Research Article

The Application of Computer Augmented Reality Technology in Ancient Murals

Mingming Zhang [ORCID]

Sejong University, Seoul 05006, Republic of Korea

Correspondence should be addressed to Mingming Zhang; 18403304@masu.edu.cn

In recent years, the country has given the green light to the development of augmented reality technology. Various preferential policies have sprung up such as mushrooms after rain, and augmented reality technology has gradually been applied to all aspects of our lives. In terms of cultural relic protection, the research and protection of ancient murals are undoubtedly not the most difficult. In terms of protection, murals are very fragile, and there are many factors that need to be paid attention to, such as air humidity, temperature, microorganisms, and the shedding of mud layers. In terms of research, the study of frescoes requires a lot of careful observation, but exhaled breath or body temperature that researchers get too close to can have unforeseen consequences. Based on this, the article expounds the composition of the augmented reality technology system. Through image recognition, three-dimensional tracking registration, virtual-real combination, and other technologies, the image obtained by the camera is recognized, and then the environment where the mural is located is a virtual scene through three-dimensional tracking technology. We build and finally use the combination of virtual and real technology to make the ancient murals vividly presented in front of us for our researchers to study more conveniently. The use of Gaussian function, SURF algorithm, FREAK algorithm, RANSAC algorithm, and so on to improve the augmented reality technology system makes the presentation of ancient murals more detailed and realistic. After experiments, we came to the conclusion that the ancient murals processed by augmented reality technology are far more detailed and more complete than the pictures taken by ordinary cameras. Processing ancient frescoes is also far less time-consuming than the existing image processors.

1. Introduction

Augmented reality technology is a kind of virtual scene image or information generated by computer science that can be vividly presented in the real scene so that the user can see through the terminal in real time so as to obtain more information [1–5]. Augmented reality technology can be applied to a variety of application fields through the support of three-dimensional tracking registration technology, virtual-real integration technology, and other technologies [6–9]. As early as the 20th century, augmented reality technology appeared in the United States. With the continuous development of science and technology, our country has also begun to use augmented reality technology in various fields, such as the military [10–12]. However, its application in the cultural field is still relatively small, and there is a lot of room for development. The ancient murals have high requirements

for preservation conditions, and the research conditions are also very harsh. It often occurs in the process of research due to improper operation or natural beauty. Precious ancient frescoes have since disappeared [13, 14]. Through augmented reality technology, all aspects of ancient frescoes are recorded and saved in a virtual scene built by a computer, and then, through three-dimensional tracking technology, they can appear in front of us in a complete and detailed manner for researchers to study [15]. The application of augmented reality technology is of great significance for the application and exploration of ancient murals.

2. Application Analysis of Computer Augmented Reality Technology

Thus, the system bridge of real scene, virtual scene, and user interaction is realized. It is generally composed of image

recognition technology, three-dimensional tracking and registration technology, and virtual-real combination technology. The composition mode diagram is shown in Figure 1.

In this augmented reality system, we first need the information needed to move from the real scene through the camera and analyse and establish a virtual-real scene space through image recognition technology. Then, the spatial positioning of the current camera and the specific posture maintained by the current camera are obtained by using the three-dimensional tracking technology. We use the computer to design the virtual model corresponding to the current state, and then according to the relevant data parsed and calculated by the three-dimensional tracking and registration technology, the real scene and the virtual scene are merged and then displayed in the virtual and real combination of the system. Finally, the customer needs the lesion to be displayed in the terminal and form a complete augmented reality system. The operation process of the image recognition technology is shown in Figure 2.

Because feature matching refers to establishing a corresponding relationship between the two images in the same real-world scene or with the same target, the anti-interference and accuracy of global feature detection are lower than those of local feature detection, and its environmental requirements for the image are very high, resulting in great limitations in the application, so the local feature detection method is adopted.

Using image recognition technology, through a series of processes of image acquisition, image preprocessing, image feature extraction, image feature matching, and output results, we can obtain more accurate and useful information than the image information obtained at the beginning. Remove the worthless information as our follow-on 3D registration tracking technology offers computational convenience and greater accuracy. The 3D tracking and registration technology is an essential technology for augmented reality technology to correctly realize superposition and virtual-real integration. The 3D tracking and registration technology is usually applied to the tracking and positioning of objects in real scenes, rendering virtual scenes to the real scene through the spatial coordinate system. The precise location of each corner of reality can be divided into three categories by the technology, equipment, and completion methods used. The most widely used tracking registration technology is the software-based registration technology, that is, the tracking registration technology developed through computer vision. This is by obtaining a spatial positioning and superposition model that can be converted between the virtual coordinate system and the real coordinate system. Whether the marker is placed in the real environment can be divided into marker three-dimensional registration technology and nonmarker registration technology.

After obtaining the target keyframe in the offline section, the camera's external reference and projection matrix of the keyframe are solved according to the matching point, and the initialization structure is established. The parameters of the target keyframe are saved during frame-by-frame analysis and the establishment of the target 3D structure.

Locate the keyframe settlement to initialize the 3D trace registration information.

Figure 3 shows the three-dimensional tracking registration technology.

