

RESEARCH

Open Access



Saliency area detection algorithm of electronic information and image processing based on multi-sensor data fusion

Xinyu Zhang^{1,2} and Kai Ye^{2*}

*Correspondence:

kaiye_edu@163.com

² School of Automation
Science and Technology,
Xi'an Jiaotong University,
Xi'an 710049, Shaanxi, China
Full list of author information
is available at the end of the
article

Abstract

Researched in the 1980s, multi-sensor data convergence has become a hot issue. Not only does it differ from general signal processing, or single to multiple sensor surveillance and measurement, on the other hand, it is a higher level of integrated decision-making processes based on multiple sensor measurement outcomes, this paper is based on the study of the saliency area detection algorithm of electronic information and image processing based on multi-sensor data fusion, based on the improved FT algorithm and LC algorithm using multi-sensor data fusion technology, a new LVA algorithm is proposed, and these three algorithms are evaluated in an all-round way through various algorithm evaluation indicators such as PR curve, PRF histogram, MAE index, and recognition image rate. The research results show that the LVA algorithm proposed in this paper improves the detection rate of saliency maps by 5–10%.

Keywords: Multi-sensor data fusion, Electronic information, Image processing, Saliency detection

1 Introduction

1.1 Background

The definition of sensor data fusion can be summarized as the integration of local data resources provided by multiple sensors of the same or different types distributed in different locations, and the use of computer technology to analyze them to eliminate the possible redundancy and Contradiction, complement each other, reduce its uncertainty. The detection of salient areas of images is of great significance in computer vision, psychology, biology, and other fields. In real scenes, image salient area detection also has good application prospects. Through the continuous efforts of researchers, significant area detection has yielded rich results. However, the salient area detection still faces some difficulties. When the target area of the image is very complicated, it is difficult for the commonly used detection algorithms to detect the complete salient area. When the background of the picture is very complicated, the result of the salient area detection will contain a lot of noise. The generalization performance of the algorithm is not good enough. These are the research directions of image salient area detection. With the

advancement of technology, pictures and videos have become the most important information carriers. Today, more than 70% of the data on the Internet is image or video data. About 500 million pictures are uploaded to the Internet every day. On average, 1.3 billion hours of videos are uploaded to the YouTube website every minute. By 2022, the total number of global cameras will reach 44 trillion. While people's lives are greatly enriched, people's requirements for computer vision are constantly increasing. Researchers not only hope that computers can replace humans to complete those mechanical and cumbersome information processing labor, but also hope that computers can process visual information more intelligently, accurately, and efficiently. At this time, it becomes more and more important to use various methods to detect salient areas in videos and images.

1.2 Significance

Image salient area detection is also called saliency detection. Its main purpose is to allow the computer to simulate human physiological reactions, analyze and calculate the image, and obtain the most intuitional and most noticeable part of the image. For humans, the information in the image is mainly concentrated in the salient area of the image, that is, in the foreground of salient area detection. If accurate salient area detection can be achieved, it is possible to directly operate on salient areas in the image, so that computing resources can be saved, the processing speed can be accelerated, and the processing efficiency of the image can be greatly improved. Therefore, image salient area detection is a basic task and a hot task in the field of computer vision. The salient area detection can extract the main content of the picture while conforming to human intuition. In large-scale image processing, salient area detection can improve processing efficiency and reduce the demand for computing resources. The salient area detection has important significance in computer vision.

1.3 Related work

In industrial processes, various sensors are increasingly used to measure and control processes, machines, and logistics. One way to process the results of large amounts of data generated by hundreds of different sensors in an application is to use an information fusion system. Information fusion systems, for example, for status monitoring, please combine different information sources (such as sensors) to generate the status of a complex system. The result of this information fusion process is regarded as an indicator of the health of a complex system. Therefore, the information fusion method is applied, for example, to automatically notify the reduction of production quality or detect possible dangerous situations. Considering the importance of sensors in the previously described information fusion systems and generally in industrial processes, defective sensors will bring several negative effects, which may lead to machine failures, for example, when machine wear cannot be detected sufficiently in advance. Jan Friedrich posited a method to detect faulty drivers by computing the agreement among sensor values. The proposed sensor defect discovery algorithm exemplarily uses the structure of a multilayer group-based factorization algorithm for sensor fusion. Desired results of the method for defects detection under different test cases are given in the paper, together with the ability of the proposed method to detect a variety of typical transducer faults [1]. However, they did not study issues such as saliency area detection based on this algorithm.

Yue L et al. solved the problem of target detection in a dynamic environment in a semi-supervised data-driven environment with low-cost passive sensors. The key challenge here is to achieve a higher probability of correct detection and a lower probability of false alarms at the same time under the constraints of limited computing and communication resources. Generally, due to limited training scenarios and assumptions about signal behavior in a static environment, changes in a dynamic environment may severely affect the performance of target detection. To this end, a binary hypothesis testing algorithm based on feature clustering extracted from multiple sensors that may observe the target is proposed. First, by using a recently reported feature extraction tool called symbol dynamic filtering, features are extracted from the time series signals of different sensors. Then, these features are grouped into clusters in the feature space to evaluate the uniformity of sensor response. Finally, a target detection decision is made based on the measurement results of the distance between the paired sensor clusters. The proposed procedure has been experimentally verified for moving target detection in a laboratory environment. In experiments, in the presence of varying ambient light intensity, multiple homogeneous infrared sensors with different directions have been used. Experimental results show that the proposed target detection program with feature-level sensor fusion is robust and superior to the target detection program with decision-level and data-level sensor fusion [2]. But their experiment is only about target monitoring of feature-level sensor data fusion, and only talks about image recognition technology, and there is no related end description of image processing technology. The track surface image captured by the line scan camera is susceptible to uneven illumination, stray light, changes in the smoothness of the track surface, etc., which will reduce the detection accuracy of the track surface peeling. In order to solve this problem, Hu Z et al. proposed a visual saliency-based detection algorithm for light track surface peeling. First, locate the rail surface area to eliminate interference from the surrounding area. Then, a two-dimensional differential Gaussian filter is used to reduce noise. The filtered image is processed by the block local contrast measurement estimator, which can enhance the contrast of the spallation area and generate a saliency map. Finally, a threshold is applied to locate the spalling area. The experimental results of researchers such as Hu Z show that the algorithm has a detection accuracy of 93.5% under uneven illumination and various track surface smoothness conditions and has good robustness [3]. But their research did not improve related algorithms based on multi-sensor data fusion, so their research still has some shortcomings.

