)^2$	$f(i+1,j-1)^2$
$f(i-1,j)^2$	$f(i,j)^2$	$f(i+1,j)^2$
$f(i-1,j+1)^2$	$f(i,j+1)^2$	$f(i+1,j+1)^2$

Fig.4 The Eight-neighborhood Square Kernel

By adjusting the operator parameters, the algorithm can achieve excellent results in edge highlighting and localization, denoising and execution efficiency.

The convolution operation of the square kernel is performed on the image to enhance the distinction between the texture area and the smooth area. Then, the square of pixel gray mean in the neighborhood is regarded as the adaptive segmentation parameter to extract the contour information of interest, which are expressed by Equations (6)-(8).

$$Q(K)=\begin{cases} True, & M(x,y)\geq K*Avg^2(x,y) \\ False, & Others \end{cases} \quad (6)$$

$$M(x,y)=\sum_{\substack{x-1\leq x'\leq x+1\\ y-1\leq y'\leq y+1}} f(x',y')^2 \quad (7)$$

$$Avg(x,y)=\frac{1}{9}\sum_{\substack{x-1\leq x'\leq x+1\\ y-1\leq y'\leq y+1}} f(x',y') \quad (8)$$

Where Q is a predicate logic that represents the authenticity of the pixel as the target contour feature. $M(x,y)$ is the result of the square kernel convolution at the pixel. $Avg^2(x,y)$ is the square of the average gray level of pixels in the neighborhood, and K is the threshold parameter. If changing the parameter K, then Q will extract different contour information. In the texture area, the values of $M(x,y)$ and $Avg^2(x,y)$ are quite different. While in smooth areas or isolated noise areas, $M(x,y)$ and $Avg^2(x,y)$ are similar in value. According to the contour model of the segmentation target, it can select the appropriate value of K to achieve a better edge detection result $Q(K)$. Finally,

according to the value of $Q(K)$ at each pixel, an image $E(x,y)$ containing the salient contour information of the target is obtained by:

$$E(x,y)=\begin{cases}1, & \textit{The value of } Q(K) \textit{ is true}\\ 0, & \textit{The value of } Q(K) \textit{ is false}\end{cases} \quad (9)$$

To detect complex image edge texture and high noise level, the adaptive convolution kernel will have the following advantages:

1) To square the adaptive convolution kernel in each pixel with the neighborhood, it will amplify the difference between the edge part and the smooth area in the image, and it will have a better effect on highlighting edge features.

2) Generally, a global thresholding is performed after edge detection to segment salient edge features. However, it is difficult to obtain the ideal effect by adjusting the detection parameters and the segmentation threshold when the gray level of the image is fluctuating due to uneven illumination. The average gray level of pixels in the neighborhood is regarded as the detection parameter to overcome the problem of local gray level instability caused by noise and uneven illumination, which improves the adaptability of edge detection.

3) The adaptive convolution kernel is flexible and universal for different edge models and detection requirements. Using different parameters *K*, kernel shape and size are adopted to detect the edge information of specific features (such as edge width, shape and contrast).

Based on the flexibility and universality of the adaptive convolution kernel, detection parameters should be set for different edge models and detection requirements to obtain better detection results. For these reasons, a higher parameter *K* should be set to detect salient edges while reducing noise. To detect the edge information in a specific direction, such as the edge in the vertical direction, the horizontal kernel can be used. On the other hand, to detect the contour of the subject area, and ignore the detailed texture, a larger kernel size should be adopted, and a higher parameter *K* should be set at the same time.

3 Algorithm Structure Design and Process

The flow chart of the proposed algorithm is presented in Fig.5.

Firstly, the salient objects of the original image are detected, and the saliency image $s(x,y)$ is obtained by band-pass filtering in frequency domain. In this process, the high-frequency noise is filtered out, and the gray level of the background is reduced, so that the contrast between the segmentation target and the background is improved. Then, the convolution based on the spatial-domain adaptive kernel is performed on $s(x,y)$, and the salient contour image $E(x,y)$ is obtained. Furthermore, by analyzing the grayscale distribution of $s(x,y)$ and the approximate contour *E*, the target grayscale feature *I* is obtained. Finally, the targets are segmented through the seeded region growing algorithm. The parameters such as seeds position, grayscale index and adjacency mode in the seeded region growing algorithm are determined by the target salient contour *E* and grayscale feature *I*.

