Advances in Educational Technology and Psychology (2021) 5: 203-217
Clausius Scientific Press, Canada
DOI: 10.23977/aetp.2021.57027
ISSN 2371-9400

Design of Intelligent Education Service Robots for Big Data

Xiaojun Liu[1,2a*], Chunrun Guo[1,b]

[1]*School of Electromechanical and Automobile Engineering, Huanggang Normal University, Huanggang 438000, Hubei, China*

[2]*Hubei Zhongke Research Institute of Industrial Technology, Huanggang 438000, Hubei, China*

[a]*18623582@qq.com*, [b]*2019280340130@smail.hgnu.edu.cn*

**Corresponding author*

Keywords: Big Data, Intelligent Education Service Robot, Saliency Region Extraction Algorithm, Visual System, Voice Interaction System

Abstract: With the development of robot technology, service robots have gradually come into people's lives, which also makes human-computer interaction more and more frequent. However, most of the current robot control algorithms have low accuracy and difficult operation, and people can't wait to find more effective algorithms. In order to design the most convenient and accurate intelligent education service robot, this paper uses SF algorithm and improved PCNN to establish a hybrid model, and proposes an improved saliency region extraction algorithm based on education service robot. The algorithm compares the standard database with the real environment, and the PR curve of the proposed algorithm is improved by about 5%. It is 10% higher than the SF algorithm in the AUC index, and the comprehensive F value is improved by 3.4% ~ 7.4%. This paper fully demonstrates that the saliency area generated by the proposed algorithm is closer to the true value, which can effectively suppress the high-brightness background area in the detection results of the SF algorithm. It also verifies that the PCNN model with the neuron propagation stimulation mechanism as the core more effective simulation of biological vision systems. Combining compressed sensing technology, this paper proposes a speech recognition scheme that is easy to implement in hardware. The algorithm performance of the robot is verified, and the optimal effect parameters are selected through comparative experiments. The method uses Chinese phonetic phrase (sentence) test to obtain good recognition results, and it can be used as an effective improvement scheme for the speech input of the proposed robot voice interaction system.

1. Introduction

As a high-speed development field, robots are widely used in industries, services, medical

treatment, transportation, and entertainment. The education service robot is designed for teaching, with the purpose of stimulating students' interest in learning, and the purpose of cultivating students' comprehensive ability. Through the modular construction method, students' practical ability and innovation ability are strengthened. Among these many ways of interaction, language and vision are the most direct, convenient and natural expressions for humans to communicate. Compared with traditional human-computer interaction, voice interaction and visual interaction break through the dependence on fingers and screen, and enhance the active communication and response of the machine, especially for the elderly who are slow to respond, the mentally retarded, the blind and the illiterate. It improves the efficiency of human-computer interaction and can satisfy the expression of human emotions.

With the development of robot technology, service robots have gradually come into people's lives, which also makes human-computer interaction more and more frequent. In [1], the authors evaluated the benefits of social robots to support children's learning in an educational environment over a longer period of time, and evaluated the effects of personalization and adaptation to the social behavior of robots. The results show that although children with personalized and non-personalized conditions are learned, children's learning of novel topics increases when interacting with robots that personalize their behavior, suggesting that this benefit extends to other class-based performance. Other evidence obtained indicates that it provides the first evidence to support the personalization of peer-to-peer robot behavior, which has a positive impact on learning when personalization of peer-to-peer robot behavior is embedded in the learning environment for a long time. In [2], the authors constructed a modular training pathway for robotic-assisted radical prostatectomy (RARP). The author's goal is to determine the sequence of procedure steps to be learned before the surgeon can perform the complete procedure without the intervention of a tutor. The article designs a new modular training route and develops an evidence-based modular training approach that can promote the safe introduction of RARP to novice doctors. Robotics competitions can play an important role in education due to their inherent multidisciplinary approach, which can motivate students to bridge across different technical fields. It can also play an important role in research and development, as its results are expected to be transferred later to practical problems in manufacturing or servicing robots. The author proposes a possible method to build a robot that can participate in competitions, which can be used as a reference for current and new teams [3]. In [4], the author introduced the development of a reconfigurable virtual environment for multi-robot programming, control, simulation, and monitoring, and its application in the education process. During the development process, special attention was paid to the realization of modularity, scalability, and openness, which represent the basic requirements for reconfigurability. It represents an integrated environment that can continuously adjust the actual laboratory environment and robot usage. By combining different operating modes and virtual environments, it is possible for the author to adjust the virtual environment based on the students' knowledge and experience. In [5], the authors conducted a randomized controlled trial of supportive education services for 33 veterans with high educational goals and post-traumatic stress disorder (PTSD). The severity of PTSD symptoms and recovery attitudes cannot predict the impact of supported educational interventions. The findings suggest that supported education services can have a measurable impact on the time to reach educational goals. Future research will require longitudinal studies and research on attrition issues and capture the impact on other educational outcomes, such as successful course completion.

