Real-Time Target Tracking and Localization in Augmented Reality Scenes with Virtual Reality-Based Computer Vision Technology

Linlin Zhai[1] and Kejun Wang[2]

[1]School of Science, Henan University of Engineering, Zhengzhou, Henan 450001, China, hjz@haue.edu.cn
[2]School of Science, Henan University of Engineering, Zhengzhou, Henan 450001, China, qhmao@haue.edu.cn

Corresponding author: Linlin Zhai, hjz@haue.edu.cn

Abstract. Augmented reality is a technology that synthesizes computer assisted virtual object information into the user's perceived real world. On the basis of studying real-time target tracking and positioning related technologies, this article studies a computer vision method based on virtual reality. In order to improve tracking and registration performance, this article focuses on the tracking and registration techniques in computer-aided augmented reality technology. We have implemented a landmark free tracking and registration algorithm based on natural feature points. Using computer vision image theory to solve registration problems in augmented reality and applying it to camera calibration. In order to extract the target text content from the video, the first problem is to determine the position of the corresponding text information in the image, that is, the position of the text. The focus of this study is on the localization and segmentation of images and characters in videos. A text region location method based on reaction diffusion equation is proposed. Wavelet moment feature, wavelet co-occurrence matrix feature and Local binary patterns feature are used to describe the texture of characters. Preprocess computer video images and combine them with SVM (Support Vector Machine) and K-means clustering methods to achieve computer-aided text segmentation. Research has shown that compared to comparative algorithms, the texture feature-based text region precise localization method proposed in this paper improves the recall rate by 6.831% and the false detection rate by 4.127%. Select text area images with edge density below average. That is to say, merging text regions with relatively simple backgrounds to achieve text enhancement. Compared with existing calibration methods, this method is more flexible and reliable.

Keywords: Virtual Reality; Computer Vision Technology; News Videos; Image Text Tracking; Augmented Reality Scenarios
DOI: https://doi.org/10.14733/cadaps.2023.S13.149-160

1 INTRODUCTION

Augmented reality (AR) is an emerging research field developed on the basis of computer-aided virtual reality (VR). In recent years, virtual reality has received extensive and in-depth research in various industries. Virtual reality refers to a computer assisted virtual environment that can provide people with various sensory stimuli. It emphasizes that from the perspective of sensory effects, users are completely immersed in the virtual environment, including visual, auditory, tactile stimuli, and even olfactory and gustatory senses. However, this precisely limits its application in certain aspects. Its goal is to enhance users' understanding of the real world through virtual information or scenes, rather than replacing the real world with virtual worlds. At the same time, factors such as computer assisted computing speed and scene modeling quality can affect the display of virtual scenes, and instead limit users' ability to observe and perceive the real world. Augmented reality technology emerged in this context. Fang et al. [1] conducted an object description of the target network tracking algorithm, which used template branching for network feature method search. By tracking targets in a 3D network, it performs feature extraction of individual target objects and network embedding module analysis. It utilizes a sub network for boundary box classification and obtains the final three-dimensional boundary box of the target object. Han et al. [2] conducted a visual framework construction analysis on product model design. Through validation testing of augmented/virtual reality 3D models, the competitive behavior of manufacturing environment patterns was analyzed.

The existing character recognition technology also provides a mature computer-aided technology foundation for extracting characters from videos and images. In terms of text retrieval, people have done a lot of work, and the technology of searching for corresponding content through keywords has become very mature. By utilizing the computer-aided advantages of natural language and mature text retrieval and search engine technology, users can easily retrieve the required video information from the Internet, which is in line with people's retrieval habits and relatively simple to implement. If computer assisted and video image processing related technologies can be utilized to automatically detect, locate, segment, and recognize these texts, it will be convenient and effective for post management and application expansion of news videos, such as news video classification, video content analysis and understanding, and comprehensive use of text and content for video retrieval. Therefore, this study has great practical value.