The specific process of the fast image segmentation algorithm based on salient features model and spatial-frequency domain adaptive kernel is summarize as follows:

Step 1. **Salient Target Detection in Frequency Domain**

A. Salient feature analysis in frequency domain:

Discrete Fourier Transform (DFT) of the image to obtain the spectrum. Then the center of the spectrogram is regarded as the center of the circle to construct a circular area with a radius of *R*, Adjusting the size of *R*, the frequency band information of the salient area is all located in the circular domain of $R \geq \omega_m$.

To improve the robustness of the algorithm, the sum of the amplitudes of the frequency components in the circular domain is usually greater than 95% of the full spectrum components, which means that the frequency bands in the circular domain contains salient objects, and the high-frequency noise components are outside the circular domain.

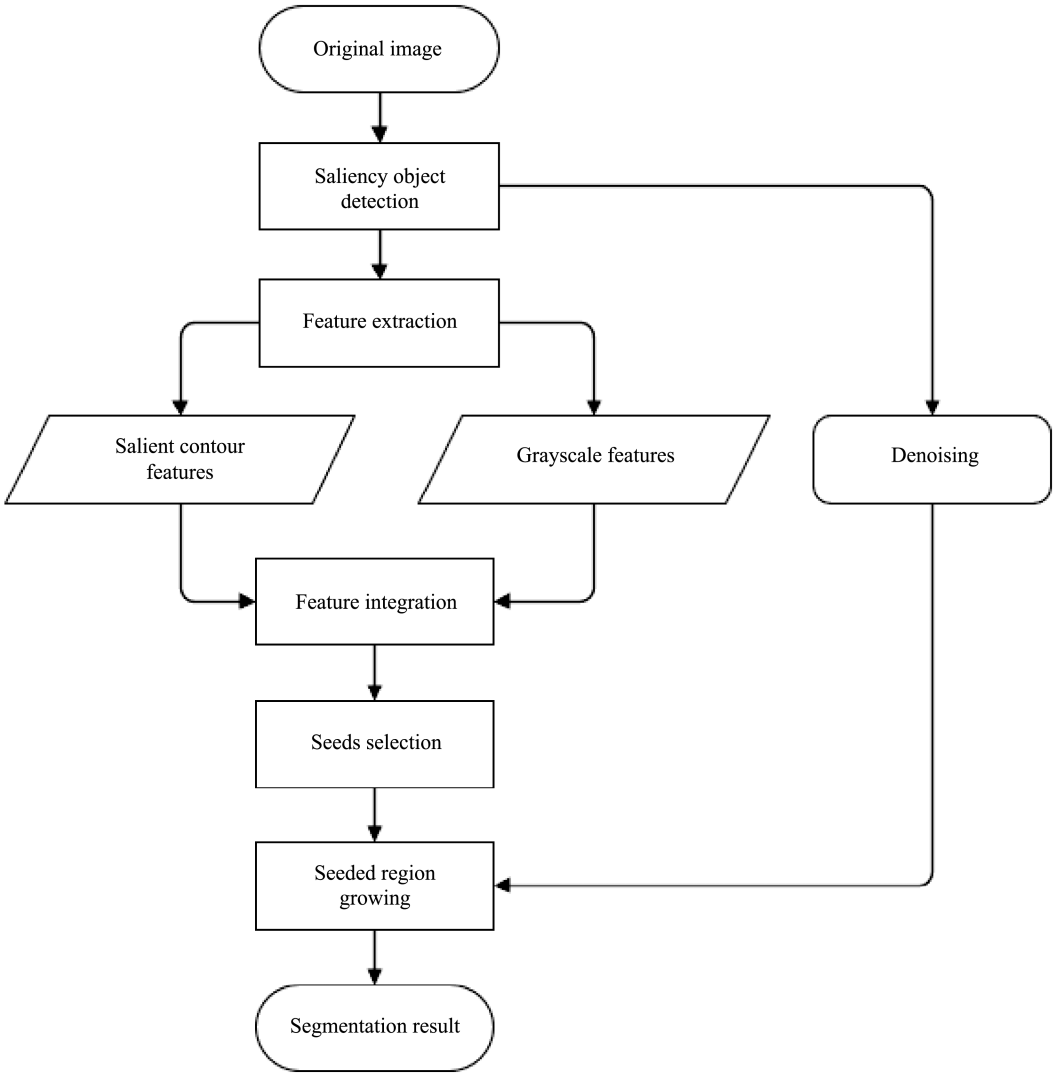


Fig.5 The Flow Chart of the Proposed Algorithm

B. Improvement of the Gaussian low-pass filter:

The radius value *R* of the circle domain is considered as the value of the high cut-off frequency ω_{hc} to construct a frequency domain Gaussian low-pass filter, and then the origin component of the low-pass filter is set to zero.