2.1. Augmented Reality Computing Model

2.1.1. Coordinate Calculation in Three-Dimensional Space

(1) Pixel Coordinates vs Image Coordinates. For a certain point P_w on the three-dimensional space, the two-dimensional homogeneous coordinates of its projection point p_0 under the pixel coordinate system $(O_0 uv)$ can be represented by $p_0 = (u, v, 1)^T$, and the two-dimensional homogeneous coordinates of its projection point p_1 under the image coordinate system $(O_1 xy)$ can be represented by $p_1 = (x, y, 1)^T$. The relationship between two points can be expressed by formulas (1) (2), where (u_0, v_0) is the origin of the coordinate system.

$$u = \frac{x}{\Delta x} + u_0, \tag{1}$$

$$v = \frac{y}{\Delta y} + v_0. \tag{2}$$

In the above formula, the variation of each pixel of the pixel coordinate system on the x and y axes is the Δx sum Δy. On this basis, the transformation relationship of these two coordinates is shown in the following formula:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\Delta x} & 0 & u_0 \\ 0 & \dfrac{1}{\Delta y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{3}$$

(2) Image Coordinates and Camera Coordinates. From the coordinate system $(O_1 xy)$ to the coordinate system where the camera is located, $(O_C X_C Y_C Z_C)$ the points in p_1 the image coordinate system that can be represented by homogeneous coordinates $p_1 = (x, y, 1)^T$ are marked in the camera coordinate system $P_C = (X_C, Y_C, Z_C, 1)^T$. The distance note is f also commonly referred to as the focal length of the camera. Then, the transformation relationship between the image coordinate system and the camera coordinate system can be expressed by the following formula, which is expressed as follows:

$$Z_C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix}. \tag{4}$$

For a point in the real world coordinate system in three-dimensional space, P_W it can be represented by $P_W = (X_W, Y_W, Z_W, 1)^T$ the real world coordinate system
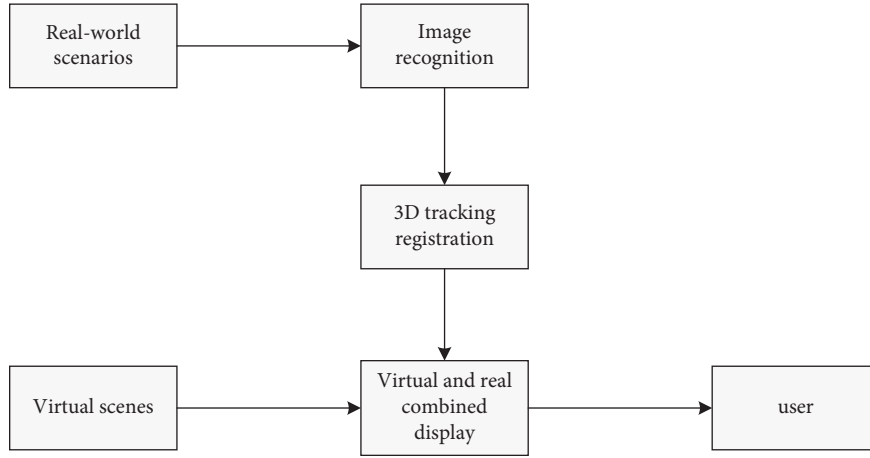
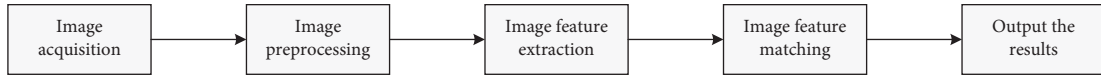FIGURE 1: System composition diagram of augmented reality technology.

FIGURE 2: Operation process diagram of image recognition technology.
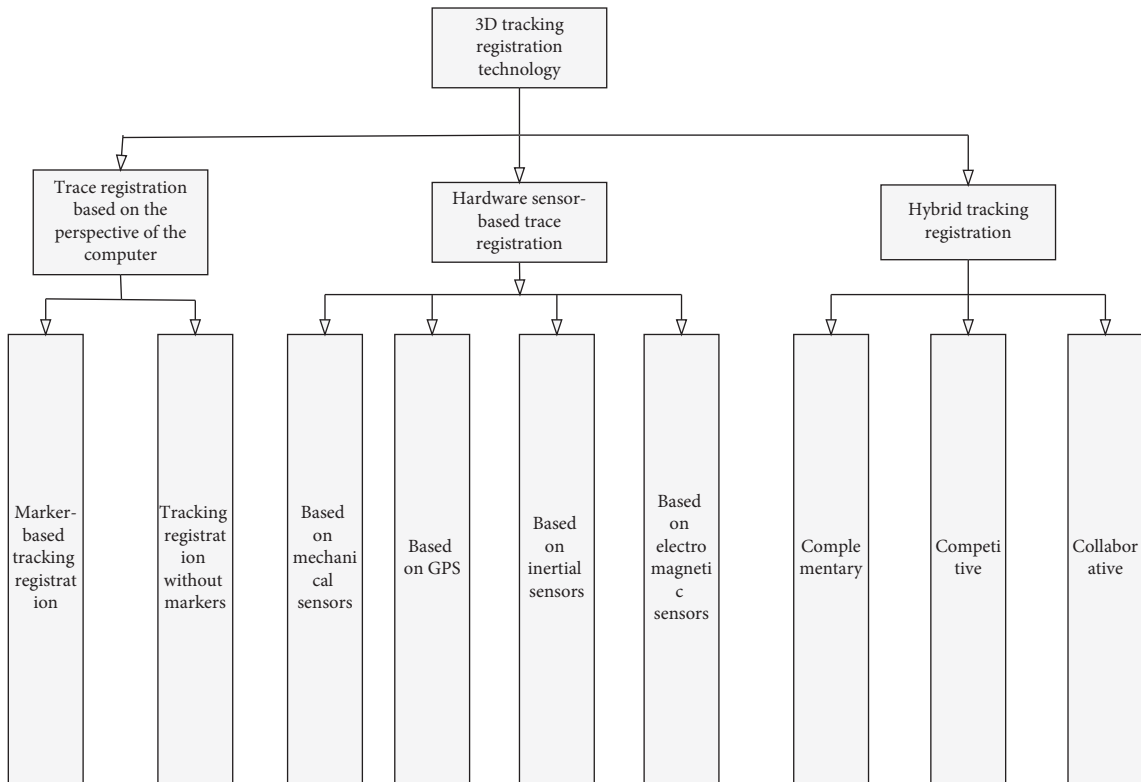