2 RELATED WORK

Lampropoulos et al. [3] designed a digital interaction method based on computer assisted semantic enhancement of the real world. By quickly obtaining logistics information and embedding it into appropriate space and practical frameworks, a scene localization system was constructed under a computer model. Li et al. [4] studied and analyzed a computer assisted automatic data synthesis attitude network, and established an instance segmentation network. Liu et al. [5] are dedicated to training computer algorithms for tracking and matching unlabeled mobile devices. Virtual image drift analysis was conducted by drawing and labeling the positions of augmented reality. Mahmood et al. [6] conducted computer-aided information retrieval analysis on augmented reality. Search for matching models in the planning model to enhance vision. Nafea et al. [7] studied the most suitable computer object monitoring algorithm for AR. Graphics power monitoring is carried out on traditional physical computer technology. Prit et al. [8] analyzed the stability concept of computer AR learning linear control through different design changes and reverse flow. Runji et al. [9] analyzed the auxiliary effect of computer augmented reality assistance technology on the manufacturing industry. Increase the implementation of intelligent physical environments by overlapping virtual information. Shewail et al. [10] analyzed the augmented reality information tracking of virtual world engines. According to the practice interval set by the computer, read different directional changes and perform augmented reality indoor tracking. The comparison of image monitoring algorithms and communication technologies highlights their advantages. CAD based virtual vision technology can simulate real-world reference objects. In order to find the target pairing standard parameters of the mock object more accurately, Sadeghi et al. [11] analyzed the tag tracking technology in the machine learning state. Its research has provided great help for the visualization standards of virtual environments.

Wei and Han [12] established a computer-aided central surveying and mapping system for information equipment management. The system adopts MVC mode to achieve the development of organizational functional module related technologies. In response to the problems of data loss and chaotic records in the past, computer informatization construction has been carried out, greatly improving the management efficiency of the laboratory. Younis et al. [13] proposed a real-time monitoring and tracking technology based on intelligent computer vision assistance. This system can monitor, classify, and track objects based on the motion compensation of the motion camera. Yue [14] proposed a more effective human motion description algorithm for feature monitoring and localization of different intelligent terminals. Zhao and Chen [15] analyzed the stereo matching problem of feature points in computer assisted real-time images. And by correcting and calibrating the image, better disparity images can be obtained. As users explore the medium of bidirectional remote applications, more and more computer-aided visual recognition technologies are being applied to the user experience. Zhou and Güven [16] proposed a fine-grained recognition method for computer-aided mobile augmented reality. This method utilizes the deep RGB characteristics of video sparsity to perform visual state analysis on various pose data. A special hardware program is used to monitor the Information set.

3 RESEARCH METHOD

3.1 Computer Assisted Augmented Reality Scene Image Analysis

The main issue in current research on computer-aided augmented reality systems is the registration. The key is to clarify the transformation relationship between the coordinate systems of virtual objects and real objects. Identifying a correct, fast, and robust computer-aided tracking and registration method is one of the keys to augmented reality technology. This is also a major challenge for the current development of augmented reality systems. So far, people have studied various methods such as computer-aided visual tracking, inertial tracking, optical tracking to assist real-time tracking, as well as their hybrid tracking methods and corresponding computer-aided algorithms.

From the implementation flowchart of a visual based augmented reality system, it can be seen that the computer-aided augmented reality system includes the implementation of four subsystems. That is, tracking and positioning systems, graphic image processing systems. The graphics and image processing subsystem analyzes and processes real-world scene images provided by tracking and positioning systems. After computer-aided processing, positioning information of cameras and real scenes can be obtained. At the same time, it can also provide the required virtual graphics for virtual reality fusion display systems. The subsystem displays computer assisted fusion images in front of users through display devices.

In order to enhance the display of augmented reality scene images under computer assistance, the minimum value in different directions is defined as the gray change of the pixel, that is, the E value. If it is still greater than a given threshold, the pixel can be considered as a corner. The value of E is calculated as follows:

$$E_W(x,y) = \sum_{u,v} W_{u,v} \left| f(u+x, v+y) - f(u,v) \right|^2 \tag{1}$$

$$E(x,y) = \min \; E_W(x,y) \tag{2}$$

W represents sliding window; $W_{u,v}$ represents the weight value at the inner point of the window, and $f(\cdot)$ is the gray value of the pixel (u,v); Respectively, the coordinate values relative to the pixel points in the sliding process of the window. Given the corner detection threshold T_E, if $E(x,y) > T_E$, the pixel (x,y) is considered as a corner.