C. Gaussian low-pass filtering in spatial domain:

According to Equations (5), the standard deviation σ of Gaussian low-pass filtering in spatial domain is calculated, and the original image is Gaussian low-pass filtering to obtain the low-pass filtered image $f_{\omega_{hc}}(x,y)$. In order to improve the efficiency of image filtering, the two-dimensional convolution is replaced by a combination of one-dimensional convolution along the rows and columns. For example, to realize a Gaussian low-pass filter with $\sigma=\sqrt{4/(3\pi)}$, it is feasible to design $\frac{1}{16}[1,4,6,4,1]$ and $\frac{1}{16}[1,4,6,4,1]^{\mathrm{T}}$ convolution kernels of one-dimensional Gaussian filter for image convolution.

D. Image difference:

The arithmetic mean f_{μ} of the original image element is calculated, and the arithmetic mean image $f_{\mu}(x,y)$ with the same size as the original image is

obtained. According to Equations (5), the corresponding pixel gray levels are subtracted for $f_\mu(x,y)$ and $f_{\omega_{hc}}(x,y)$, then the absolute value of the results is taken to obtain the saliency image $s(x,y)$ of the salient objects.

Step 2. **Salient Feature Extraction in Spatial Domain**

A. Convolution operation with adaptive kernel:

According to the texture features of interest in $s(x,y)$, the shape, size and scale factor *K* of the adaptive kernel are designed, which is described in Section 2.2. For targets with different edge model, the value of *K* can be fine-tuned to obtain salient and clear edge feature, and its value is usually set to 9 to 10. Then, convolution is performed on $s(x,y)$, as shown in Equations (6) ~ (8), and the salient edge feature $Q(K)$ is obtained.

B. Denoising:

The Blob detection of $Q(K)$ is performed, and the Blob area is classified according to the size of Blob area. In this process, the gray level of Blob with small area is set to zero, so as to remove small area noise.

C. Salient contours extraction of the target:

Edges of the denoised image are refined, and then all breakpoints in the refined edges are found. The edge connection between the breakpoints with the smallest Euclidean distance is carried out to obtain the approximate contour image $E(x,y)$.

D. Connected region detection:

The inner contour and the outer contour are tracked in contour image $E(x,y)$ and each contour is marked with a corresponding label. Pixels in $E(x,y)$ are scanned and corresponding region label R_i is given to the pixels between the inner and outer contours.

E. Grayscale features extraction of the target:

In the saliency image $s(x,y)$, the average gray level of the pixels in the corresponding area R_i is calculated, and the result is used as the target gray feature I_i:

$$I_i = \frac{1}{N_{R_i}} \sum_{(x,y)\in R_i} s(x,y) \tag{10}$$

where N_{R_i} is the number of pixels contained in the area R_i, *i.e.*, the area of the region R_i.

Finally, the connected regions are classified. Generally, the region R_i with average gray level less than 20 in the image $s(x,y)$ can be defined as the background region B, and the remaining regions are defined as the salient target region S_i.

Step 3. **Segmentation of Target Area**

A. Selection of seed points:

Firstly, slight fuzzy denoising is carried out on the salient enhancement image $s(x,y)$, and then all the pixels in the blurred image are traversed to select the corresponding seed points. The proposed algorithm selects pixels that meet the following two requirements as seed points for seeded region growing:

1) The seed points are located inside the salient region S_i.

2) The gray level of seed points is the average gray level I_i of target region S_i.