FIGURE 3: Three-dimensional tracking registration technology classification diagram.

$(O_W X_W Y_W Z_W)$. The points P_C in the camera coordinate system $(O_C X_C Y_C Z_C)$ can be represented by homogeneous coordinates $P_C = (X_C, Y_C, Z_C, 1)^T$. The transformation relationship between these two coordinate systems can be expressed by the following formula:

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = T \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}. \tag{5}$$

The R in the above formula which refers to the coordinate axis is a 3×3 rotation matrix when the real world coordinate system is transformed to the camera coordinate system. t is the vector that represents the coordinate translation after changing from the real world coordinate system to the camera coordinate system. T is the matrix used to represent the transformation relationship between the real-world coordinate system and the camera coordinate system. This transformation relationship can also be expressed using the following formula:

$$Z_C \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\Delta x} & 0 & u_0 \\ 0 & \dfrac{1}{\Delta y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \tag{6}$$

$$= K \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = KT \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = M \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}.$$

2.1.2. Feature Point Detection Calculation. A scale is often used to describe zooming in on a distant object, and the image of the distant object becomes blurry. This is often used to simulate human eye vision. When the feature points of two objects on the image can be matched on many scales, it can be proved that the two are the same object.

In order to achieve the invariance of the algorithm scale, we established an image pyramid to build the scale space. Each layer of the pyramid is made up of numerous images, and the scales between the images increase by a constant multiplier. The starting image of each group is doubled by the penultimate third-level image sample of the next group of images, and the image sizes between the different groups form a continuous pyramid, thus ensuring the invariance of the scale.

(1) Gaussian Function. The scale space function of the image can be expressed $G(x, y, \sigma)$ by the convolution of the scale-variable Gaussian function $L(x, y, \sigma)$ and the input image as $I(x, y)$, as shown in the following formula:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y). \tag{7}$$

$*$ represents convolution, and the variable-scale Gaussian function $G(x, y, \sigma)$ can be represented by the following formula:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-x^2+y^2/2\sigma^2}. \tag{8}$$

It σ represents the scale factor, and its size determines the blurriness of the image. The feature point detection in this

paper is the feature point detection of the SURF algorithm, which is interest point detection based on the Hessian matrix. Assuming that there is a pixel p in a $p(x, y)$ given image I, and the Hessian matrix definition at this point and scale σ can be expressed by the following formula:

$$H(p, \sigma) = \begin{bmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{bmatrix}. \tag{9}$$

Here, L_{xx}, L_{xy}, L_{yy} are the convolution of the Gaussian second derivative and the original image. Because the computational complexity of the Gaussian function is also slow, we usually use approximate values to represent it, reducing the amount of calculation to speed up its formula as shown in the following formula:

$$\text{Det}(H) = D_{xx}D_{yy} - (\omega D_{xy})^2. \tag{10}$$

Among them, D_{xx}, D_{yy}, D_{xy} represents L_{xx} the L_{xy} approximate value obtained by filtering the corresponding values of L_{yy}

In fact, there are many algorithms for feature detection of target images, which are selected according to different features and different occasions. The SURF algorithm has a high advantage over other algorithms in terms of the accuracy, scale, and rotation of feature point detection, and the algorithm has also been greatly improved in speed, so we chose the SURF algorithm as the main algorithm of the augmented reality system.

(2) SURF and F REAK Algorithms. In the past, we used the traditional SURF algorithm to describe floating-point features. Although the hot SURF has been updated and accelerated, it still cannot meet our requirements. It needs to use a lot of memory and time for processing, describing, and matching. Therefore, we use the FREAK algorithm, which is closer to the process of the retina of our human eye receiving image information and processing it. It uses the gradient calculation of the sampling points in parts, and the sampling points are combined into pairs by comparing the sampling points. The pixel intensities to get our descriptors can be expressed by the following formula:

$$G = \sum_{0 \le a < N} 2^a T(P_a). \tag{11}$$

N in the formula represents the length of the descriptor, in which $T(P_a)$ represents the binary test, where P_a represents a large number of sampling points. The expression can be represented by the following formula:

$$T(P_a) = \begin{Bmatrix} 1, \cdots\cdots I(P_a^{r_1}) - I(P_a^{r_2}) > 0 \\ 0, \cdots\cdots\cdots\cdots\cdots \text{otherwise} \end{Bmatrix}, \tag{12}$$

wherein $I(P_a^{r_1})$ represents the pixel value of the previous pixel point P_a and $I(P_a^{r_2})$ represents the P_a pixel value of the latter point.