Wavelet transform can detect transient phenomena in normal signals and show their components, and solve many difficult problems that Fourier transform can't solve. Wavelet decomposition can be understood as the inner product of the selected basic wavelet function $\psi(t)$ and the signal $x(t)$ to be analyzed under different scales α after displacement τ, namely:

$$WT_x(\alpha,\tau) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{+\infty} \psi^*\left(\frac{t-\tau}{\alpha}\right) dt, \ \alpha > 0 \tag{3}$$

WT_x is the wavelet decomposition of signal $x(t)$ at various scales, and $\psi^*(t)$ is the conjugate of $\psi(t)$.

Fuzzy clustering algorithm is considered as the generalization of standard clustering algorithm. Assuming that the feature vector $V = v_1, \cdots, v_N$ represents the data set, and the expected category is c-class, $2 \le c \le N$, the goal of the algorithm is to find the fuzzy classification set U containing N element data sets:

$$\begin{cases} U | u_{ik} \in [0,1] & \forall_{i,k} \\ \sum_{i=1}^{c} u_{ik} = 1 & \forall_k \\ 0 < \sum_{k=1}^{N} u_{ik} < N & \forall_i \end{cases} \tag{4}$$

Where u_{ik} represents the degree to which the feature vector f_k belongs to the i class. If a large amount of image data is only organized by the linear table structure, the retrieval time will increase linearly, and finding a suitable similarity measure can improve the retrieval efficiency. The basic framework of the retrieval system is shown in Figure 1.

The first goal of implementing augmented reality is to integrate the computer assisted virtual environment with the real scene around the user. Make users feel confident from a sensory perspective that the virtual environment is a component of the real scene around them. Accurate virtual reality registration technology is a key technology for implementing augmented reality systems. Computer assisted tracking technology is a key link in virtual reality registration, and calibration technology provides accurate data sources for computer tracking technology. Virtual reality registration is essentially a mapping transformation between the coordinate systems in the figure above. As long as the transformation matrix between the coordinate systems in the figure is obtained, the virtual object can be correctly integrated into the real scene.

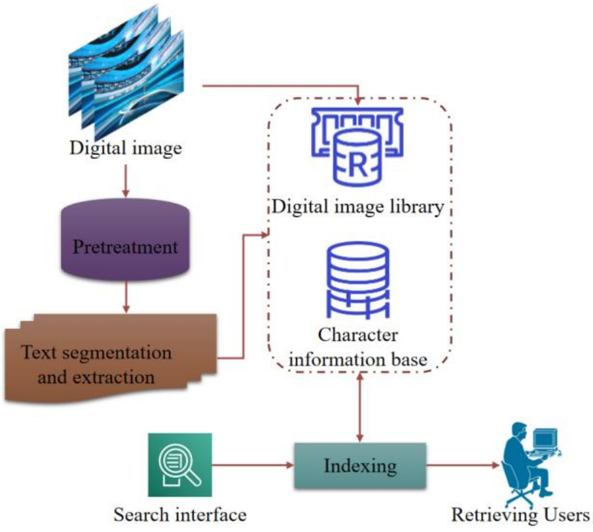

Figure 1: Basic framework of information retrieval system based on characters in image.

3.2 Computer-Aided Tracking Technology in Augmented Reality

The registration technology based on computer-aided visual tracking is built on the foundation of modern computer technology and image processing technology. The principle is to determine the camera's pose relative to the target in a real scene by analyzing and processing one or more images. This system does not require complex tracking devices, usually only one or two cameras. Due to the need to handle a large number of image processing operations. Therefore, it is easy to cause significant system latency, but with the continuous improvement of computer hardware technology, this problem has been greatly improved. Due to the simple tracking principle of computer-aided visual tracking technology, the system is easy to implement, and the tracking process directly processes images, with errors limited to the image plane space. Therefore, in current computer-aided augmented reality systems, vision-based tracking and registration techniques are often used.