Big data is an information asset that is characterized by such a high volume, speed, and diversity that it requires specific technologies and analytical methods to turn it into value. Formally proposed definitions can make the development of big data concepts more coherent, as it only relies on the

basic elements of the latest technology and is consistent with the most popular definitions currently in use. In [6], the author describes the latest communication technologies and smart-based applications used in smart city environments. The author proposes a future business model for smart city big data and identifies challenges in business and technology research. This research can provide a benchmark for researchers and industry in the future development and development of smart cities in the context of big data. In [7], the author explored various methods to integrate big data analysis with network optimization to improve the quality of user experience. The author first proposes a framework for big data-driven (BDD) mobile network optimization. The author then describes the characteristics of big data collected not only from user devices but also from mobile networks. In [8], the author introduced a knowledge fusion taxonomy to understand the relationship between traditional marketing analysis (TMA), big data analysis (BDA), and new product success (NPS). The author draws on the vast amount of information and knowledge from different stakeholders in the digital economy. This taxonomy is designed to help companies develop strategies to integrate knowledge from the fields of marketing and big data. Research shows that knowledge fusion is not an automatic means to improve NPS, it requires strategic choices to obtain its benefits. In [9], the author introduced a method for identifying and omitting less informative observations in an online and data adaptive manner. Given the stream data, the author sequentially finds the relevant maximum likelihood estimator using first and second order random approximation algorithms. Compared with traditional methods, numerical tests on real and synthetic datasets confirm the effectiveness of the proposed data adaptation method. In [10], the author first provided a functional framework for identifying areas for the acquisition, management, processing, and mining of Internet of Things (IoT) big data, and defined and described several related technologies based on their main features and functions. The article analyzes the current research in the application of the Internet of Things. In addition, it identifies the challenges and opportunities related to the research of Internet of Things big data.

In order to design the most convenient and accurate intelligent education service robot, this paper uses SF algorithm and improved PCNN to establish a hybrid model, and proposes an improved saliency region extraction algorithm based on education service robot. The algorithm first takes the brightness feature map obtained by the SF algorithm as the input neuron of the PCNN, and outputs it as an internal neuron after the modulation and demodulation unit; then the dot multiplication result of the internal neuron and the binary initial saliency map generated by the SF algorithm as an input signal, the ignition pulse unit is improved to optimize the ignition range. Combining compressed sensing technology, this paper proposes a speech recognition scheme that is easy to implement in hardware. The algorithm performance of the robot is verified, and the optimal effect parameters are selected through comparative experiments.

2. Method

2.1 Related Theoretical Foundations of Big Data Technology

Big data refers to a huge collection of data that cannot be managed and processed with existing general technology. An information asset is defined to say that its own data cannot be directly analyzed by statistics, and it needs to be combined with new cloud computing processing methods to centrally process the data before it can be found for analysis and decision-making. Due to the special significance of big data, the definition of big data has changed over time. From the "3V" proposed by the earliest Mata Group (now Gartner) to the "5V" proposed by IBM this year. ”Features: huge amount (Volume, super petabytes of data), authenticity (Veracity, data availability

quality), speed (Velocity, frequency of data generation and update), diversity (Variety), and value (Value, high cost value).

Big data driving is a new operation management model. In the era of big data Internet, no matter which industry, a large amount of data will be generated in the development process. However, data is hidden, and individual data cannot be used as a whole anyway. If big data information is used to make decision management for the whole or individual, and the effective quantitative use of data becomes an important part of the development of the new era. Big data decisions reflect a more "precision" behavioral direction. Therefore, most management and development thinking will change the data-centric and data-driven new model. The development of the Cyber-Physical Convergence System (CPS) is in parallel with the development of big data technology. CPS system characteristics require computing devices to be able to quickly process large amounts of data such as sensors, system logs, etc. in order to feed back to the physical device in real time. The number of device integration units in CPS systems is increasing. The integrated processing unit has been unable to complete the calculation tasks quickly and in real time. The use of big data processing technology to process the huge amount of data has become an effective solution to the bottleneck of CPS development.

2.2 Speech Interaction Technology Based on Education Service Robot

(1) Compressed sensing of speech signals

Signal sparsity is a prerequisite for compressive sensing applications. However, in practical applications, we only need to determine whether the signal is compressible, that is, to sort the absolute value of the coefficient obtained after the signal has been transformed in a known domain in descending order, to determine whether the result is exponentially decaying and eventually approaches zero the result of. Here we take the research subject's speech signal as an example. Suppose there is an N-dimensional speech signal $x=[x_1,x_2,\ldots,x_N]^T$ that is sparse on an orthogonal basis $\Psi=[\varphi_1,\varphi_2,\ldots,\varphi_N]$, then x can be linearly represented by:

$$x=\sum_{i=1}^{N}\theta_i\varphi_i=\Psi\theta \tag{1}$$

In the above formula, $i=1,2,3,\ldots,N,\varphi_i\in R^{N\times 1}$, $x_i\in R^{N\times 1}$ and $\theta_i=\langle x,\varphi_i\rangle$ are called the sparse coefficients of the discrete signal x on the orthogonal basis Ψ. These coefficients need to satisfy:

$$\|\theta\|_p=\left(\sum_{i=1}^{N}|\theta_i|^p\right)^{1/p}\le K \tag{2}$$

Where $0<P<2$, $0<K$, $\theta=[\theta_1,\theta_2,\theta_3,\ldots,\theta_N]^T$ is the sparse vector and $\|\theta\|_p$ is the lp normal form of θ. In theory, if and only if $\|\theta\|_0=K(K\ll N)$, that is, θ has only K unknown non-zero elements, then the speech signal x can be said to be K-sparse on the orthogonal basis Ψ, and Ψ is called the sparse basis of the discrete signal x above formula is the sparse representation of x in the known domain, but in practical applications, $\|\theta\|_0$ cannot be solved, and we can obtain the optimal value by solving $\|\theta\|_1$.