1.4 Innovation

In this paper, a saliency detection algorithm combining multiple features and an improved method of FT algorithm are proposed. In this paper, the frequency weighting method is used to increase the weight of the low-frequency part of the image to realize the improvement of the FT algorithm. Aiming at the image scene of a single small salient object, this article makes full use of the advantages of frequency, color, location information, etc., in processing such images, combines the frequency-based improved FT algorithm and the contrast-based LC algorithm, and optimizes and enhances the location information for the post-processing of significant differences, an LVA algorithm for saliency detection combining multiple features is proposed.

2 Introduction to methods and related concepts

2.1 Multi-sensor data fusion

Human beings are a complex multi-sensor information fusion system, which is performing information fusion all the time. Multi-sensor information fusion is the use of multiple sensors to obtain relevant information, and perform data preprocessing, correlation, filtering, integration, and other operations to form a framework that can be used to make decisions, so as to achieve identification, tracking, and situation assessment. [4]. In summary, the multi-sensor data fusion system includes the following three parts:

1. Sensor. Sensors are the cornerstone of a sensor data fusion system. Without sensors, data cannot be obtained. Multiple sensors can obtain more comprehensive and reliable data.
2. Data. Data are the processing object in the multi-sensor data fusion system and the carrier of fusion. The quality of the data determines the upper limit of the fusion system performance, and the fusion algorithm only approaches this upper limit.
3. Fusion. Fusion is the core of a multi-sensor data fusion system. When the quality of the information cannot be changed, fusion is to mine the information to the greatest extent and make decisions based on the data.

Data fusion performs multi-level processing on multi-source data. Each level of processing abstracts the original data to a certain extent. It mainly includes data detection, calibration, correlation, and estimation [5]. Data fusion can be divided into three levels according to the degree of abstraction of data processing in the fusion system:

1. Pixel-level fusion. This method is currently the most widely used fusion method. Directly use the original image data, design algorithms to process and integrate the pixels one by one to achieve the purpose of image fusion. The data processed at the pixel level are all raw data, without any conversion, the information expressed is more accurate. However, because each pixel of the original data is processed, the amount of data becomes larger, and the fusion efficiency is low.
2. Feature level fusion. This method is based on the feature information of the image itself for fusion. The algorithm is used to extract feature information such as the edge and contour of the image, and then the information is fused. This method reduces the amount of data in the fusion process and improves the fusion efficiency, but it has higher requirements for feature extraction, and the quality of feature extraction directly affects the fusion effect [6, 7].
3. Decision-level integration. This type of method is the most complicated. It is necessary to prepare expert knowledge related to the image content before image processing, and then perform targeted adjustment and processing of the image. This fusion method is relatively abstract and requires higher expert knowledge, but due to its strong pertinence, the fusion effect is also more ideal [8, 9].

Shannon entropy is defined as "the averaged amount of information after excluding information redundancy", which is defined by Shannon as "something that can

remove uncertainty". Most modern scholars have proposed the opposite view to Shannon's data definition, that is, information is the increase of certainty. In accordance with information theory, the multidimensional information created by the fusion of multiple one-dimensional pieces of information is more informative than any one-dimensional piece of information, which is the theoretical ground for the fusion of multi-sensor data. Below we give the proof from the perspective of Shannon entropy [10].

Suppose the Shannon entropy $H(X)$ of the random variable X is a function of the probability distribution P_1, P_2, \dots, P_n . According to the definition of Shannon entropy:

$$H(X) = - \sum_{j=1}^n P_j \log P_j \tag{1}$$

Among them, $0 \leq P_j \leq 1$, easy to get:

$$H(P_1, P_2 \dots P_n) \geq 0 \tag{2}$$

If and only if each item on the right side of the equal sign in formula (1) is 0, the equal sign in formula (2) is true:

$$\sum_{j=1}^n P_j = 1 \tag{3}$$

Combining formula (2) and formula (3), we can see that when $P_j = 1$ and $P_k = 0$, formula (2) takes the equal sign.

Assuming that the Shannon entropy of random variables X and Y are $H(X)$ and $H(Y)$, respectively, their joint Shannon entropy is $H(XY)$. According to the additivity of Shannon entropy, we can know:

$$H(XY) = H(X) + H(Y|X) \tag{4}$$

Suppose the Shannon entropy $H(Y)$ of the random variable Y is a function of $P_1, P_2 \dots P_m$, and the conditional transition probability of X and Y is P_{ij} . Combining formulas (1) and (4), the Shannon of the two-dimensional random variable can be obtained. Entropy expression:

$$\begin{aligned} & H(P_1P_{11}, P_1P_{12}, \dots, P_1P_{1n}, P_2P_{21}, P_2P_{22}, \dots, P_2P_{2n}, P_mP_{m1}, P_mP_{m2}, \dots, P_mP_{mn}) \\ &= H(P_1, P_2, \dots, P_n) + \sum_{j=1}^m P_j H(P_{j1}, P_{j2}, \dots, P_{jn}) \end{aligned} \tag{5}$$

From the non-negativity of Shannon entropy and $0 \leq P_j \leq 1$, the formula (6):

$$H(P_1P_{11}, P_1P_{12}, \dots, P_1P_{1n}, P_2P_{21}, P_2P_{22}, P_2P_{2n}, \dots, P_mP_{m1}, P_mP_{m2} \dots P_mP_{mn}) \geq H(P_1, P_2, \dots, P_n) \tag{6}$$

Generalizing to the scenario of n random variables X_1, X_2, X_n , from the additivity of Shannon entropy, formula (7) can be obtained:

$$H(X_1X_2 \dots X_n) = H(X_1) + H(X_1|X_2) + \dots + H(X_n|X_1X_2 + \dots X_{n-1}) \tag{7}$$

Shannon entropy is a quantity that describes the uncertainty of a system or variable, not the amount of information in the system, but when a random variable takes a specific value, the value with information is equal to the Shannon entropy [11]. The larger the Shannon entropy, the larger the amount of information that the random variable has when it takes a specific value. Combining formulas (6) and (7), it can be inferred that the multi-dimensional information fused by multiple single-dimensional information contains more information for a specific target than any single-dimensional information [12].