Computer assisted visual tracking is based on modern computer technology and image processing technology. Computer assisted visual tracking refers to the use of technologies such as image processing. By detecting, extracting, recognizing, and tracking targets in image sequences, motion parameters such as position, velocity, acceleration, and trajectory of moving targets are obtained. In augmented reality systems, the purpose of computer-aided tracking is to process target features in image sequences. By utilizing changes in target features to obtain camera motion parameters, the determination of posture. Developing accurate, robust, and real-time tracking algorithms is a problem that we need to solve. The lack of logo information, occlusion, changes in lighting, and motion blur caused by rapid head movement all pose significant challenges to tracking problems. Unlike traditional tracking techniques in computer-aided vision, augmented reality tracking also includes establishing the pose of the tracked object. In this paper, the idea of robot visual servo is applied to augmented reality, and a tracking algorithm based on reaction diffusion equation is proposed.

There are many research results on the traveling wave solution of reaction-diffusion equation. The classical research method is to transform the traveling wave equation into the corresponding ordinary differential equation, and then use the phase plane analysis method to obtain its existence and uniqueness. For the reaction-diffusion equation with known existence of traveling wave solution, it is natural to consider the stability of its traveling wave solution.

$$\frac{\partial u(x,t)}{\partial t} = \Delta u(x,t) + f(u(x,t)) \tag{5}$$

Among them, the $f(u) = u(1-u), x \in R, t > 0$ model can describe the traveling wave phenomenon of alien invasive species or the propagation process of excellent genes of animals in one-dimensional infinite habitat.

In the real world, there are many phenomena that need to be described by partial differential equations. For example, the spatial structure plays a very important role in reflecting the spread of contact infectious diseases, not only because the space is non-uniform, but also because the object studied is moving in space.

A well-known law for calculating the flux of a flow is that the direction of the flow is always from high concentration to low concentration, that is, the direction of negative gradient, and the flux of its random movement at one point is approximately proportional to the local gradient of the individual quantity at this point, that is, it satisfies Fick's law, that is

$$J(t,x) = -D\frac{\partial}{\partial x}u(t,x) \tag{6}$$

Formulae (5) and (6) can be obtained:

$$\frac{\partial}{\partial t}u(t,x) = D\frac{\partial^2}{\partial x^2}u(t,x) + f(t,x) \tag{7}$$

Thus, a system described by the reaction-diffusion equation is deduced.

The location and extraction of title subtitles in news video images is essentially image segmentation. Compared with the background of news video images, the title subtitle area has rich edge information, so the title subtitle area position can be determined by using the edge information in edge detection images. Usually, news video images are RGB images. In order to simplify the operation, the formula (8) is used to convert them into grayscale images.

$$f(x,y) = 0.299 \times R(x,y) + 0.587 \times G(x,y) + 0.114 \times B(x,y) \tag{8}$$

After rough detection of potential text areas, it is necessary to further classify them and accurately locate the position of text lines. In the precise positioning step, text is considered as an object that satisfies a specific texture. In this paper, a method of text region location based on computer aided reaction diffusion equation is proposed. Wavelet moment feature, wavelet co-occurrence matrix feature and Local binary patterns feature are used to describe the texture of characters. Then, the isometric mapping algorithm is used to reduce the dimensionality of the feature vectors, and the reduced feature vectors are used as input features for SVM (Support Vector Machine) to classify candidate text regions.

The computer aided gray image shows different spatial gray distribution. This article uses the moment features of detail wavelet coefficients as texture features to classify text regions and background regions. The statistical characteristics of wavelet coefficients describe the overall distribution of a signal, but cannot reflect the local characteristics of the signal, especially the relationship between them. As is well known, the statistical tool used to extract second-order texture information from computer images is the co-occurrence matrix.

Energy (angular second moment):

$$F = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left[C(i,j|d,\theta)\right]^2 \tag{9}$$

When the elements in the co-occurrence matrix are concentrated, the F value is large, which indicates that the texture pattern is uniform and changes regularly.