B. Seeded region growing:

Centered on each seed point (x_i, y_i), the algorithm determines whether the pixel point (x_k, y_k) in the eight-neighborhoods of the seed point meets the growth criterion Q_S. The grayscale difference between the seed point and the inspected point is considered as the criterion in the proposed algorithm, which is expressed by Equation (11).

$$Q_S = \begin{cases} coincident, & I(x_i,y_i) - I(x_k,y_k) \le M \quad k = 1,2,...,8 \\ non-coincident, & else \end{cases} \tag{11}$$

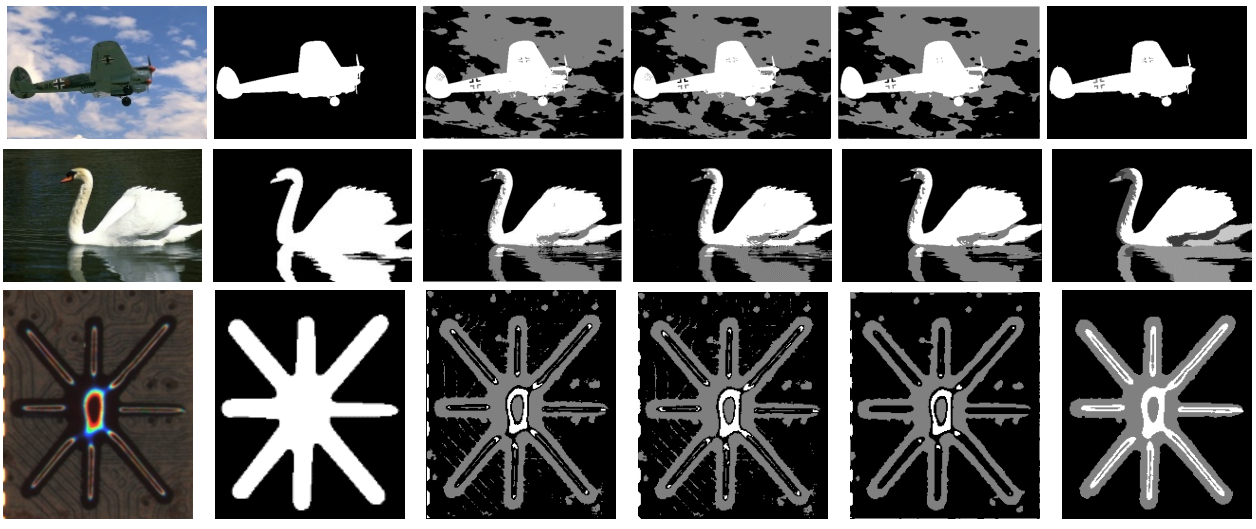
where *i* is the number of seed points, *k* is the serial number of neighborhood points, *M* is the grayscale difference threshold. In general, for high contrast images, the grayscale difference is less than 20, while for low contrast images, the grayscale difference is less than 10, which meets the criteria. If the point

meets the criteria, it is classified as a seed point (foreground pixel), otherwise it is classified as a background pixel. Then continue to evaluate the neighboring pixels of other seed points, iterative this process until each pixel is assigned. After seeded region growing, the pixel set classified as seed points is considered as segmentation target, and the remaining pixel sets are the background.

4 Experimental Results and Analysis

In order to verify the effectiveness of the proposed algorithm, the proposed algorithm is applied to the test image of Berkeley Segmentation Data Set and the sample image of glue defect detection, then the results are compared with current segmentation algorithms of OTSU[3], K-Means[23] and FLICM [24]. All the following experiments are operating on a PC (I5-10400F at 2.90GHz, 16GB of Ram, single NVIDIA GTX1650S GPU) with Win10, Visual Studio 2017, and OpenCV 3.4.0 framework.

The image sizes of the three samples are 481×321, 481×321, and 200×240, respectively. Fig.6(b) shows the result of manual segmentation, which serves as an ideal segmentation for reference. In order to segment the salient objects from the background area while preserving the salient texture in the detected target, the number of initial clusters in the OTSU, K-Means and FLICM algorithms compared in this experiment are all set to three. As shown in Fig.6(c), (d), (e), the three algorithms can preserve the edge and texture information of the salient objects completely, but some feature regions inside the target are incorrectly assigned into the background area, such as the logo on the airplane fuselage, part of the aircraft nose area, part of the dark area of the swan head and the reflective areas of the glue, respectively. As for the segmentation of the first image, three algorithms all fail to suppress the non-salient texture in background area. For the second image, there is still some noise in the background area after segmentation by OTSU and K-Means algorithms. The third image is a sample picture for defect detection in the glue dispensing process. In this image, the contrast between the segmented target area and the background area is not obvious. Therefore, the segmentation of the OTSU and K-Means on the edge of the glue is not ideal, and burrs occur obviously on the edge of the glue after segmentation due to the interference of the background texture.