(2) Selection of observation matrix

According to the compressed sensing theory, the measurement matrix Φ is the key to obtaining the measurement value y. To achieve the purpose of reconstructing the original signal, the measurement matrix must ensure that two different K-sparse coefficients cannot be projected into the same vector x, that is, it must meet the RIP characteristics:

$$(1-\delta_k) \le \frac{\| \Phi x\|_2^2}{\| x\|_2^2} \le (1+\delta_k) \tag{3}$$

Assuming the compression ratio r=4, the observation sequence $y=\{y_1, y_2, \ldots, y_m \mid m=1,2,3,\ldots,M\}$, and the speech signal $x=\{x_1, x_2, \ldots, x_n \mid n=1,2,3,\ldots,N\}$, where i = 1,2,3 ... N, then the row step matrix can be expressed as:

$$\begin{cases} y_1 = x_1 + x_2 + x_3 + x_4 \\ y_2 = x_5 + x_6 + x_7 + x_8 \\ \vdots \\ y_m = x_{(4m-3)} + x_{(4m-2)} + x_{(4m-1)} + x_{(4m)} \end{cases} \tag{4}$$

The work of this paper is mainly to design an educational service robot that can intelligently recognize speech, so as to achieve a speech recognition and control robot solution that improves efficiency and saves system resources. Therefore, the selected observation matrix should also meet the requirements for easy hardware implementation: 1) The required storage space is small, the elements are real numbers, which only involve addition and subtraction calculations; 2) the irrelevance with the sparse orthogonal basis matrix is strong; 3) satisfying real-time sampling and reconstruction. The calculation of the matrix only involves addition and subtraction, and at the same time meets the RIP characteristics, and has good irrelevance. Although the random matrix has a better reconstruction performance than the structured matrix, this article does not consider reconstruction, but performs experiments based on the observations, and only needs to consider the observation performance of the observation matrix.

2.3 Hybrid model of visual area based on pulse coupled neural network (PCNN)

This paper compares the visual hybrid model based on PCNN with seven algorithms including IT, GB, SR, AC, CA, FT and SF for performance comparison.

(1) Saliency filtering algorithm (SF)

The concept of color independence is defined as the difference between an element and its neighboring elements, which represents the degree of discrimination between the corresponding element and other elements, which can be equivalent to a description of the salient area of the image. The SF algorithm specifically defines the color independence of superpixel i as the sum of the weighted distances of i and other superpixels j in the CIE Lab color space. The specific description is as follows:

$$U_i = \sum_{j=1}^{N} \left\| c_i - c_j \right\|^2 \cdot \underbrace{\omega\left(p_i, p_j\right)}_{\omega_{ij}^{p'}} \tag{5}$$

In the formula, U represents the color brightness value; i, j = 1,2, .., N, i ≠ j; N represents the number of superpixel blocks; p represents the spatial position of the superpixel; c represents the

color in CIE Lab space ; Indicates the weight coefficient, which is related to the distance of each superpixel unit in the superpixel space, that is, the farther superpixel unit has a lower significance contribution to the central superpixel unit. If the above formula is used to estimate the color independence of all superpixels directly, the operation complexity is O (N2). In order to reduce the calculation complexity, formula $\omega_{ij}^{(p)} = \frac{1}{Z_i}\exp\left(-\frac{1}{2\sigma_p^2}\left\|p_i - p_j\right\|^2\right)$ can be used to reduce the calculation complexity to O (N) by Gaussian weights. Zi is the normalization factor that guarantees $\sum_{j=1}^{N}\omega_{ij}^{(p)} = 1$

.

(2) Improved saliency area extraction algorithm (ISRE)

The proposed ISRE algorithm is a hybrid algorithm model based on PCNN and based on the detection results of the SF algorithm. The specific implementation of the algorithm includes six steps:

Step 1. Use the SF algorithm to coarsely segment the original image to obtain the initial saliency map OSM and brightness feature map IFM. The OSM is a saliency map extracted by the SF algorithm, while the IFM is generated on the basis of the OSM. Obtaining IFM is an important step in the overall algorithm, because the brightness feature is a prerequisite for the Human Vision System (HVS) to perceive all external visual information. Without the brightness feature, the visual system cannot perceive and describe color, shape, orientation, Features such as movement and depth distance. Using the brightness feature as the original stimulus information is most consistent with the human visual system;

Step 2. Set the PCNN input neurons to the pixels in the brightness feature map IFM to simulate the stimulation information of biological neurons;

Step 3. At the input of the PCNN, a synapse weight is used to form a local stimulus for each externally stimulated neuron as an input to the modulation unit;

Step 4. Connect the modulation unit in PCNN to generate an internal activity item;

Step 5. Use the threshold of 75% of the overall pixel value to binarize the OSM to obtain the OSM_C binarized segmentation map OSM_C. Multiply the output of Step 4 and OSM_C to generate a more accurate ignition range. Improvement and optimization of PCNN ignition pulse input mode;

Step 6. After multiple iterations, a binary saliency map is output.