According to the existing feature point information, a matrix is established, each row corresponds to a key point FREAK binary character representation, and the variance of the binary distribution matrix is calculated by calculating the mean of each column of the matrix. Sort each column of the

matrix according to the maximum variance, with the nearest variance ranking the farther in front and the less close to the one behind. Finally, save the best column, iterate over the remaining columns to calculate the covariance of the best column, leaving the column with the smallest covariance, and finally filter to get the desired binary descriptor.

In this model, the vector direction of the feature points is calculated by selecting the sampling points with center symmetry and using the local gradient summation method. We assume that O the local gradient information is calculated by the formula (13) to obtain the direction of the feature points. The formula is as follows:

$$O = \begin{bmatrix} 0_x \\ O_y \end{bmatrix} = \frac{1}{M} \sum_{P_0 \in G} \left(I\left(P_0^{r_2}\right)\right) \frac{P_0^{r_1} - P_0^{r_2}}{P_0^{r_1} - P_0^{r_2}}. \quad (13)$$

In this formula, M is the number of points in the set G, G is the set of all sampling points used to calculate the direction of the feature point, $P_0^{r_1}$ and $P_0^{r_2}$ are points passing through the sampling center, $I(P_0^{r_1})$ denotes a pixel value of a preceding one of the two passing sample points. $I(P_0^{r_2})$ denotes a pixel value of a following one of the two passing sample points.

After sampling, we need to make a rough match of the collected data, which is called rough matching. The metric of rough matching is usually Hamming distance, and the formula for calculating Hamming distance between feature point descriptors is shown in the following formula:

$$\text{Ham}\left(\text{de sc}_1, \text{de sc}_2\right) = \sum_0^k x_i \oplus y_i. \quad (14)$$

3. An Improved Model of Augmented Reality Technology Based on the RANSAC Algorithm

It should be due to the existence of the threshold that the points that can be matched are only close to the true and correct match, and the correctness of the points cannot be fully guaranteed. Therefore, we decided to use the RANSAC algorithm in the last step of the feature detection and matching algorithm process, that is, use the random sampling consensus algorithm to screen and purify the obtained feature matching points. This algorithm can help us estimate the mutual transformation relationship between the coordinate points of the feature points between the two images, which can be called a homography matrix H. The homography matrix can help us find the specific location of the image in both the virtual scene and the real scene. The homography relationship between images is shown in the following formula:

$$s \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = H \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad (15)$$

s The scale parameter representing the matched feature point (x, y) is the matched feature point coordinates of the target

image and the matched feature point coordinates in (x', y') the virtual scene or the real scene.

The robustness test and the accuracy of the image recognition algorithm are used to compare the feature detection and feature matching algorithms to accurately describe the accurate description and matching of features between images, in which AM measures the average matching number and the average matching rate. The average matching rate represents the ratio between the total number of matching points N and the total number of image feature points M. The higher the ratio is, the more accurate the feature points detected by the algorithm are, as shown in the following formula:

$$\text{AM} = \frac{N}{M}. \quad (16)$$

The average matching rate AP is the ratio between the total number of exact matching points and the total number N_c of matching points N of feature points. The ratio represents the matching accuracy of the algorithm. Since only the correct number of matching points can be called the number of feature matching points, the higher the ratio, the higher the accuracy of the algorithm, as shown in the following formula:

$$\text{AP} = \frac{N_c}{N}. \quad (17)$$

In a system with augmented reality technology, both the establishment of 3D structures and the analysis of 3D registration information require specific information about camera fixation. In a flat image, if the real-world coordinates and the plane on which the camera coordinates are located coincide, the conversion relationship between two-dimensional plane coordinates is shown in the following formula:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \quad (18)$$

If it H is a homography matrix, the form of the formula changes as shown in the following formula:

$$H = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix}. \quad (19)$$

Because R the transposed matrix of is an orthogonal matrix, then, according to the correlation properties of the orthogonal matrix, there is a formula (20) as follows:

$$r_1^T r_2^T = 0, r_1^T r_1 = r_2^T r_2 = 1. \quad (20)$$

Combining formulas (19) and (20) together results in the formula as follows:

$$h_1^T K^{-T} K^{-1} h_2 = 0, h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2. \quad (21)$$

The first two frames selected are used as the basis to construct the initial 3D structure. Name the planes of the two frames as I_1 and I_2, assuming that p the projection points of a point outside the two planes on the two planes are, respectively, the p_1 homogeneous coordinates of the image

pixels of these two $\overline{p_1}, \overline{p_2}$ points p_2. p The spatial coordinates of the point are p homogeneous coordinates $\overline{p} = (p^T, 1)^T$, and the camera parameter matrix is K. According to this information, we can obtain the formula as follows:

$$\lambda_1 \overline{p_1} = K[I|0]\overline{P}, \lambda_2 \overline{p_2} = K[R|t]\overline{P}. \tag{22}$$

R sum is t obtained by the translation of the coordinate system rotation matrix sum of t the second frame through the first frame. By combining the two formulas in the formula (2) R2), the formula can be obtained as follows:

$$\lambda_2 K^{-1} \overline{p_2} = \lambda_1 RK^{-1} \overline{P_1} + t. \tag{23}$$