(a) (b) (c) (d) (e) (f)

Fig.6 Segmentation Results Comparison of Different Algorithms. (a) Original Image, (b) Ground True, (c) OTSU, (d) K-Means, (e) FLICM, (f) Proposed

As for the segmentation of the proposed algorithm, in order to detect significant targets in the frequency domain, the filtering method proposed in Step 0 of the Section 3 are adopted in frequency domain filtering. In the feature extraction of spatial domain, the eight-neighborhood kernel is used. For the three images mentioned above, to obtain clear edge information, 9.3, 10 and 9.5 are evaluated and selected as the scale coefficient K of contour recognition. In order to accurately segment the target area, 20, 20 and 10 are tested and selected as the threshold M of grayscale difference in seeded region growing. Considering the segmentation performance of the target of interest shown in Fig.6, the proposed algorithm can completely preserve the edge of the salient object, and can distinguish the internal texture of the target from the background texture, and it also can weaken the background texture and noise in the segmentation better.

As shown in Table 1, *Jaccard Similarity Coefficient* (JSC), *Hausdorff Distance* (HD) [25] and the execution time are adopted to evaluate the algorithms and quantify their accuracy.

The area overlap ratio of ideal segmentation and algorithm segmentation is measured by *Jaccard Similarity Coefficient*:

$$JSC = \frac{|GT \cap ST|}{|GT \cup ST|} \tag{12}$$

where GT is the target area of ideal segmentation, while ST is the target area of actual segmentation from the algorithm.

Evaluating the segmentation indicators of the first two pictures, the four segmentations can obtain a high degree of area overlap, and *Jaccard Similarity Coefficient* is above 0.87. However, regarding low-contrast image of glue dispensing process samples, the OTSU, K-Means, and FLICM only consider the grayscale feature of the pixels, resulting in the incorrect segmentation of the internal region of the target, and a low *Jaccard Similarity Coefficient*. While the proposed algorithm improves the seeded region growing algorithm based on grayscale feature by considering the prior target contour in segmentation. Compared with the other three algorithms, the accuracy of the target area segmentation of the proposed algorithm is obviously improved, and the *Jaccard Similarity Coefficient* is above 0.90, which further verifies the advantage of the proposed algorithm.

The coincidence of the target boundary between the ideal segmentation and the algorithm segmentation is measured by the *Hausdorff Distance* expressed by Equations (13)-(15):

$$h(G,S) = \max_{g_i \in G}\left\{\min_{s_j \in S}\left\|g_i - s_j\right\|\right\} \tag{13}$$

$$h(S,G) = \max_{s_i \in S}\left\{\min_{g_j \in G}\left\|s_i - g_j\right\|\right\} \tag{14}$$

$$HD(G,S) = \max\left\{h(G,S), h(S,G)\right\} \tag{15}$$

Table 1 Performances Evaluation of Different Algorithms

Algorithms	Plane			Goose			Glue		
	JSC	HD	Time(ms)	JSC	HD	Time(ms)	JSC	HD	Time(ms)
OTSU	0.927	13.000	1891.952	0.874	13.591	1617.122	0.660	6.394	1829.772
K-Means	0.934	3.597	922.892	0.871	12.394	932.552	0.652	7.797	509.350
FLICM	0.896	3.597	40115.000	0.879	11.997	47571.200	0.591	4.000	8994.630
Proposed Method	0.971	3.597	354.168	0.924	8.000	413.471	0.962	1.000	233.402

where G is the boundary point set of the target region in ideal segmentation, S is the boundary point set of target region in the algorithm segmentation, g and s are the points on point set G and point set S, $h(G,S)$ and $h(S,G)$ is the one-way *Hausdorff Distance* from point set G to point set S and from point set S to point set G, $HD(G,S)$ is the *Hausdorff Distance* between point set G and point set S.