Input unit: The weight W is composed of a 3 × 3 kernel matrix. The values in the matrix are determined by calculating the boundary distance from the central pixel to the surrounding 8 neighboring pixels. The specific formula is as follows:

$$\mathrm{W}_{ijkl} = \frac{1}{\sqrt{(i-k)^2 + (j-l)^2}} \tag{6}$$

In the formula, (i, j) represents the pixel coordinates of the central pixel in the image, and (k, j) represents the pixel coordinates of the pixels of the 8 domains, respectively.

Connect the modulation unit: Connect the modulation unit to the PCNN, and use the external biological neuron stimulus Fij received at the input as the main input signal to perform joint modulation coupling with the local stimulus input Lij generated by the weight matrix W. The modulation process is as follows Show:

$$U_{ij}[n] = F_{ij}[n]\left\{1 + \beta L_{ij}[n]\right\} \tag{7}$$

In the formula, Uij is the internal activity term generated by the modulation unit of the neuron, β is the connection strength coefficient, which is used to measure the strength of the connection between synapses in the neural network. The specific value is determined by formula $\beta = \frac{1}{\sqrt{STD_{ij}} + 1}$.

3. Experiment

3.1 Data Source

In this paper, three representative databases, ASD, SED2 and ECSSD, will be selected based on the visual blending algorithm proposed by PCNN and the existing algorithms for overall image vision evaluation.

(1) ASD: The database consists of 1,000 representative images found in the 10,000 images of the MARA10K database, and also provides one thousand truth-valued images of saliency object regions labeled manually.

(2) SED2: The database consists of 100 images containing two objects.

(3) ECSSD: The database also contains 1,000 images. The composition of the images is more complicated than that of the ASD database. It contains the foreground and background areas of various scenes.

In terms of voice interaction testing, the data set used in this chapter is a voice phrase (sentence) recorded in real-time using MATLAB software and recorded in text form, and its content is the voice interaction system robot control test voice instructions. The entire data set has 220 examples of Chinese male voice phrases (sentences). The detailed information is shown in Table 1.

Table 1: Speech data set details

Dataset name	Category	Number	Examples	Signal to noise ratio	Sampling frequency	Duration / article
robot_cmd	11	20	220	16	16	1

3.2 Experimental Protocol

To meet the needs of the voice interaction system test, before training, we processed 3 data sets into 6 data sets: (1) 4 sets of data sets for compressed sensing sampling and MFCC feature extraction. In this type of data set calculate the first-order and second-order dynamic features of the MFCC feature parameters to obtain two sets of data, and then perform endpoint detection on these two sets of data to remove the irrelevant speech segments to obtain two sets of data sets; (2) Perform compressed sensing without MFCC features extracted 1 set of data sets. This type of data set is directly trained using compressed sensing technology to obtain observation sequences. (3) A set of data sets that only perform MFCC feature extraction and do not perform compressed sensing sampling. This class uses Nyqui directly. The speech segments obtained by star's theorem are subjected to MFCC feature extraction to obtain a data set.

Referring to the visual area detection algorithm evaluation scheme, the ISRE model binary saliency map presented in this paper is based on the SF algorithm detection results as a template, and the ISRE model gray saliency map ISREM is mapped. In the three standard databases of ASD, SED2, and ECSSD, the experimental evaluation of the proposed algorithm and the existing seven algorithms is performed in both qualitative and objective quantitative ways. The seven evaluated algorithms are compared in order of chronological order, including IT , GB, SR, AC, FT, CA and

SF.

3.3 Evaluation Programs and Indicators

(1) P-R curve and ROC curve

First, define the saliency map of each algorithm as SM (Salient Map), and quantize it to [0,255]; then set a fixed threshold value every other value from 0 to 255, and binarize all saliency maps. The segmentation result is defined as Segmentation Salient Map (SSM); finally, the saliency truth map GT (Ground Truth) manually labeled with each image is compared with SSM, and the calculation result is drawn into the corresponding evaluation curve. Among them, the calculation formula of the ordinate Precision and the abscissa Recall of the P-R curve is as follows:

$$Precision = \frac{SSM \cap GT}{SSM}, \quad Recall = \frac{SSM \cap GT}{GT} \tag{8}$$

The ROC curve was originally an evaluation method to evaluate the quality of the classifier in the binary classification problem. When applied to the evaluation of the saliency detection algorithm, it can be similar to the strategy of drawing a PR curve, and the detected saliency map will be binary with a varying threshold. And evaluate the binary saliency saliency map and truth values containing only 0 and 1 as binary classification problems. The corresponding formulas for the ordinate True Positive Rate (TPR) and abscissa False Positive Rate (FPR) of the corresponding ROC curve are as follows:

$$TPR = \frac{DSR \cap GT}{DSR}, \quad FPR = \frac{DSR \cap \overline{GT}}{\overline{GT}} \tag{9}$$

(2) AUC. AUC is an index to measure the area under the ROC curve. An ideal model has an AUC of 1 and an AUC of 0.5 based on random guessing. When evaluating the AUC index of an algorithm, the closer the value is to 1, the better the algorithm performance.