LBP (Local Binary Pattern,) is a good texture description operator, which is simple in calculation, can represent the local features of the image, and has monotonicity of gray scale and rotation invariance. According to the different positions of pixels, the LBP value of this window is obtained, which is an integer in the range, and 256 values represent 256 unique texture patterns. The calculation formula of LBP operator is as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c) 2^n \tag{10}$$

For a $(P, R) = (8,1)$ neighborhood template, the encoding calculation process of lbp is shown in Figure 2:

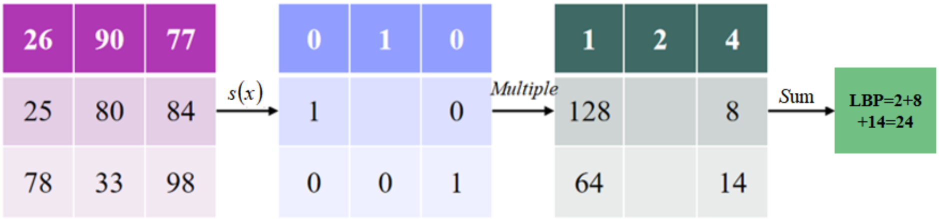

Figure 2: LBP calculation process.

4 ANALYSIS AND DISCUSSION OF RESULTS

All experiments in this paper are implemented on IntelCPU2.24G and 512MRAM microcomputers. The video images are collected by the camera, and the resolution of the video images is 1080×1080. The experiment of histogram segmentation is carried out on MATLAB. The background update rate β is the update rate of the reference background. If the background update rate β is set too large, the target will soon become the background, and the purpose of segmentation will not be achieved. If it is set too small, it will not be able to reflect the current background. In the actual experiment, the background update rate $\beta = 0.04$, the matching threshold $\delta = 25$.

This paper defines a kind of LRH (Local Region histogram), which adopts multiple horizontal and vertical projections from coarse to fine on the pyramid level of the original image and the image with half resolution, and synthesizes the results to determine the specific ranks of text blocks. Figure 3 and Figure 4 depict a newly generated LRH image compared with the processed edge image.

The text area forms a continuous area with high function value with a certain width in LRH diagram. C is a fine-tuning coefficient, which can be adjusted according to the statistical characteristics of different video sources.

According to the selected video clips in this paper, after a large number of experimental comparisons, the empirical values with better experimental results are selected, and $C = 0.8$ is selected for horizontal LRH images; Select $C = 0.4$ for vertical LRH chart. Finally, according to the size ratio of the text, a simple constraint condition is formulated to remove the marked rectangular box from the non-text area.

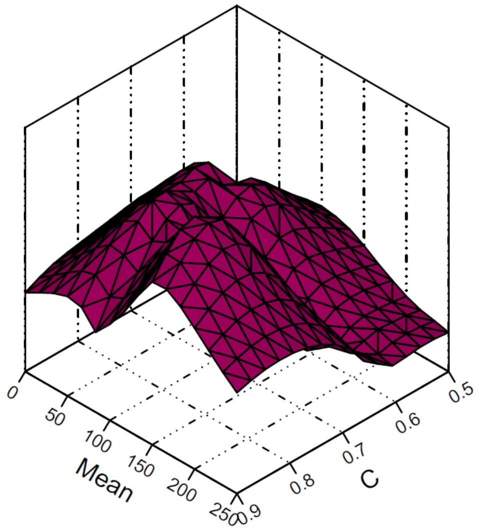

Figure 3: Original gray scale map.

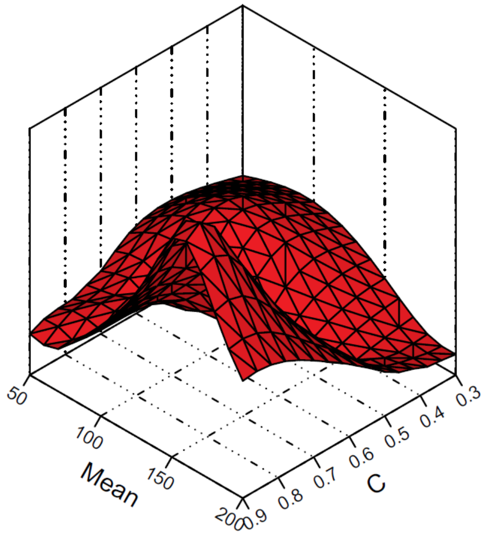

Figure 4: The processed edge map.

The internal parameters of a computer obtained from one image are significantly different from those obtained from another perspective. That is, using the same computer, the same calibration reference, etc. Therefore, if multiple image registration processes are considered simultaneously in the same minimization process, it is bound to improve calibration accuracy. Among them, some test images and positioning results are shown in Figure 5.