OTSU is based on the maximum variance between classes and K-Means is based on the feature distance from the pixel point to each cluster center. Since the local characteristics of the edge are not considered, the two algorithms cannot preserve the target edge contour stably. FLICM considers the local spatial information of the pixels, and the coincidence of the target boundary is improved. In contract, the proposed algorithm combines the contour features of the target region in the segmentation, which is superior to the other three algorithms in preserving the boundary information of the target regions. Inferred from the segmentation performance, the proposed algorithm is feasible.

In terms of the execution time, multiple iterative calculations are performed in OTSU, K-Means and FLICM, which are used to find clustering centers in the entire graph, so the algorithms are time-consuming. The proposed algorithm pre-segments the salient regions, then it calculates the characteristic parameters of each salient regions, and finally performs the iterative calculation of segmentation. Thus, the efficiency of the proposed algorithm is significantly higher than other three algorithms. In summary, the above experimental results and indicators show that the proposed algorithm is effective and adaptive.

6 Conclusion

In this paper, a fast image segmentation algorithm based on salient features model and spatial-frequency domain adaptive kernel is proposed. Combining the characteristics of the image in spatial and frequency domain, the salient contour and grayscale feature of the regions of interest are extracted, and the target region is segmented accurately and quickly. The major advantages of the proposed algorithms are summarized as follows:

1) The saliency model of the image is intuitively analyzed from the Fourier spectrum, which can stably detect the salient target area.

2) As for salient feature extraction of the target, adaptive convolution kernel is flexible and universal for different edge models and detection requirements.

3) The accuracy and efficiency of image segmentation are improved by combining target grayscale feature and contour feature.

The experimental results show that the proposed segmentation algorithm can accurately segment the target area, and it preserve the target contour and salient texture more completely. Moreover, the proposed algorithm executes efficiently is obvious, which is especially suitable for online image detection with high requirements for segmentation efficiency.

Acknowledgment

This work was supported by National Natural Science Foundation of China [grant numbers 61573233]; Natural Science Foundation of Guangdong, China [grant numbers 2021A1515010661]; Special projects in key fields of colleges and universities in Guangdong Province [grant numbers 2020ZDZX2005]; Innovation Team Project of University in Guangdong Province [grant numbers 2015KCXTD018]. The authors also would like to thank Dongguan Attach Point Intelligent Equipment Co. Ltd, China, for providing with experimental facilities and testing assistance.

References

[1] Pare S, Kumar A, Singh G K, et al. Image Segmentation Using Multilevel Thresholding: A Research Review[J]. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 2020, 44(1):1-29.

[2] Shrivastava N, Bharti J. Automatic Seeded Region Growing Image Segmentation for Medical Image Segmentation: A Brief Review[J]. International Journal of Image and Graphics, 2020, 20(3): 2050018.

[3] Kuze N M H. (2010). Determination of Tropospheric

Aerosol Characteristics by Spectral Measurements of Solar Radiation Using a Compact, Stand-alone Spectroradiometer, *Applied Optics*, 49(8), pp.1446-1458.

[4] Pei Yang, Wei Song, Xiaobing Zhao, Rui Zheng, Letu Qingge. An improved Otsu threshold segmentation algorithm[J]. International Journal of Computational Science and Engineering,2020,22(1): 146-153.

[5] Adapa D, Raj A N J, Alisetti S N, et al. A supervised blood vessel segmentation technique for digital Fundus images using Zernike Moment based features[J]. PLoS ONE, 2020, 15(3): e0229831.

[6] D. Zhao, L. Liu, F. Yu et al., Chaotic random spare ant colony optimization for multi-threshold image segmentation of 2D Kapur entropy, Knowledge-Based Systems (2020), doi: https://doi.org/10.1016/j.knosys.2020.106510.

[7] Ahror Belaid, Djamal Boukerroui. Local maximum likelihood segmentation of echocardiographic images with Rayleigh distribution[J]. Signal, Image and Video Processing, 2018, 12(6): 1087-1096.

[8] Kang C C, Wang W J, Kang C H. Image segmentation with complicated background by using seeded region growing[J]. AEUE - International Journal of Electronics and Communications, 2012, 66(9):767-771.

[9] Li Q, Wei Z, Zhao C. Optimized Automatic Seeded Region Growing Algorithm with Application to ROI Extraction[J]. International Journal of Image and Graphics, 2017, 17(04):1750024.