The left and right sides of the formula (23) are simultaneously cross-multiplied by t, and the formula (24) can be obtained by representing the matrix form of the cross multiplication by $\hat{t}$ as follows:

$$\frac{\lambda_2}{\lambda_1} \hat{t} \left(K^{-1} \overline{p_2} \right) = \hat{t} RK^{-1} \overline{p_1}. \tag{24}$$

The projection matrix is obtained M after passing the formula $\lambda_2 x_2 = \lambda_1 Rx_1 + t$, and the three-dimensional coordinates are multiplied by both sides of the reequation at the same time x_2 can get $0 = \lambda_1 x_2 Rx_1 + x_2 t$ any pair of feature points obtained are then paired with coordinates, the rotation matrix R and the translation vector t, which can be obtained λ_1 and λ_2.

By simplifying the original formula by defining the matrix F and matrix E and setting it as the homogeneous coordinates of the image, $xx = K^{-1}\overline{p}$ our final formula can be obtained as follows:

$$\overline{p_2}^T F \overline{p_1} = x_2^T E x_1 = 0. \tag{25}$$

At $\overline{P_2}^T K^{-T} \hat{r} RK^{-1} \overline{P_1} = 0$ Both sides of the equation are with at the same time $K^{-1}\overline{p_2}$ multiply the dots, remove the constant term to get: $E = K^{-T}\hat{r}$ In defining the base matrix and the essential matrix E are: $E = \hat{t}R$ finally bring in the formula (24) and you get (25)

4. Experimental Simulations

An important indicator of augmented reality technology is the average number of matched feature points, the average matching rate, and the average accuracy rate in the feature detection experiment. These data are very important for augmented reality technology. Through many different transformations such as rotation transformation, blur transformation, brightness transformation, and scale transformation of multiple scenes and one scene, we can obtain the following data. Comparing the data before and after optimization, we can visually view the optimized results, as shown in Table 1 and Figure 4.

From the data comparison chart, we can clearly see that in the various transformations of the image, the average number of matches obtained by the SURF algorithm and the FREAK algorithm in different transformations is quite different, and it will be better for a while. Compared with the previous two algorithms, the average

Table 1: Comparison results of the average number of matches under various transformations.

Transform name	SURF	FREAK	The method of this paper
Scaling	652	387	578
Rotation transformation	813	1000	902
Perspective change	673	874	768
Luminance transformation	637	978	735
Fuzzy transformation	1146	749	1089
Compression transform	600	937	723

number of matches obtained in this method has less variation and is more stable.

In the above changes, in addition to the average number of matches, we can also obtain an important indicator, the average feature point matching rate, which is referred to as the average matching rate. The changes in the average matching rate in the above transformation are shown in Table 2 and Figure 5.

From the data map, we can see that compared with the method in this paper and the SURF algorithm, the FREAK algorithm has a low average matching rate in the transformation of various images, and the robustness test results are also poor. The average matching rate of the method in this paper and the SURF algorithm are not much different in general, and the effect is stronger in the potential angle transformation and the compression transformation, where the average matching rate is higher. On the whole, the average matching rate of this method is the highest.

In the reenhanced color system, the error mainly occurs in the tracking registration error caused by the relative motion between the camera and the scene during the online tracking registration, which can be calculated by reprojecting the error err to be measured instead err

The calculation formulas are as follows: $err = 1/n \sum_{i=1}^{n} p_{mi} - p_i$ thereinto n represents the number of feature point pairs currently calculated for camera pose, p_i represents the coordinates of a two-dimensional feature point of a scene image, p_{mi} represents projected feature coordinates.

Finally, we make an intuitive comparison through the data of the average accuracy rate in each transformation. The results are shown in Table 3 and Figure 6.

According to the experimental data graph, it can be known that the traditional FREAK algorithm has a lower overall accuracy rate than other algorithms, and the robustness test results cannot achieve the desired effect. While the SURF algorithm has lower accuracy in the case of rotation transformation, it achieves better results in other image transformations and has higher accuracy. On the whole, the method in this paper is the best, followed by SURF. Finally, after a series of experiments, the pictures of ancient murals we hope to get are shown in Figure 7.

4.1. Time Consumption Comparison of Experimental Simulation with Different Algorithms. Time consumption is a key indicator to measure the real-time performance of an algorithm. By comparing the running time of the algorithm in
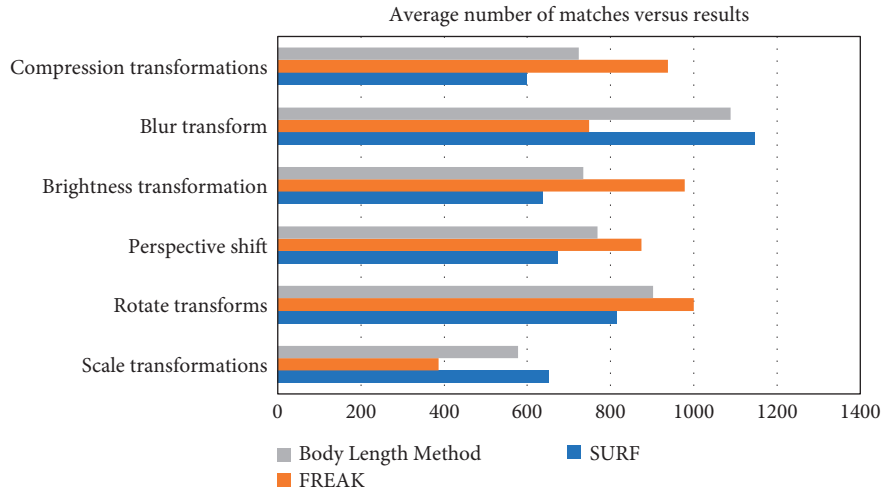
Figure 4: Average match count comparison results.