The functional model diagram of multi-sensor data fusion is shown in Fig. 1.

As a matter of fact, multi-sensor information fusion is a functional simulation of the human mind's integrated processing of complex problems. When compared with single sensor, the use of multi-sensor information fusion technology can enhance the system survivability, reliability and robustness of the whole process, credibility of data, accuracy, duration and spatial coverage, realize real-time and information utilization, etc., in addressing the problems of exploration, tracking, and object identification.

Translated with <http://www.DeepL.com/Translator> (free version).

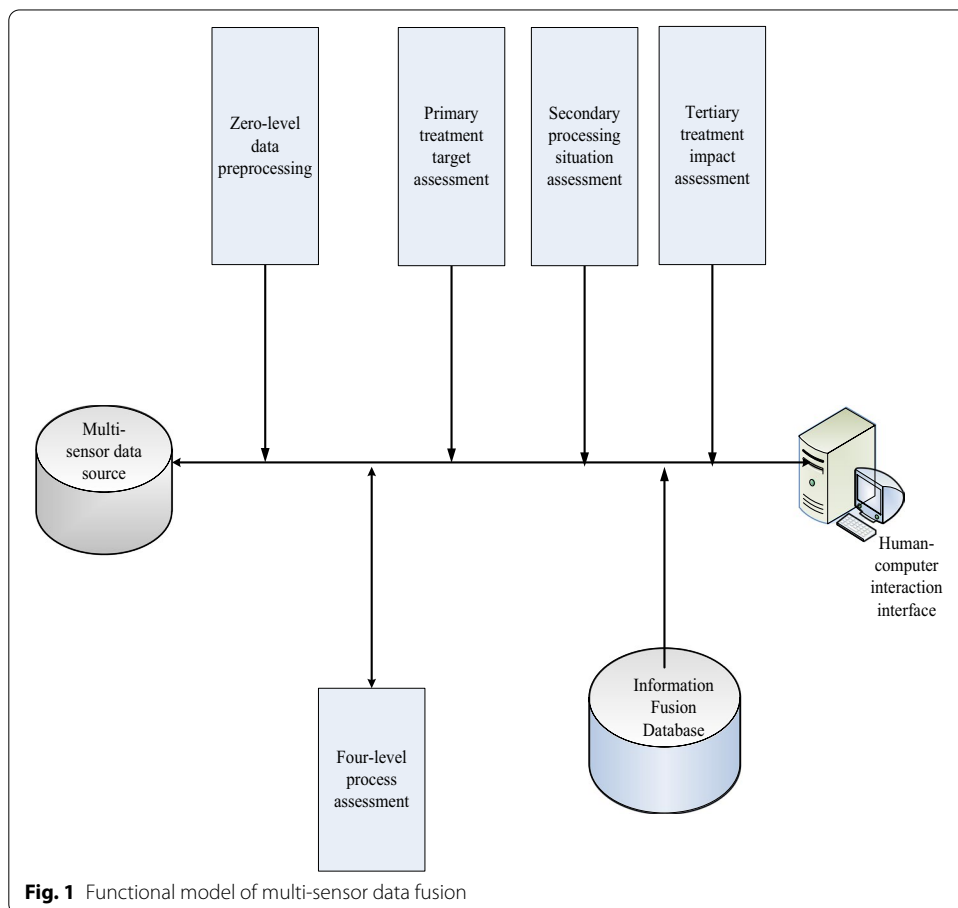


Fig. 1 Functional model of multi-sensor data fusion

2.2 Image processing

Image processing refers to the use of a computer to process the image to be recognized to meet the subsequent needs of the recognition process. It is mainly divided into two steps: image preprocessing and image segmentation [13]. Image preprocessing mainly includes image restoration and image transformation. Its main purpose is to remove interference and noise in the image, enhance the useful information in the image, and improve the detectability of the target object. At the same time, due to the real-time requirements of image processing, it is necessary that the image is re-encoded and compressed to reduce the complexity and computational efficiency of subsequent algorithms. The existing image segmentation methods mainly include edge-based segmentation, threshold-based segmentation, and region-based segmentation [14].

2.3 Significant area detection

The salient area detection aims to find the most salient target area in the picture. When observing a picture, a specific target object often attracts our attention immediately. When processing image scene information, it is possible to obtain priority processing target areas through saliency area detection, so as to rationally allocate computing resources and reduce the amount of calculation. Therefore, detecting the saliency area of the image has higher application value. Generally speaking, saliency area detection is divided into two types: top-down and bottom-up methods. The top-down approach is task-driven and requires the use of high-level information for supervised training and learning. This method has complex cross-disciplinary issues, because it probably requires a combination of neurology, physiology, and other related subject areas. The bottom-up method is data-driven, which mainly uses low-level information such as color contrast and spatial layout characteristics to obtain the saliency target area. This method is simple and fast to operate. Relevant studies in recent years have shown that this type of salient detection method has good results and has been widely used in image segmentation, target recognition, visual tracking, and other fields [15].

1. Bottom-up significant area detection

Bottom-up data-driven salient area detection has nothing to do with human cognition. The salient value is calculated by extracting the underlying features of the image. These features can be color, direction, brightness, or texture. Bottom-up salient area detection methods can be divided into salient area detection methods based on local contrast and salient area detection methods based on global contrast.