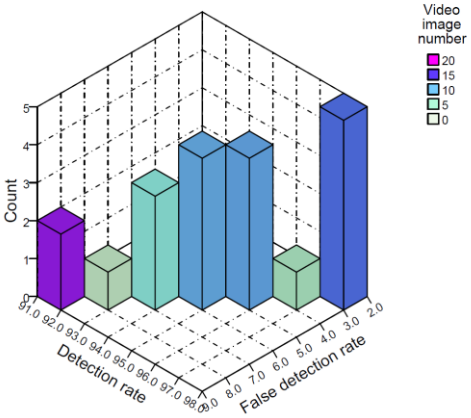

Figure 5: Histogram of false detection rate.

From the figure, it can be seen that the computer-aided character image localization method proposed in this article has a high detection rate and low false detection rate, and can accurately and effectively locate characters in complex backgrounds. Using the moving window, after removing the non-character part, the mathematical morphology operation is applied to the position of the character image area, and the image points of the obtained computer assisted character image are connected. Finally, locate the characters in the image by calculating the range of the connection domain.

In the experiment, the detection and localization results are divided into three categories: correctly located subtitle areas, incorrectly located subtitle areas, and incorrectly located non subtitle areas. For different news video images, Figure 6 shows the statistical data of computer-aided experimental results.

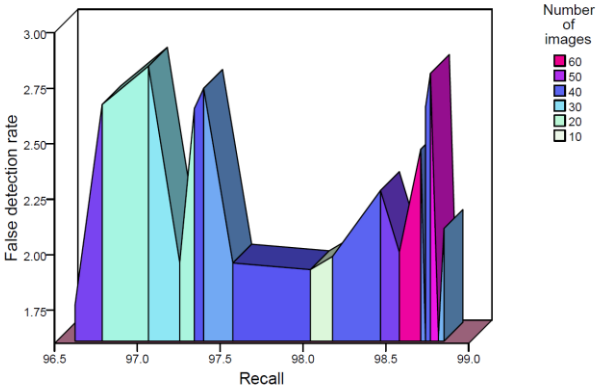

Figure 6: Statistical chart of recall rate and false detection rate.

Figure 6 shows the statistical chart. The computer assisted Sobel operator used in this article can only handle single channel images used for edge detection. Therefore, in general, we first need to extract the R, G, and B components of the subtitle detection result image, and cut the color multi-channel image into three single channel images. Calculate the edges of each component subgraph and select the largest edge subgraph to save. You can also judge based on the filtering effect and gradually increase the size until a satisfactory computer-aided filtering effect is achieved.

Since the sampled frame sequence is obtained through 0.5s sampling, the detector should also be able to detect static artificial characters in subsequent sampled frames. On the contrary, if the text area is not verified in the detection of subsequent sampling frames, it is considered an unexpected error detection.

As shown in Figure 7 and Figure 8, the statistical results of these two computer-aided algorithms are listed. Although the operation time of the algorithm in this article is relatively high, it has a high recall rate and a low error detection rate. Compared with the comparison algorithm, this paper proposes a text area accurate location method based on the reaction diffusion equation, which improves the recall rate and reduces the false detection rate by 6.831% and 4.127% respectively.

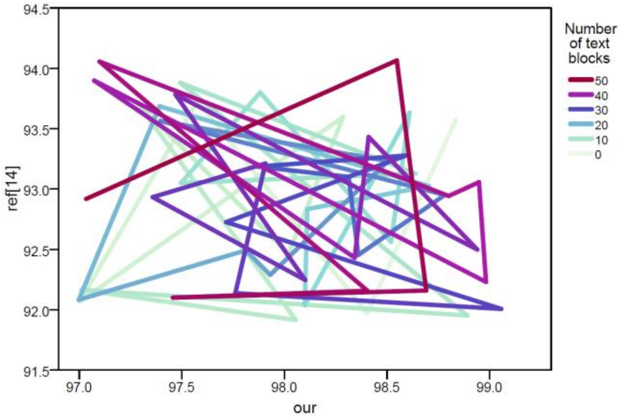

Figure 7: Statistics of recall and comparison results.