Table 2: Average match rate comparison results under various transformations.

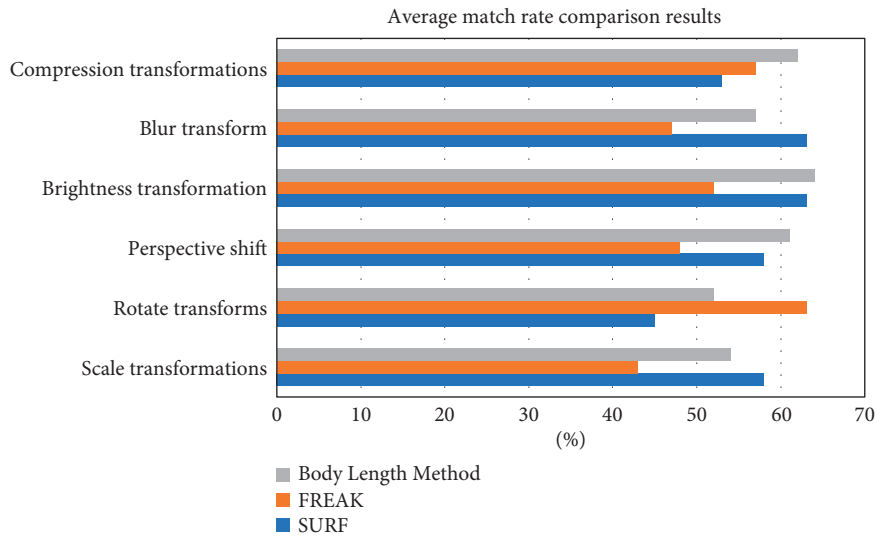| Transform name	SURF (%)	FREAK (%)	The method of this paper (%)
Scaling	58	43	54
Rotation transformation	45	63	52
Perspective change	58	48	61
Luminance transformation	63	52	64
Fuzzy transformation	63	47	57
Compression transform	53	57	62

Figure 5: Comparison of average match rates.

Table 3: Comparison results of average accuracy under various transformations.

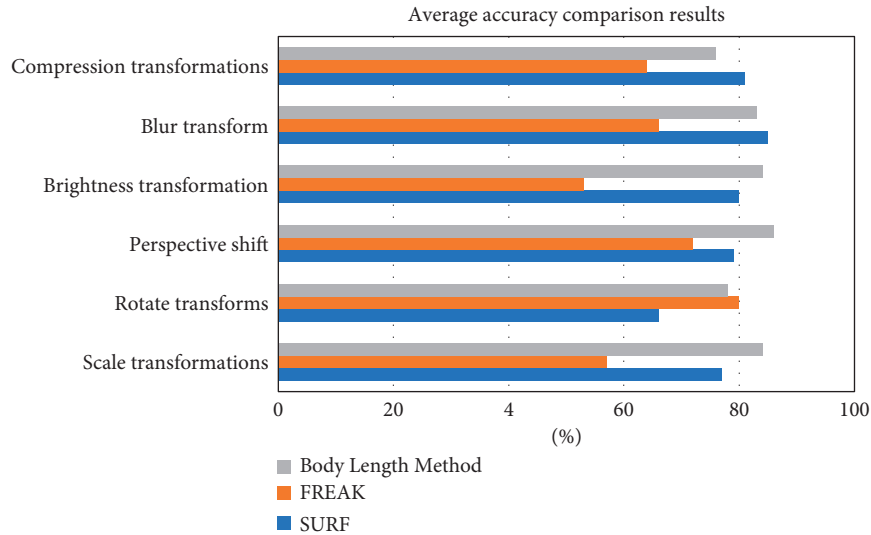| Transform name	SURF (%)	FREAK (%)	The method of this paper (%)
Scaling	77	57	84
Rotation transformation	66	80	78
Perspective change	79	72	86
Luminance transformation	80	53	84
Fuzzy transformation	85	66	83
Compression transform	81	64	76

Figure 6: Average accuracy comparison results.

Figure 7: Experimental map of ancient frescoes.

feature point detection and feature point matching and purification, the less time-consuming, the better the real-time performance of the algorithm.

By using the average data in the image database to calculate the time-consuming time of using the SURF algorithm and the FREAK algorithm and the time calculated by the feature detection method in this paper, the results are shown in Table 4 and Figure 8.

Compared to the three methods, the SURF method takes the most time. Compared with the FREAK algorithm, the SURF algorithm has better data such as the average accuracy rate and average matching rate, but each experiment takes more time than the FREAK algorithm. Since the improvement of the SURF algorithm has a long bottleneck, we can choose the method in this paper, which takes less time and is more accurate.

The time-consuming nature of 3D registration is an indicator of our evaluation of registration methods. We use the target recognition registration method and the target tracking registration method to compare the experimental results and obtain the following data as shown in Table 5 and Figure 9.