2. Top-down saliency area detection

The top-down salient area detection is a task-driven computing model. For example, if you want to find a person, you will pay attention to human-shaped objects when you look at the image; when you want to find a dog, you will ignore the person in the image and pay attention to the dog. Therefore, the top-down salient area detection is generally to find a certain kind of things [16]. Most of the top-down salient area detection requires training on a large amount of image data, which is computationally intensive and will get different results due to different tasks, which is not universal.

3 Algorithm design of saliency area detection based on multi-sensor data fusion

We collected data from our school’s database and related data sets on the Internet and collected data in three data sets. In this section, we will explain and analyze the three detection algorithms. First, we compare the characteristics of the three detection algorithms. In order to let reader see the characteristics of the three algorithms more clearly, we draw these characteristics become a table, the specific description is shown in Table 1.

3.1 Improvements to the FT algorithm

In the frequency domain, an image can be divided into high-frequency components and low-frequency components. The high-frequency components mainly reflect the detailed information of the image, such as texture, and the low-frequency components mainly reflect the overall information of the image, such as contours [17, 18]. In terms of frequency, Achanta et al. have proposed a frequency adjustment algorithm with good performance. The frequency adjustment algorithm is also called the FT algorithm. It mainly considers the saliency of the image from the frequency domain, adjusts the frequency components through a band-pass filter, and then calculates the saliency value to obtain the saliency map. The FT saliency detection algorithm is simple to implement [19], but has high performance and high reference value. In different scenarios, the FT algorithm has a better effect on images with smaller salient objects. However, in the FT algorithm, although the passband of each DoG filter is different, the status of each DoG filter is the same. This causes the saliency map of the FT algorithm to be easily disturbed by the performance of the object [20]. In order to play more of the role of the low-frequency part while suppressing the influence of the high-frequency part, this chapter will improve the FT algorithm through the frequency weighting method and increase the weight of the low-frequency part. First, set the weight of the DoG filter with the lower passband to be larger, as shown in formula (8):

$$FN(x, y) = \sum_{n=0}^{N-1} Wn[G(x, y, \sigma\rho^{n+1}) - G(x, y, \sigma\rho^n)] \tag{8}$$

In order to simplify the formula, we can set Wn to $n+1$, so that when the filter frequency decreases, the weight linearly increases, formula (8) can be rewritten as formula (9):

Table 1 Description of the respective characteristics of the three detection algorithms

Algorithm	Contrast type	Multi-scale	Whether based on biological structure	Whether to consider frequency	basic unit	Whether to consider the space factor
FT	Global	No	No	No	Pixel	No
LC	Global	Yes	No	Yes	Pixel	No
LVA	Global	No	Yes	Yes	Pixel	No

$$FN(x, y) = \sum_{n=0}^{N-1} (n+1)[G(x, y, \sigma\rho^{n+1}) - G(x, y, \sigma\rho^n)] \quad (9)$$

After $FN(x, y)$ filtering, the significant value of the improved FT algorithm can be expressed by formula (10):

$$S(x, y) = \sum_{n=0}^{N-1} ||I_\mu - I_{\omega_{hc}}^n(x, y)|| \quad (10)$$

Through the improvement and optimization of the FT algorithm, we can get the saliency map shown in Fig. 2.

3.2 LC algorithm

The LC algorithm is based on contrast-based saliency detection, which compares three levels of saliency description, the three levels are: salient point, salient area, salient view [21]. In this algorithm, the color statistics of the image are used to display the color contrast information of the scene. Through the saliency map at the pixel level, the points of interest can be found from the pixels with the local maximum saliency value, and the saliency area is in the saliency. Based on points. Among the three levels of division [22], salient points represent the most interesting points in the image; salient areas represent the most important parts and possible regions of interest in the image; salient views represent the main information and composition of the image. The calculation process of the LC algorithm is simple, but the number of comparisons is very large, and several techniques need to be used to speed up the calculation. The LC algorithm has a feature that it will highlight the importance of rare colors. If the rare colors of the image are in the salient area, the LC algorithm can be used [23].

The saliency map of the LC algorithm is based on the color contrast of the image pixels. The algorithm has a linear calculation complexity relationship with the number of pixels. At the same time, the algorithm does not limit the features [24]. The features mentioned in this article can theoretically calculate the contrast. For the original image I , the saliency value of the pixel I_k is defined as shown in formula (11):

$$S(I_k) = \sum_{\forall I_i \in I} ||I_k - I_i|| \quad (11)$$

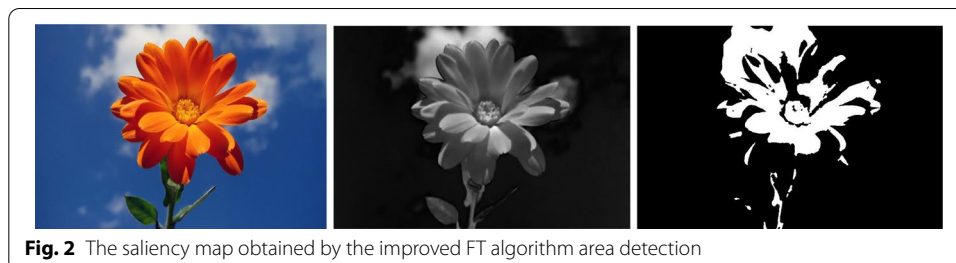


Fig. 2 The saliency map obtained by the improved FT algorithm area detection

Among them, the range of I_i is 0–255, $\|*\|$ represents the color distance measurement standard, formula (11) can also be extended to formula (12):

$$S(I_k) = \|I_k - I_1\| + \|I_k - I_2\| + \dots + \|I_k - I_n\| \tag{12}$$

Among them, n is the total number of pixels in the original image. Given an input image, the color value of each pixel I_i is known. Then formula (12) can be simplified to formula (13):