(3) Comprehensive F-number. Although the PR curve and ROC curve can quantitatively evaluate the performance of each algorithm, when multiple algorithm curves are drawn together, the curves between algorithms with similar effects will be distributed in the dense axis area, which is not conducive to observing one algorithm. For specific performance, you can segment the saliency algorithm detection result by selecting an adaptive threshold instead of traversing all the thresholds from [0,255]. Calculated as follows:

$$T = \frac{2}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} SM(i, j) \tag{10}$$

Where W and H are the width and height dimensions of the saliency map SM. Using the formula to get a set of average precision and recall for all pictures in the database, the comprehensive F-measure (F-measure) can be obtained as an indicator of overall performance, as shown in the following formula:

$$F - measure = \frac{\left(1 + \lambda^2\right) Precision \times Recall}{\lambda^2 \times Precision + Recall} \tag{11}$$

(4) MAE. MAE is an evaluation index to measure the difference between the saliency map SM and the true value obtained by the algorithm. First, SM is normalized to the interval 0,1 to obtain

SM. Then, the difference between the pixel positions corresponding to the true value is calculated pixel by pixel. The absolute value is averaged, and the MAE calculation formula is as follows:

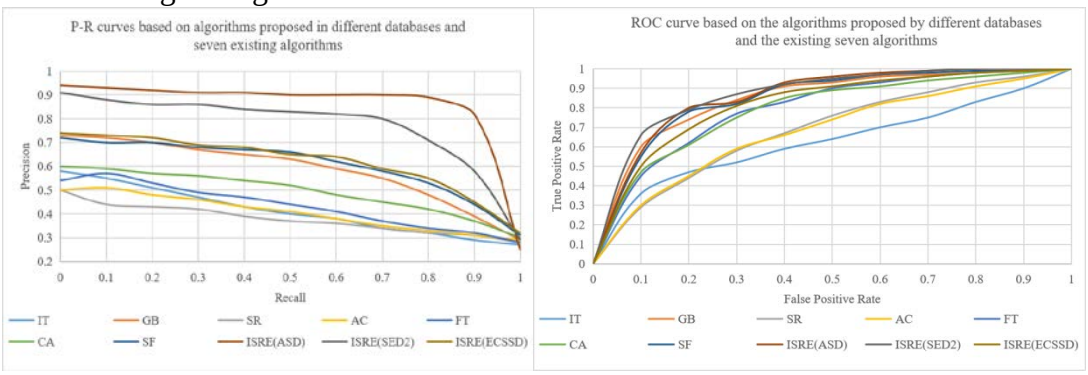
$$MAE = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} \left| SM'(i, j) - GT(i, j) \right| \tag{12}$$

4. Results and Discussions

4.1 Performance Analysis of Intelligent Education Service Robot

(1) P-R curve and ROC curve

The P-R curves and ROC curves of each algorithm based on the three databases are shown in Figure 1 (a) and Figure 1 (b), respectively. By comparing with the seven existing saliency detection algorithms on the PR curve and the ROC curve, it can be seen that on the three unique databases, the algorithms proposed in this paper have improved on the SF algorithm, which is consistently better than remaining six algorithms.



(a) P-R curve (b) ROC curve

Figure 1: P-R, ROC curves based on three database algorithms and seven existing algorithms

Since the algorithm in this paper is mainly an improvement on the SF algorithm, the proposed algorithm and the SF algorithm are individually evaluated in detail for the performance of the P-R curve. The proposed algorithm is mainly based on the SF algorithm to effectively suppress the strong brightness background interference in the image. From the observation of the PR curves of the three algorithms in the three databases, it can be found consistently that the maximum accuracy of the proposed algorithm (the PR curve horizontal when the coordinate is zero, the value of the ordinate) is improved by about 5%. Similarly, as the recall rate (abscissa of the PR curve) gradually increases, that is, as the fixed threshold gradually decreases from 255 to 0, the proposed algorithm checks the accuracy rate (the ordinate of the PR curve) is still better than the SF algorithm, which indicates that the value of the background area generated by the proposed algorithm is still below the threshold, which strongly proves that the proposed algorithm is better than the SF algorithm. The foreground area is significant and the background area is effectively suppressed.

At the same time, if the overall performance of each algorithm on the three databases is individually measured, the performance effects are ranked from excellent to poor in the order of ASD> SED2> ECSSD, which exactly validates the evaluation of the overall difficulty of the three databases. However, if you look at the performance of each algorithm in each database, it does not necessarily follow the order of the overall difficulty of the database. For example, the CA algorithm

can highlight and extract significant objects and large surrounding areas. Therefore, the ECSSD database with a more complex background performs better , but the SED2 database in multi-saliency objects has a poor monetization, which fully illustrates the necessity of unified evaluation of each algorithm in multiple unique databases.

(2) AUC value

In order to more intuitively reflect the evaluation effect of the ROC curve on each algorithm, the Area Under ROC Curve (AUC) is further obtained. A histogram of the calculation results based on the three database AUC values is shown in Figure 2.