[10] Y. Liu, Z. Xie and H. Liu. An Adaptive and Robust Edge Detection Method Based on Edge Proportion Statistics[J]. IEEE Transactions on Image Processing[J], 2020, 29: 5206-5215.

[11] P Arbeláez, Maire M, Fowlkes C, et al. Contour Detection and Hierarchical Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.

[12] Amit Saxena, Mukesh Prasad, Akshansh Gupta,Neha Bharill, Om Prakash Patel, Aruna Tiwari, Meng Joo Er, Weiping Ding, Chin-Teng Lin. A review of clustering techniques and developments[J]. Neurocomputing, 2017, 267: 664-681.

[13] Bong C W, Rajeswari M. Multi-objective nature-inspired clustering and classification techniques for image segmentation[J]. Applied Soft Computing, 2011, 11(4): 3271-3282.

[14] Felzenszwalb P F, Huttenlocher D P. Efficient Graph-Based Image Segmentation[J]. International Journal of Computer Vision, 2004, 59(2):167-181.

[15] Achanta R, Shaji A, Smith K et al. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(11): 2274-2282.

[16] David Stutz, Alexander Hermans, Bastian Leibe. Superpixels: An evaluation of the state-of-the-art[J]. Computer Vision and Image Understanding, 2018, 166: 1-27.

[17] Soomro N Q, Wang M. Superpixel Segmentation: A Benchmark[J]. Signal Processing Image Communication, 2017, 56: 28-39.

[18] Kashif Hussain Memon, Dong-Ho Lee. Generalised kernel weighted fuzzy C-means clustering algorithm with local information[J]. Fuzzy Sets and Systems, 2018, 340: 91-108.

[19] Chisti, Md. Khwaja Muinuddin, S. Srinivas Kumar and G Prasad. Defects Identification, Localization, and Classification Approaches: A Review[J]. IETE Journal of Research, 2021: 1-14.

[20] Liu Z, Qu B. Machine vision based Online detection of PCB Defect[J]. Microprocessors and Microsystems, 2021, 82(9):103807.

[21] Ali, Borji, Ming-Ming, et al. Salient object detection: A survey[J]. Computational Visual Media, 2019, v.5(02): 3-36.

[22] Zhang Y, Mao Z, Li J, et al. Salient region detection for complex background images using integrated features[J]. Information Sciences, 2014, 281: 586-600.

[23] Baptiste Magnier. Edge detection: a review of dissimilarity evaluations and a proposed normalized measure[J]. Multimedia Tools and Applications, 2018, 77(8): 9489-9533.

[24] Rahul Ratnakumar, Satyasai Jagannath Nanda. A low complexity hardware architecture of K-Means algorithm for real-time satellite image segmentation[J]. Multimedia Tools and Applications, 2019, 78(9): 11949-11981.

[25] Krinidis, Stelios, Chatzis, et al. A Robust Fuzzy Local

Information C-Means Clustering Algorithm[J]. IEEE Transactions on Image Processing, 2010, 19(5):1328.

[26] Zhaobin Wang, E. Wang, Ying Zhu. Image segmentation evaluation: a survey of methods[J]. Artificial Intelligence Review, 2020, 53(8): 5637-5674.

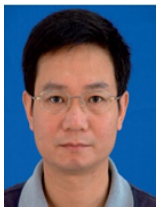
Author Biographies

WU Fupei received the Ph.D. in mechanical engineering from South China University of Technology, Guangzhou, China, in 2009. He is now an associate professor in Department of Mechanical Engineering, Shantou University, Shantou, China. His main research interests include Automatic optical inspection and 3D measurement

E-mail: fpwu@stu.edu.cn

LIANG Jiaye received a B.Sc. degree from Guangzhou University in 2018. He is now a M.Sc. candidate in mechanical engineering at Shantou University. His main research interests include machine vision and image processing.

E-mail: 18825081623@163.com



LI Shengping received a B.Sc. degree from Wuhan University of Technology in 1987, a M.Sc. degree from Beijing University of Technology in 1992 and a Ph.D. degree from Huazhong University of Science and Technology in 1995. He is now a professor at Shantou University. His main research interests include adaptive control, robust control, robust design theory and application, machine vision.

E-mail: spli@stu.edu.cn