$$S(a_m) = \sum_{n=0}^{255} f_n \|a_m - a_n\| \tag{13}$$

The calculation of the LC algorithm is usually slow, and the complexity is $O(N^2)$. In practical applications, the calculation process needs to be optimized. The f_n in formula (13) can be described by a histogram. The complexity of the histogram is $O(N)$. Since $a_n \in [0,255]$, the range of the color distance measurement standard $\|a_m - a_n\|$ is also $[0,255]$. Therefore, for this fixed range we can construct a distance matrix D [25, 26]. In this matrix, the element $D(x,y) = \|a_m - a_n\|$ is the color difference between a_x and a_y , and the significant value of pixel I_k can be calculated by formula (14):

$$S[I_k] = S[a_m] = \sum_{n=0}^{255} f_n D(m, n) \tag{14}$$

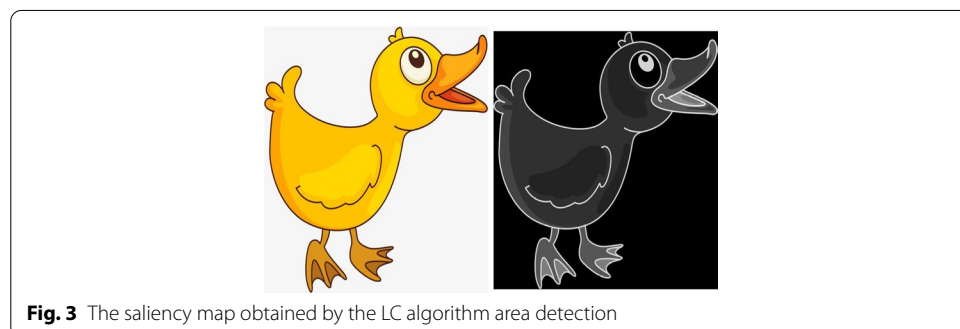
Through the calculation and optimization of the above formulas, the saliency map of the LC algorithm shown in Fig. 3 can be obtained.

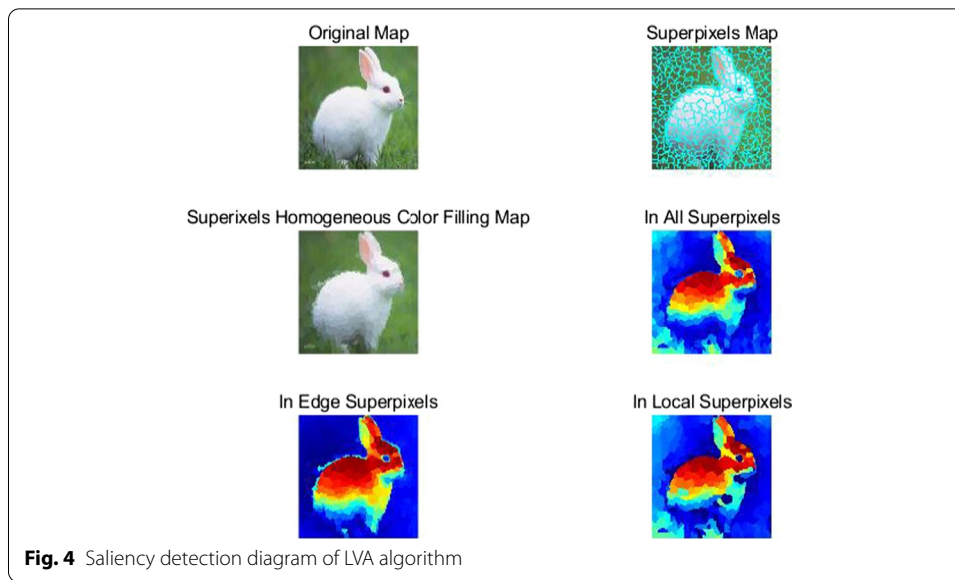
3.3 LVA algorithm design based on multi-sensor data fusion

The LVA algorithm is the link vector algorithm. By the detailed descriptiveness in the previous two subsections, the whole process of LVA algorithm can be divided on the following four cascading steps.

(1) Using the improved FT algorithm and LC algorithm to obtain the primary saliency map and.

(2) Combining the advantages of the improved FT algorithm and the LC algorithm, recombine the saliency maps S_{FT} and S_{LC} into a new saliency map S . Here, the linear combination method is used, and the calculation formula is shown in (15):



**Table 2** System and software environment

Operating System	Windows10 64bit Professional Edition
RAM	16 GB
CPU	Intel(R)Core i5-7300u
Software Environment	Matlab R2010a

$$S = a \cdot S_{FT} + b \cdot S_{LC} \quad (15)$$

The current research results have not studied the selection of coefficients. Through a large number of experiments, the value of a is 0.5, and b is 0.5.

(3) Combining the position information of the image to help the salient image extraction. We weight the position information of S to get the salient image S_1 .

(4) Use the method to enhance the saliency image S_1 , and finally obtain the saliency image S_2 of the LVA algorithm.

The saliency map detection diagram of the LVA algorithm is shown in Fig. 4.

4 Comparative analysis of three image salient area detection algorithms

Next, we will test and analyze several aspects of the LVA algorithm designed in this article, and compare it with the classic FT algorithm and LC algorithm. This article uses PASCAL, ECSSD, DUO-OMRON these three data sets with a large number of single object images, which meet our test conditions. The system and software environment are shown in Table 2.