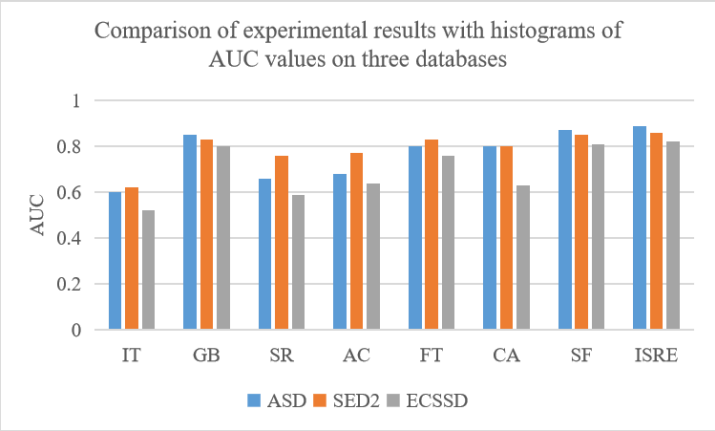


Figure 2: Histogram comparison experiment results of AUC values on three databases

As can be seen from Figure 2, the algorithm proposed in this paper achieved the highest AUC index on all three databases, the ASD database is 0.8918, the SED2 database is 0.8676, and the ECSSD database is 0.8142, indicating that the detection results of the proposed algorithm are closer to the true value. At the same time, the GB algorithm and the CA algorithm that are observed during the visual qualitative visual comparison evaluation are not good at extracting the edge information of the salient objects, but both algorithms tend to detect a larger area around the salient objects. From the comparison results of the AUC index, both algorithms perform better. Especially the GB algorithm, its overall performance is second only to the SF algorithm and the algorithm proposed in this paper, and its performance is better than the SF algorithm in the ECSSD database. The effective improvement of the proposed algorithm to the SF algorithm is explained from the side. Especially for the ECSSD database which is more difficult overall, the proposed algorithm improves the AUC index by 10% compared with the SF algorithm. Observing Figure 1, you will find that the convergence speed of the results of each algorithm is different. When the results of some algorithms are similar to the AUC index, you can further examine the corresponding integration curve and evaluate the corresponding algorithm more specifically.

4.2 Performance analysis of improved algorithm based on PCNN

(1) Integrated F value and MAE value

The F value is a comprehensive evaluation index of the recall rate and precision rate. The higher the value, the higher the accuracy of the detection result of the algorithm, that is, the unit of true value is 1, and the closer the F value is 1, it indicates the accuracy of the algorithm extraction area better (0.9 is better than 0.8). The MAE value is obtained by normalizing the detection result of the algorithm to the [0,1] interval and comparing it with the closeness of the binarized true value (0 or

1). The smaller the value is, the larger the value of the background area is, and the greater the difference between the two is (0.9 / 0.1 is better than 0.8 / 0.2), which indicates that the algorithm's detection results are more differentiated, and it is easier to find a simple threshold.

In the three databases, two formulas are used to calculate the comprehensive F value and MAE value of the existing algorithm and the proposed algorithm, respectively, and the comparison is displayed in the form of a histogram, as shown in Figure 3 and Figure 4.

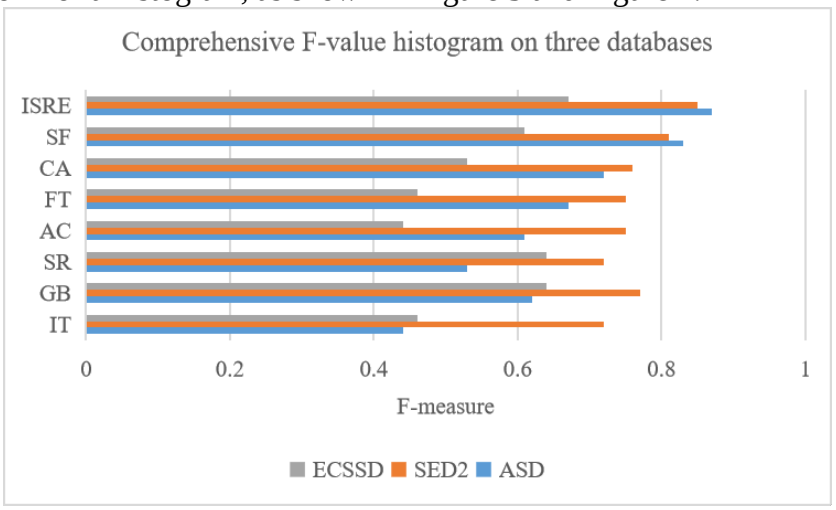


Figure 3: Histogram of comprehensive F-values on three databases

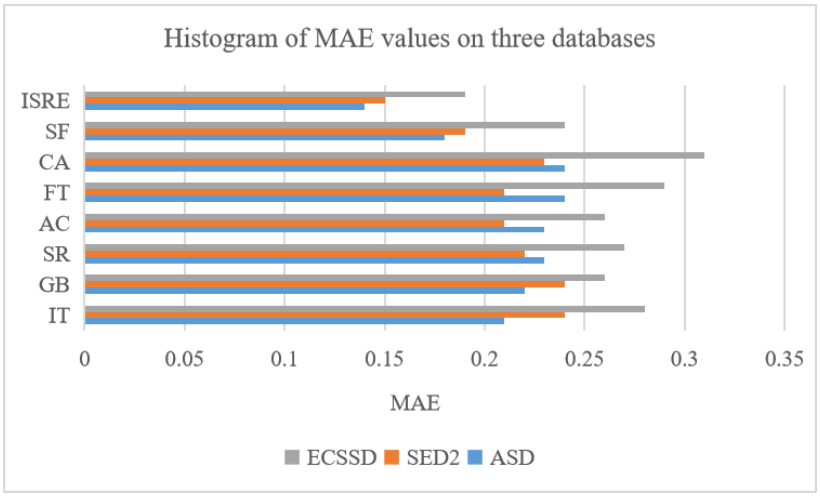


Figure 4: Histogram of MAE values on three databases

As can be seen from the comparison of the pictures in Figure 3 and Figure 4, the algorithm proposed in this paper has obvious advantages over the earlier proposed algorithms such as IT, GB, SR, etc., and performs consistently better than the original SF algorithm on the three databases, of which the comprehensive F value increased by 3.4% ~ 7.4%, and the MAE value decreased by 15.4% ~ 21.9%. Because PCNN is used to effectively suppress background interference, the more improved the proposed algorithm is, the more obvious the ECSSD database is for a large number of complex background images. In order to more intuitively show the evaluation results of the proposed algorithm and the SF algorithm, the specific performance of the three evaluation indicators of AUC value, F value, and MAE value are compared in Table 2.