Through the time-consuming results obtained by the frame-by-frame feature detection and matching method, we can find that the time-consumption of target recognition and registration is much more than that of target tracking registration. Because the more time-consuming it is, the greater the load on the system and the more the system will be stuck. Therefore, the target recognition registration method is not feasible.

The data are our standard test library using the PASCAL VOC2007 image library as the image feature, which contains a large number of scenes and transformations, and we also calculate the image data sequence in this image library on average. The algorithms used are also two-way matching algorithms that combine the ratio of nearest neighbors and neighbors, which make the experimental data convincing and representative.

Finally, in the system constructed according to the augmented reality technology, we compare the speed of processing image frames of the main modules in the system, and the data obtained are shown in Table 6 and Figure 10.

TABLE 4: Time consumption comparison of each algorithm.

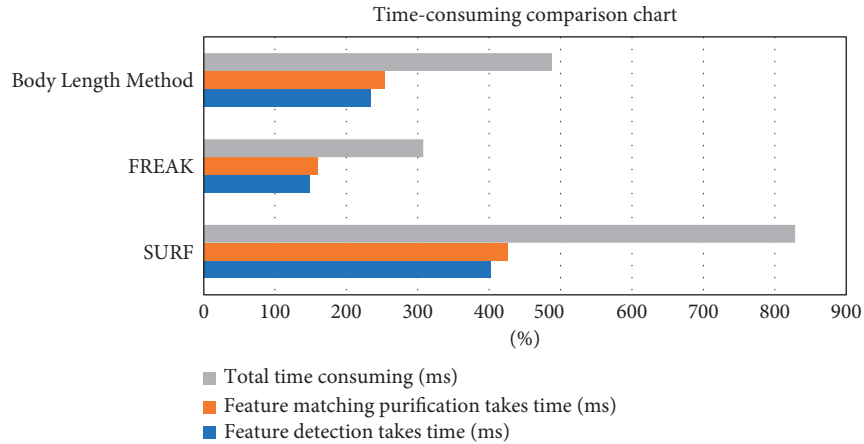| Algorithm	Feature detection time (ms)	Feature matching purification time (ms)	Total time (ms)
SURF	402.6	425.8	828.4
FREAK	147.5	160.2	307.7
The method of this paper	234.4	253.7	488.1

FIGURE 8: Time consumption comparison chart.

TABLE 5: Time consumption comparison of identifying registrations and tracking registrations.

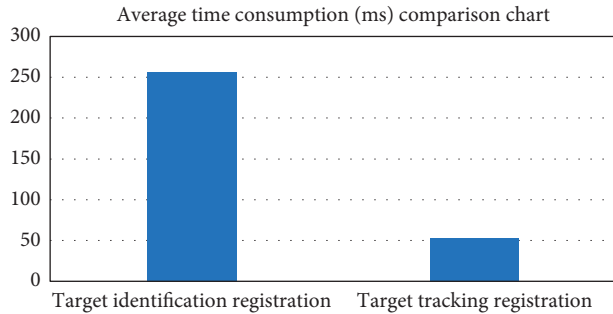| Process	Average time (ms)
Target recognition registration	256.6
Target tracking registration	53.1

FIGURE 9: Average time consumption comparison charts.

TABLE 6: The time consumption table of the number of frames processed by the three modules.

Scene image frame	Target image recognition (ms)	3D tracking registration (ms)	Fusion display of virtual and real (ms)
Initialization frame	238.3	18.3	23.6
Online tracking frame	42.6	10.5	11.2

According to the table, the three modules take the longest time to process the initial frame, and target image recognition is the most time-consuming one. At this time, we need to maintain the stability of the camera to ensure the integrity of the initial frame because the system uses augmented reality technology to superimpose the image solution bits by using feature tracking, thereby maintaining the fast registration of the system. The effect can be seen from the processing time of the online tracking frame on the table. The rendering of the ancient murals after quick processing is shown in Figure 11.
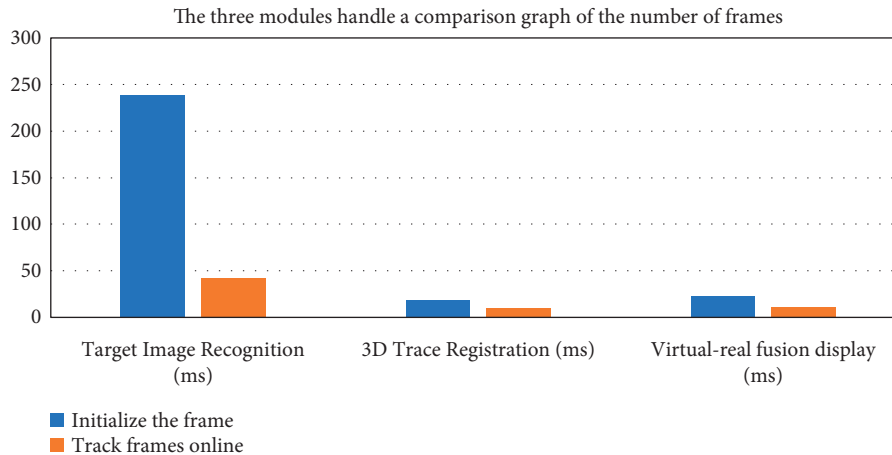
The three modules handle a comparison graph of the number of frames