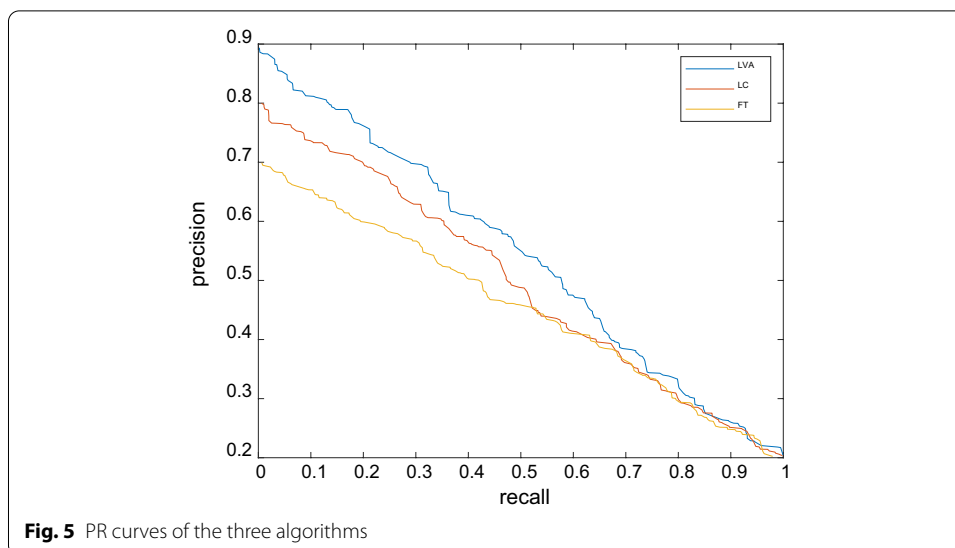


Fig. 5 PR curves of the three algorithms

4.1 PR curve

The difference between the saliency map and the manually labeled standard map is an important criterion for judging the pros and cons of a salient area detection algorithm. The saliency map is binarized by the fixed threshold method to obtain 256 binary images; then each binary image is compared with the standard image, and the accuracy of the saliency is determined by the accuracy and recall curves [27]. The PR curves of the three algorithms are shown in Fig. 5.

It can be seen from Fig. 5 that the abscissa represents the recall rate, and the ordinate represents the precision. In the PR-curve, when $T_f=0$, the entire saliency map is recognized as a saliency region, that is, the pixel value of all pixels is 1. When, no matter which detection method is used on which test set, $R=1$, that is, the recall rate is 1. When $T_f=1$, the entire saliency map is recognized as the background area, that is, the pixel value of all pixels is 0. At this time, no matter which detection method is used on which test set, $R=0$, that is, the recall rate is 0. The P value and the R value have a mutually restrictive relationship. The higher the P value, the more pixels in the detected saliency area belong to the saliency target area, and vice versa, the more they belong to the background area. But we can't blindly pursue the accuracy rate, which will lead to a lower recall rate. In the same way, we cannot blindly increase the recall rate, so the precision rate will continue to decrease. The precision recall curve clearly shows that the LVA algorithm is better than the improved FT algorithm and the LC algorithm. This also shows that the pixels belonging to the salient area in the image used in this article are on average 25%.

4.2 P-R-F diagram analysis of three algorithms

Regarding the Precision and Recall of the experimental results, it must be that they are larger, which means that the experimental results are more ideal, but in realities there will always be contradictions between fish and bear's paws. In order to balance such contradictions, we need to make certain combinations and trade-offs on all aspects of the requirements, and the more widely popular method is F-Measure to measure. Because

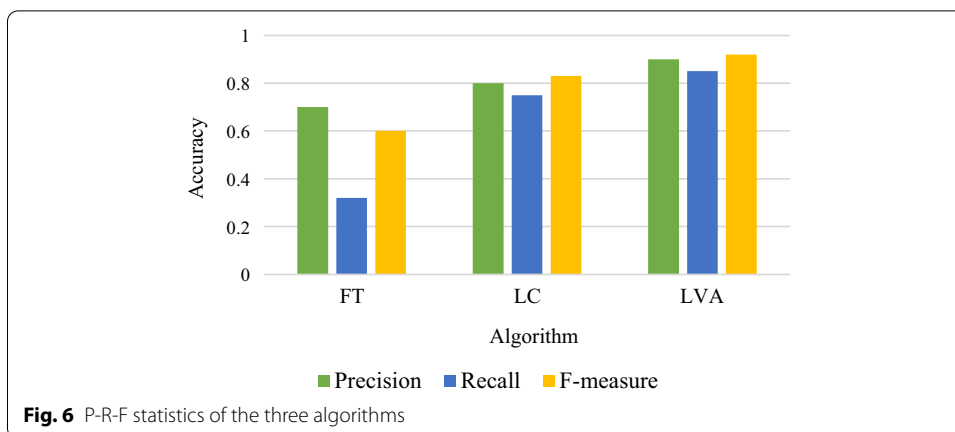


Fig. 6 P-R-F statistics of the three algorithms

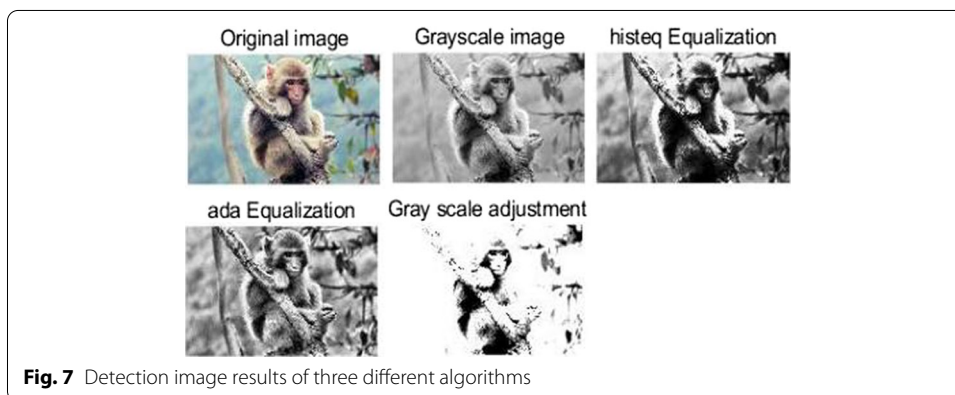


Fig. 7 Detection image results of three different algorithms

F-Measure is an evaluation index proposed after comprehensive consideration of these two indicators, it is a harmonic average of Precision and Recall, which is used to comprehensively reflect the overall index [28]. The calculation method of F-measure is related to the way of saliency map binarization, and when saliency map binarization uses image-related fixed thresholds, the value of F-measure can be obtained by consulting the PR curve. The statistical results are shown in Fig. 6.