Table 2: Comparison of AUC value, F value and MAE value index between ISRE algorithm and SF algorithm

	AUC			F-measure			MAE		
	ASD	SED2	ECSSD	ASD	SED2	ECSSD	ASD	SED2	ECSSD
ISRF	0.892	0.868	0.814	0.870	0.852	0.662	0.137	0.145	0.195
SF	0.874	0.855	0.736	0.839	0.823	0.616	0.175	0.181	0.230
%	2%	1.5%	10.6%	3.8%	3.4%	7.4%	21.9%	19.5%	15.4%

(2) Algorithm execution efficiency

In order to measure the real-time performance of each saliency algorithm in actual use, on a computer with a uniform hardware configuration (Intel Core i5-4590 3.30GHz quad-core CPU, 4G RAM), the performance of the algorithms proposed in this chapter and the seven saliency algorithms are compared. The processing speed is evaluated, that is, the average time required for each algorithm to process an image in the three standard image databases as a whole is shown in Table 3.

Table 3: Comparison of execution efficiency of different algorithms

	IT	GB	SR	AC	FT	CA	SF	ISRE
Time	0.283	0.651	0.046	0.137	0.083	43.52	0.192	0.247
Code	C	M+C	M	C	C	M+C	C	C

As can be seen from Table 3, the speed of algorithm execution is greatly affected by the programming language (C = C / C ++, M = Matlab) and the form of algorithm model construction. The CA algorithm model designs a large number of nearest neighbor searches and uses a mixed language, so the operation speed is very slow. Except the CA algorithm, the other algorithms generally have good operation speed, especially the SR and FT algorithms. The construction is simple. Although the detection result is not ideal, the execution speed is optimal, which can be applied to the occasions that do not require high detection accuracy. In general, the ISRE algorithm mentioned in this chapter is affected by the secondary operation of the PCNN model. The algorithm speed is lower than that of the SF algorithm.

4.3 Performance Analysis of Voice Interaction Technology Based on Intelligent Education Service Robot

This section analyzes the most important speech recognition rate, speech interaction stability, and system response time of the speech interaction system. The user's test of the robot's voice instructions is not a full test of a set of sentences, but a test of mixed voice commands and multiple consecutive identical voice commands. Each voice command is performed on-site through a microphone connected to the embedded main control board the results are collected and then processed by the system in real time and the predictions are performed. For the relatively quiet environment in the laboratory, each group of test voice instructions listed in Figure 5 will be divided into 4 groups for test statistics, which is also for analysis. The outdoor environment tests are grouped according to the school statistics. A total of 4 schools are tested. The number of test cases in each school is different, but they have reached hundreds.

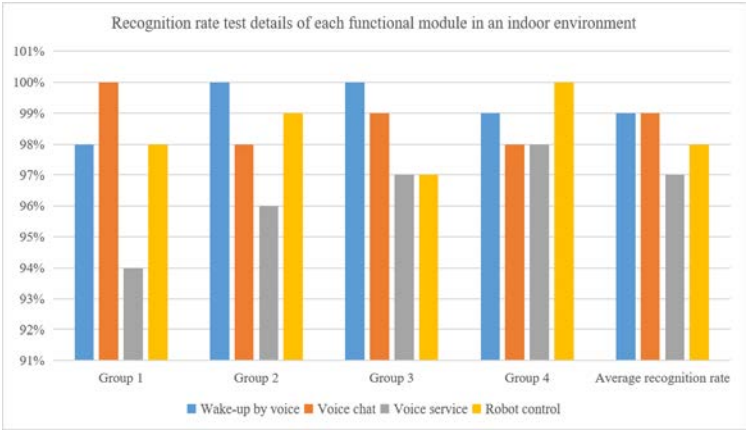


Figure 5: Recognition rate test details of each functional module in an indoor environment

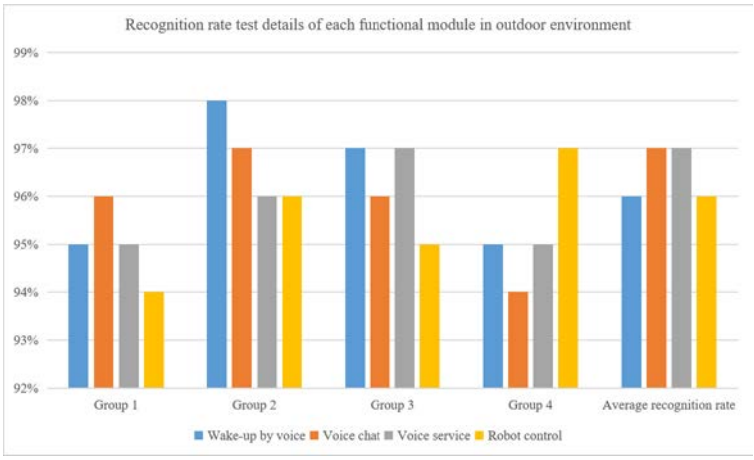


Figure 6: Recognition rate test details of each functional module in an indoor environment