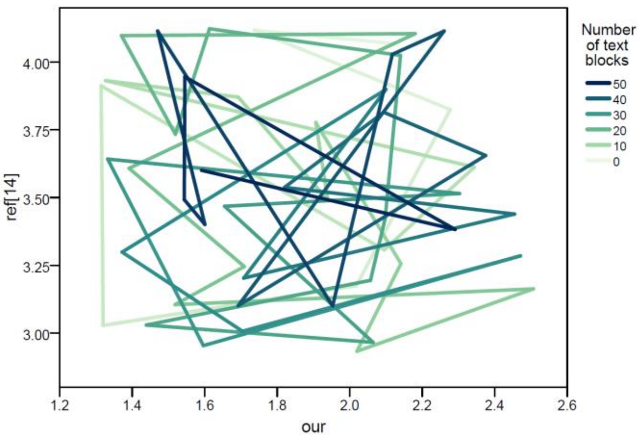

Figure 8: Statistics of false detection rate comparison results.

After introducing the texture characteristics of characters, SVM classifier eliminates a large number of non-text regions, reduces the false detection rate and keeps the hit rate. Finally, after adding the timing characteristics, the timing optimization discards short-term and moving characters. Therefore, the algorithm in this paper is very ideal.

These two methods also use the temporal redundancy information of video frames to achieve video text segmentation on the basis of multiple frame integration. Through Otsu binarization and connected domain screening of the enhanced image, the segmentation of characters and strokes is realized. Table 1 lists the performance comparison between the methods in this paper and those in other two literatures.

Algorithm	*Recognition accuracy /%*	*Recognition accuracy /%*
our	94.2995	91.7345
Literature [14]	92.585	87.8474
Literature [15]	93.2431	86.6872

Table 1: Performance comparison of three algorithms.

From the data results in the table, it can be seen that the recognition rate and accuracy of the proposed computational text segmentation method are higher than the other two methods. Because before the integration of multiple frames, text area images were classified using the edge density of computer assisted images. Select text area images with edge density lower than the average value, that is, fuse text areas with relatively simple backgrounds, to achieve the effect of text enhancement. The calibration objects used in traditional calibration methods need to have special accuracy and must have prior knowledge of the template. The computer-aided method proposed in this article has great flexibility and does not require any prior knowledge of templates. Reliable results can still be obtained without considering the initial position and internal parameters. This method can also be used for computer self-calibration and can measure internal parameters in real-time online.

5 CONCLUSION

The ultimate goal of computer-aided augmented reality systems is to achieve seamless fusion of real scene images and computer-generated virtual objects in the user environment. Accurate virtual reality registration technology is a key technology for implementing augmented reality systems, and computer-aided tracking technology is a key link in virtual reality registration. 80% of the information obtained by humans comes from vision, so vision-based tracking is widely used. Due to its high precision, high flexibility, and adaptability to various environments. The focus of this study is on the localization and segmentation of images and characters in videos. A text region location method based on reaction diffusion equation is proposed. Using the moment features of detail wavelet coefficients as texture features to classify text and background regions. Compared with the comparison algorithm, this paper proposes a computer-aided text area precise location method based on the reaction diffusion equation, which improves the recall rate and reduces the false detection rate by 6.831% and 4.127% respectively. After adding timing features, timing optimization discards short-term and moving characters. Therefore, the computer-aided algorithm proposed in this article is very ideal. Tracking and registration technology based on computer-aided vision is an important aspect of augmented reality research. Studying tracking and registration techniques based on computer-aided vision can enrich the implementation methods of augmented reality applications. It also provides a wider range of application environments and technical support.