It can be seen from Fig. 6 that the algorithm proposed in this paper is significantly better than the other two algorithms, and the highest accuracy rate is close to 1. The correct rate and recall rate of the FT algorithm are low. Although the FT algorithm can detect significant areas, it contains a lot of cluttered background and has a low resolution. The LC algorithm uses smooth prior theory, and the PR curve is higher than the FT algorithm. The whole is lower than the algorithm of this paper. The reason is that this article combines bottom-up and top-down models. The bottom-up weak saliency detection model can effectively and smoothly detect the salient regions, and the top-down strong saliency detection model goes further the cluttered background interference is suppressed, which greatly improves the detection accuracy, and the F-measure value of the algorithm in this paper is 0.92, and the highest recall rate is 0.85, which effectively shows that the algorithm in this paper can evenly and correctly detect the entire salient target part.

Table 3 Evaluation indicators related to the three algorithms

Algorithm	Information entropy	Average gradient	Spatial frequency	SSIM
FT	6.321	6.124	26.1247	0.5147
LC	6.584	6.482	26.1533	0.4124
LVA	7.029	7.213	27.1245	0.6832

4.3 Other related evaluation indicators

After we detect the salient area of the same original image according to the three algorithms, the result of the detected image is shown in Fig. 7.

From the saliency map detection results in Fig. 7, we can calculate various evaluation indicators for image fusion of these three algorithms according to the relevant formulas. The calculation results are shown in Table 3.

From Table 3, we can see that the LVA algorithm proposed in this paper performs well in most evaluation indicators, especially the SSIM indicator, which is the best in each group of data. From the previous evaluation index introduction, it can be seen that SSIM is more sensitive to the structural similarity of the fusion image and the source image, that is, to the similarity of the texture information. Therefore, through the analysis of the results of the above indicators, it can be explained that the fusion rules proposed in this chapter can be the effective preservation of image details that indicates the effectiveness of the saliency calculation method proposed in this paper.

In addition, we also calculated the average saliency map detection time of the LVA algorithm, the FT algorithm, and the LC algorithm mentioned in this article on the three data sets and calculated the average running time of the three algorithms in Table 4.

It can be seen from Table 4 that the LVA algorithm proposed in this paper takes the least average time to detect saliency maps on different data sets.

Finally, we calculated the MAE indices of the three algorithms on three different data sets, and the calculation results are shown in Table 5.

It can be seen from Table 5 that the data set has similar effects on the MAE index, and the salient area detection algorithm based on local and global salient information is also affected by the initial saliency map. When the data set becomes larger, there are relatively fewer accidents in the detection., The influence of the initial saliency map on the entire algorithm is also decreasing. The DUO-OMRON dataset contains the most pictures in the three datasets. The PASCAL-S data set contains the least pictures. However,

Table 4 Running time of the three algorithms

Algorithm	FT	LC	LVA
Time(s)	0.945	0.848	0.736

Table 5 MAE index of three algorithms

	PASCAL	ECSSD	DUO-OMRON
FT	0.621	0.485	0.678
LC	0.614	0.545	0.574
LVA	0.735	0.632	0.741

the PASCAL-S data set is the most difficult to detect. The images of the data set contain multiple targets, which is not suitable for situations such as prior knowledge [29]. But in general, the algorithm proposed in this paper can get better detection results.

5 Conclusions

With the advancement of technology, the image data in the Internet are growing rapidly, and people's requirements for computer vision are getting higher and higher. The salient area of the image contains the most important information in the image. As a visual information selection method, salient area detection is of great significance for computer vision tasks. The salient area detection can obtain salient areas in the image that meet human perception and can provide very effective preprocessing for computer vision tasks. In recent years, salient area detection has become a research hotspot in the field of computer vision. Of course, this article also has some shortcomings. The research in this article only compares sensor data fusion with the other two algorithms. The conclusions drawn are relatively simple, and the sensor data fusion and image detection methods are relatively simple. Increase the data comparison of the methods to make the conclusions more reasonable.

Abbreviations

FT: Frequency-tuned; LC: Luminance contrast; LVA: Link vector algorithm.

Authors' contributions

Xinyu Zhang: contributed significantly to analysis and manuscript preparation; Kai Ye: performed the data analyses and wrote the manuscript. All authors read and approved the final manuscript.

Authors' Information

Xinyu Zhang was born in Xi'an, Shaanxi, P.R. China, in 1982. She received the Master's degree from *XiDian University, P.R. China*. She works in School of Information Engineering, Xi'an University. Now, she studies in School of Automation Science and Technology, Xi'an Jiaotong University. Her research interests include image and signal processing, machine learning, and big data analysis.

Kai Ye professor and doctoral tutor of *School of Automation Science and Technology, Xi'an Jiaotong University*. He is a visiting professor at Leiden University. From 1999 to 2003, he received bachelor's degree and master's degree from Wuhan University; In 2008, he received doctor's degree from Leiden University. In 2008, he was engaged in postdoctoral research in the European Institute of bioinformatics; From 2009 to 2012, he was an assistant professor at Leiden University School of medicine in the Netherlands; from 2012 to 2016, he was an assistant professor at Washington University in St. Louis, USA. He has published 27 SCI papers in science, nature, NAT Med, NAT commun and other journals (average $if > 11$, ESI was cited 2), and he cited more than 2677 times of SCI; 21 SCI papers were published by communication/CO communication authors, 15 in recent five years. His current research fields: Large data mining, pattern recognition, computer algorithms, machine learning, bioinformatics, genomic variation.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This article is ethical, and this research has been agreed.

Consent for publication

The picture materials quoted in this article have no copyright requirements, and the source has been indicated.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Information Engineering, Xi'an University, Xi'an 710065, Shaanxi, China. ²School of Automation Science and Technology, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China.