From the test results of the two environments, it can be seen that the recognition rate of the outdoor environment is lower than the indoor environment, but it can still maintain a high recognition rate, but the sound of the outdoor environment is larger than the test voice instructions indoor that is, the sound volume is large. It is worth noting that the recognition function of the voice service in the voice service has a lower recognition rate than other songs. This is caused by the fact that the singer in the input voice has similar pronunciation to the song title or cannot find some unpopular songs. It can be seen that the system has good stability and can meet the needs of the service robot system. According to the data in Figures 5 and 6, there are some use cases with lower tests, which are mainly due to factors such as energy loss caused during system operation. The response time of the system is basically within 2 seconds. The high delay is mainly in the part of controlling the robot. Most modules have a response time of milliseconds. This is mainly due to the delay of the robot communication caused by the network delay and so on. Network optimization is especially important in real products.

5. Conclusions

In order to design the most convenient and accurate intelligent education service robot, this paper uses SF algorithm and improved PCNN to establish a hybrid model, and proposes an improved

saliency region extraction algorithm based on education service robot. The algorithm first takes the brightness feature map obtained by the SF algorithm as the input neuron of the PCNN, and outputs it as an internal neuron after the modulation and demodulation unit; then the dot multiplication result of the internal neuron and the binary initial saliency map generated by the SF algorithm as an input signal, the ignition pulse unit is improved to optimize the ignition range. By comparing the standard database with the real environment, this paper fully demonstrates that the binarized saliency area generated by the proposed algorithm is closer to the true value, which can effectively suppress the high-brightness background area in the detection results of the SF algorithm. The PCNN model with propagation stimulus as its core can more effectively simulate the biological vision system.

Combining compressed sensing technology, this paper proposes a speech recognition scheme that is easy to implement in hardware. The algorithm performance of the robot is verified, and the optimal effect parameters are selected through comparative experiments. This method uses Chinese phonetic phrase (sentence) tests to obtain good recognition results. It can be used as an effective improvement solution for the proposed robot voice interaction system's voice input, which can effectively reduce the energy loss caused by maintaining voice monitoring and the need for real-time Handle the shortage of storage resources caused by a large amount of voice data.

Acknowledgements

The project was supported by the Computer Subject Team Building Project, Huanggang Normal University.

References

[1] Paul Baxter, Emily Ashurst, Robin Read, James Kennedy, & Tony Belpaeme. (2017). Robot education peers in a situated primary school study: personalisation promotes child learning. Plos One, 12(5), e0178126.

[2] Catherine Lovegrove, Kamran Ahmed, Giacomo Novara, Khurshid Guru, & Prokar Dasgupta. (2016). Modular training for robot-assisted radical prostatectomy: where to begin?. Journal of Surgical Education, 74(3), 486-494.

[3] Paulo José Costa, Nuno Moreira, Daniel Campos, José Gonçalves, José Lima, & Pedro Luís Costa. (2016). Localization and navigation of an omnidirectional mobile robot: the robot@factory case study. IEEE Revista Iberoamericana De Tecnologias Del Aprendizaje, 11(1), 1-9.

[4] Maja Lutovac, Zoran Dimic, Stefan Mitrovic, Aleksandar Stepanovic, & Vladimir Kvrgic. (2016). Reconfigurable virtual environment for multirobot operations and its application in education. Telfor Journal, 8(2), 127-132.

[5] Marsha Langer Ellison, Erin D. Reilly, Lisa Mueller, Mark R. Schultz, & Charles E. Drebing. (2018). A supported education service pilot for returning veterans with posttraumatic stress disorder. Psychol Serv,15(2), 200-207.

[6] Hashem, Ibrahim Abaker Targio, Chang, Victor, Anuar, Nor Badrul, Adewole, Kayode, Yaqoob, Ibrar, & Gani, Abdullah. (2016). The role of big data in smart city. International Journal of Information Management,36(5), 748-758.

[7] Kan Zheng, Zhe Yang, Kuan Zhang, Periklis Chatzimisios, Kan Yang, & Wei Xiang. (2016). Big data-driven optimization for mobile networks toward 5g. IEEE Network, 30(1), 44-51.

[8] Xu, Zhenning, Frankwick, Gary L., & Ramirez, Edward. (2016). Effects of big data analytics

and traditional marketing analytics on new product success: a knowledge fusion perspective. Journal of Business Research,69(5), 1562-1566.

[9] Dimitris Berberidis, Vassilis Kekatos, & Georgios B. Giannakis. (2016). Online censoring for large-scale regressions with application to streaming big data. IEEE Transactions on Signal Processing, 64(15), 3854-3867.

[10] Hongming Cai, Boyi Xu, Lihong Jiang, & Athanasios V. Vasilakos. (2017). Iot-based big data storage systems in cloud computing: perspectives and challenges. IEEE Internet of Things Journal, 4(1), 75-87.