Linlin Zhai, https://orcid.org/0009-0003-5790-0886
Kejun Wang, https://orcid.org/0009-0002-7433-5279

REFERENCES

[1] Fang, Z.; Zhou, S.; Cui, Y.; Scherer, S.: 3d-siamrpn: An end-to-end learning method for real-time 3d single object tracking using raw point cloud, IEEE Sensors Journal, 21(4), 2020, 4995-5011. http://dx.doi.org/10.1109/JSEN.2020.3033034

[2] Han, Y.-S.; Lee, J.; Lee, J.; Lee, W.; Lee, K.: 3D CAD data extraction and conversion for application of augmented/virtual reality to the construction of ships and offshore structures, International Journal of Computer Integrated Manufacturing, 32(7), 2019, 658-668. https://doi.org/10.1080/0951192X.2019.1599440

[3] Lampropoulos, G.; Keramopoulos, E.; Diamantaras, K.: Enhancing the functionality of augmented reality using deep learning, semantic web and knowledge graphs: A review, Visual Informatics, 4(1), 2020, 32-42. https://doi.org/10.1016/j.visinf.2020.01.001

[4] Li, X.; Cao, R.; Feng, Y.; Chen, K.; Yang, B.; Fu, C.-W.; Heng, P.-A.: A sim-to-real object recognition and localization framework for industrial robotic bin picking, IEEE Robotics and Automation Letters, 7(2), 2022, 3961-3968. http://dx.doi.org/10.1109/LRA.2022.3149026

[5] Liu, J.; Xie, Y.; Gu, S.; Chen, X.: A SLAM-based mobile augmented reality tracking registration algorithm, International Journal of Pattern Recognition and Artificial Intelligence, 34(01), 2020, 2054005. https://doi.org/10.1142/S0218001420540051

[6] Mahmood, B.; Han, S.-U.; Lee, D.-E.: BIM-based registration and localization of 3D point clouds of indoor scenes using geometric features for augmented reality, Remote Sensing, 12(14), 2020, 2302. https://doi.org/10.3390/rs12142302

[7] Nafea, M.-M.; Tan, S.-Y.; Jubair, M.-A.; Abd, M.-T.: A Review of Lightweight Object Detection Algorithms for Mobile Augmented Reality, International Journal of Advanced Computer Science and Applications, 13(11), 2022, 1-11. https://doi.org/10.14569/IJACSA.2022.0131162

[8] Prit, K.-D.; Mantri, A.; Horan, B.: Design implications for adaptive augmented reality based interactive learning environment for improved concept comprehension in engineering paradigms, Interactive learning environments, 30(4), 2022, 589-607. https://doi.org/10.1080/10494820.2019.1674885

[9] Runji, J.-M.; Lee, Y.-J.; Chu, C.-H.: User requirements analysis on augmented reality-based maintenance in manufacturing, Journal of Computing and Information Science in Engineering, 22(5), 2022, 050901. https://doi.org/10.1007/s40684-022-00444-w

[10] Shewail, A.-S.; Elsayed, N.-A.; Zayed, H.-H.: Survey of indoor tracking systems using augmented reality, IAES International Journal of Artificial Intelligence, 12(1), 2023, 402-414. https://doi.org/10.11591/ijai.v12.i1.pp402-414

[11] Sadeghi, N.-A.; Choi, S.-M.: A survey of marker-less tracking and registration techniques for health & environmental applications to augmented reality and ubiquitous geospatial information systems, Sensors, 20(10), 2020, 2997. https://doi.org/10.3390/s20102997

[12] Wei, X.; Han, Y.: Design and implementation of interactive visual communication teaching system based on CAD technology, Computer-Aided Design and Applications, 19(S7), 2022, 56-66. https://doi.org/10.14733/cadaps.2022.S7.56-66

[13] Younis, O.; Al-Nuaimy, W.; Alomari, M.-H.; Rowe, F.: A hazard detection and tracking system for people with peripheral vision loss using smart glasses and augmented reality, International Journal of Advanced Computer Science and Applications, 10(2), 2019, 1-9. http://dx.doi.org/10.14569/ijacsa.2019.0100201

[14] Yue, S.: Human motion tracking and positioning for augmented reality, Journal of Real-Time Image Processing, 18(2), 2021, 357-368. https://doi.org/10.1007/s11554-020-01030-6

[15] Zhai, L.; Chen, D.: Image real-time augmented reality technology based on spatial color and depth consistency, Journal of Real-Time Image Processing, 18(2), 2021, 369-377. https://doi.org/10.1007/s11554-020-00988-7

[16] Zhou, B.; Güven, S.: Fine-grained visual recognition in mobile augmented reality for technical support, IEEE Transactions on Visualization and Computer Graphics, 26(12), 2020, 3514-3523. https://doi.org/10.1109/TVCG.2020.3023635