■ Initialize the frame
■ Track frames online

FIGURE 10: Time-consuming comparison chart of the number of frames processed by the three modules.
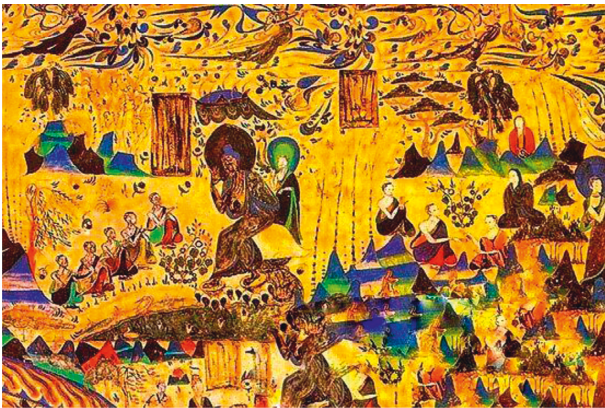

FIGURE 11: Fast processing of ancient frescoes.

5. Conclusion

Augmented reality technology is an emerging discipline that can help us express information more conveniently and intuitively through technologies such as image recognition, 3D tracking, and the combination of virtual and real. The building is then integrated with the real scene through image recognition and three-dimensional registration tracking and finally displayed on the terminal to help people explore the application of ancient murals. Among them, through the optimization method combining the scale space feature detection algorithm SURF algorithm and the binary sweep algorithm FREAK and through feature matching, the method presented in this paper combines the nearest neighbor ratio and two-way matching, which can ensure accuracy and reduce the calculation time at the same time. The 3D registration tracking module performs registration by establishing a targeted 3D structure and PnP algorithm, then settles the registration information as quickly as possible through the Gaussian function, and then changes the observation conditions of the virtual scene and the real scene. Use accurately superimposed virtual objects to achieve augmented reality effects.

Data Availability

The experimental data used to support the findings of this study are available from the author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this work.

References

[1] S. Matsutomo, T. Miyauchi, S. Noguchi, and H. Yamashita, "Real-time visualization system of magnetic field utilizing augmented reality technology for education," *IEEE Transactions on Magnetics*, vol. 48, no. 2, pp. 531–534, 2012.

[2] S. F. Javier, G. G. Jaime, and D. Laura, "Design and implementation of a GPS guidance system for agricultural tractors using augmented reality technology," *Sensors*, vol. 10, no. 11, pp. 10435–10447, 2010.

[3] M. Ivanova, "Enhancement of learning and teaching in computer graphics through marker augmented reality technology," *International Journal of New Computer Architectures and Their Applications*, vol. 1, no. 1, pp. 176–184, 2011.

[4] M. Vilkoniene, "Influence of augmented reality technology upon pupils' knowledge about human digestive system: the results of the experiment," *US-China Education Review A*, vol. 6, no. 1, p. 8, 2009.

[5] J. Chen, Y. T. Wang, and J. D. Lin, "Digital reconstruction of yuan ming yuan based on augmented reality technology," *Journal of System Simulation*, vol. 22, no. 2, pp. 424–428, 2010.

[6] M. Januszka and W. Moczulski, "Machinery designing aided by augmented reality technology," *Computer Assisted Mechanics and Engineering Sciences*, vol. 14, no. 4, pp. 621–630, 2007.

[7] J. Byeongki and Y. Janghyeok, "Competitive intelligence analysis of augmented reality technology using patent information," *Sustainability*, vol. 9, no. 4, p. 497, 2017.

[8] R. R. Rasalingam, B. Muniandy, and R. R. Rasalingam, "Exploring the application of augmented reality technology in early childhood classroom in Malaysia," *IOSR Journal of Research & Method in Education*, vol. 4, no. 5, pp. 33–40, 2014.

[9] J. Vaughn, M. Lister, and R. J. Shaw, "Piloting augmented reality technology to enhance realism in clinical simulation,"

CIN: Computers, Informatics, Nursing, vol. 34, no. 9, pp. 402–405, 2016.

[10] R. Luckin and D. S. Fraser, "Limitless or pointless an evaluation of augmented reality technology in the school and home," *International Journal of Technology Enhanced Learning*, vol. 3, no. 5, pp. 510–524, 2011.

[11] I. J. Lee, C. H. Chen, and K. P. Chang, "Augmented reality technology combined with three-dimensional holography to train the mental rotation ability of older adults," *Computers in Human Behavior*, vol. 65, no. DEC, pp. 488–500, 2016.

[12] M. Y. C. Yim, S. C. Chu, and P. L. Sauer, "Is augmented reality technology an effective tool for E-commerce? An interactivity and vividness perspective," *Journal of Interactive Marketing*, vol. 39, pp. 89–103, 2017.

[13] J. Zhang, S. K. Ong, and A. Nee, "Design and development of an in situ machining simulation system using augmented reality technology," *Procedia CIRP*, vol. 3, no. 1, pp. 185–190, 2012.

[14] Y. T. Wang, W. Zheng, and Y. Liu, "On-site digital reconstruction of yuanmingyuan using augmented reality technology," *Science and Technology Review*, vol. 24, no. 3, pp. 36–40, 2006.

[15] C. Hsu and H. C. Shiau, "The visual web user interface design in augmented reality technology," *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 2, 2013.