Received: 22 July 2021 Accepted: 28 September 2021

Published online: 12 October 2021

References

1. J.-F. Ehlenbröcker, U. Mönks, V. Lohweg, Sensor defect detection in multisensor information fusion. *J. Sens. Sens. Syst.* **5**(2), 337–353 (2016)
2. L. Yue, D.K. Jha, A. Ray et al., Information fusion of passive sensors for detection of moving targets in dynamic environments. *IEEE Trans. Cybernet.* **47**(1), 93–104 (2016)
3. Z. Hu, H. Zhu, M. Hu et al., Rail surface spalling detection based on visual saliency. *IEEJ Trans. Electr. Electron. Eng.* **13**(3), 505–509 (2018)
4. A. Cecaj, M. Mamei, F. Zambonelli, Re-identification and information fusion between anonymized CDR and social network data. *J. Ambient. Intell. Humaniz. Comput.* **7**(1), 83–96 (2016)
5. B.J. Liu, Q.W. Yang, W.U. Xiang et al., Application of multi-sensor information fusion in the fault diagnosis of hydraulic system. *Int. J. Plant Eng. Manag.* **22**(01), 12–20 (2017)
6. L. Peng, L. Bo, Z. Wen et al., Predicting drug-target interactions with multi-information fusion. *IEEE J. Biomed. Health Inform.* **21**(2), 561–572 (2017)
7. K.G. Srinivasa, B.J. Sowmya, A. Shikhar, R. Utkarsha, A. Singh, Data analytics assisted internet of things towards building intelligent healthcare monitoring systems: iot for healthcare. *J. Organ. End User Comput.* **30**(4), 83–103 (2018)
8. P. Braca, R. Goldhahn, G. Ferri et al., Distributed information fusion in multistatic sensor networks for underwater surveillance. *IEEE Sens. J.* **16**(11), 4003–4014 (2016)
9. H. Li, Research on target information fusion identification algorithm in multi-sky-screen measurement system. *IEEE Sens. J.* **16**(21), 7653–7658 (2016)
10. C. Qing, F. Yu, X. Xu et al., Underwater video dehazing based on spatial-temporal information fusion. *Multidimension. Syst. Signal Process.* **27**(4), 909–924 (2016)
11. L. Zhang, Z. Xiong, J. Lai et al., Optical flow-aided navigation for UAV: a novel information fusion of integrated MEMS navigation system. *Optik Int. J. Light Electron Opt.* **127**(1), 447–451 (2016)
12. X. Xu, D. Cao, Y. Zhou et al., Application of neural network algorithm in fault diagnosis of mechanical intelligence. *Mech. Syst. Signal Process.* **10**, 106625 (2020)
13. F. Sattar, F. Karray, M. Kamel et al., Recent advances on context-awareness and data/information fusion in ITS. *Int. J. Intell. Transp. Syst. Res.* **14**(1), 1–19 (2016)
14. B.S. Chandra et al., Robust heartbeat detection from multimodal data via CNN-based generalizable information fusion. *IEEE Trans. Biomed. Eng.* **66**(3), 710–717 (2019)
15. T. Sasaoka, I. Kimoto, Y. Kishimoto et al., Multi-robot SLAM via information fusion extended Kalman filters. *IFAC PapersOnLine* **49**(22), 303–308 (2016)
16. S. Zhang, T. Feng, Optimal decision of multi-inconsistent information systems based on information fusion. *Int. J. Mach. Learn. Cybern.* **7**(4), 563–572 (2016)
17. X. Li, E. Seignez, A. Lambert et al., Real-time driver drowsiness estimation by multi-source information fusion with Dempster-Shafer theory. *Trans. Inst. Meas. Control* **36**(7), 906–915 (2017)
18. E. Moamed, S. Abdulaziz, X.H. Yuan, Optimizing robot path in dynamic environments using genetic algorithm and Bezier curve. *J. Intell. Fuzzy Syst.* **33**(4), 2305–2316 (2017). <https://doi.org/10.3233/JIFS-17348>
19. P.K. Davis, D. Manheim, W.L. Perry et al., Causal models and exploratory analysis in heterogeneous information fusion for detecting potential terrorists. *Tetrahedron Lett.* **44**(44), 8165–8167 (2017)
20. E. Bareinboim, J. Pearl, Causal inference and the data-fusion problem. *Proc. Natl. Acad. Sci. USA* **113**(27), 7345–7352 (2016)
21. A. Ali, H.U. Farid, Z.M. Khan et al., Temporal analysis for detection of anomalies in precipitation patterns over a selected area in the Indus basin of Pakistan. *Pure Appl. Geophys.* **178**(2), 651–669 (2021)
22. H.A. Vargas, A.M. Hötker, D.A. Goldman et al., Updated prostate imaging reporting and data system (PIRADS v2) recommendations for the detection of clinically significant prostate cancer using multiparametric MRI: critical evaluation using whole-mount pathology as standard of reference. *Eur. Radiol.* **26**(6), 1606–1612 (2016)
23. E. Hisham, E. Mohamed, A.M. Riad, A.E. Hassani, A framework for big data analysis in smart cities. *Adv. Intell. Syst. Comput.* **723**, 405–414 (2018)
24. J. Park, S.C. Kim, The significance test on the AHP-based alternative evaluation: an application of non-parametric statistical method. *J. Soc. e-Bus. Stud.* **22**(1), 15–35 (2017)
25. A.P. Silva, M.N. Vieira, A.V. Barbosa, Forensic speaker comparison using evidence interval in full Bayesian significance test. *Math. Probl. Eng.* **2020**(1), 1–9 (2020)
26. D. Wittenburg, V. Liebscher, An approximate Bayesian significance test for genomic evaluations. *Biom. J.* **60**(6), 1096–1109 (2018)
27. H. Moghadam, M. Rahgozar, S. Gharaghani, Scoring multiple features to predict drug disease associations using information fusion and aggregation. *SAR QSAR Environ. Res.* **27**(8), 609–628 (2016)
28. Z. Yu, L. Chang, B. Qian, A belief-rule-based model for information fusion with insufficient multi-sensor data and domain knowledge using evolutionary algorithms with operator recommendations. *Soft. Comput.* **23**(13), 5129–5142 (2019)
29. Y.B. Salem, R. Idodi, K.S. Ettaba'a et al., High level mammographic information fusion for real world ontology population. *J. Digital Inf. Manag.* **15**(5), 259–271 (2018)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.