Research of Virtual and Real Occlusion Technology and Experimental Verification

Bo WU, De-long HU and Jin-li XU

School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan, China

*Corresponding author

Keywords: Augmented reality, Virtual and real occlusion, Background contour extraction, Occlusion recovery.

Abstract. For the wrong occlusion relationship between the virtual and real objects is generated due to the spatial relationship in augmented reality, which makes it difficult for the user to distinguish the real relationship between the virtual and real objects, the influence of immersion and interaction. The background contour of the occlusion object is extracted by the improved background modeling method and the improved shadow removal method. The occlusion relationship between the virtual and real objects is determined by depth information recovery and pixel traversal. The occlusion recovery is completed by using OpenGL and OpenCV as graphics drawing and handling tools. Experimental verification shows that the virtual and real occlusion handling method can effectively recovery the correct virtual and real occlusion relationship.

Introduction

Augmented reality technology superimposes the virtual objects into the real scenes to complete the fusion of virtual and real scenes. However, the occlusion problem is generated between the real and virtual objects due to the existence of spatial relationship, which affects the fusion effect. Therefore, it has great significance to solve the virtual and real occlusion problem.

In the last years, a large number of related research on virtual and real occlusion have been carried out by scholars. Billingherst et al. determined the occlusion relationship by modeling the real scene and comparing the depth information between the real and virtual object model [1]. Olsson et al. proposed the algorithm whose image resolution become half of the original to extract the depth information of the scene, and compared with the depth information of the virtual objects to determine the occlusion relationship [2]. Dong et al. and Behzadan et al. both used the depth perception algorithm and frame buffer technology to achieve accurate virtual and real occlusion’s judgment and recovery [3,4].

The above literatures’ methods, some have better real-time performance but the calculation amount is too large or not suitable for complex backgrounds; some methods have faster calculation speed but with poorer real-time performance. In this paper, firstly, the improved background modeling and shadow removal method is used to obtain a more accurate background contour of the object. Then, the improved occlusion judgment method is used to judge the occlusion relationship between the virtual and real objects. Finally, the occlusion relationship is recovered by using OpenGL and OpenCV.

Basic Principles of Spatial Vision

Figure 1 shows the principle of spatial vision imaging. The visual imaging involves coordinate transformation between four coordinate systems, which is world coordinate system \((O_wX_wY_wZ_w)\), the camera coordinate system \((OX,Y,Z)\), the imaging plane coordinate system \((O_{xy}xy)\) and the image pixel coordinate system \((O_{uv}uv)\) [5]. \(f\) is the focal length of the camera.
The depth information of the scene can be obtained by transformation between the world coordinate system and the image pixel coordinate system. According to the basic principle of perspective projection, the homogeneous coordinates in the world coordinate system is \( X_w = \left[ X_w, Y_w, Z_w, 1 \right]^T \), and in the image pixel coordinate system is \( x = \left[ u, v, 1 \right]^T \), so the coordinate transformation process can be showed in formula (1).

\[
Z_x \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = K \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}
\]

(1)

Where, \( f_x = f / d_x \) and \( f_y = f / d_y \) are the focal lengths when using the pixel width \( d_x \) and height as units; \( K \) is an in-camera parameter; \( T_{ew} \) is an external camera parameter; \( R \) is three-dimensional rotation matrix; \( t \) is three-dimensional translation matrix.

Research on Virtual and Real Occlusion Handling Technology

Improved Background Contour Extraction Algorithm

The PBAS (Pixel-Based Adaptive Segmenter) background modeling algorithm completes the background real-time detection by combining the selection and update of the judgment threshold which adaptive change with the background complexity. However, the calculation amount is large and the real-time detection cannot be realized when scenes change faster \([6]\). The formula of judgment threshold selection and update have been adjusted to gain an improved background contour extraction algorithm. If the current background is more complicated, the judgment threshold is raised to prevent the background pixel from being detected as the foreground pixel. Conversely, the judgment threshold is appropriately reduced to prevent the foreground pixels from being detected as background pixels. The adjustment of the judgment threshold update formula is shown in the formulas (2), (3), and (4).

\[
R \left( x_i \right) > d \left( x_i \right) \quad (2)
\]

\[
d \left( x \right) = \begin{cases} 
\frac{1}{N \sum_{i} \min_{j} \text{dist} \left( I \left( x_i \right), B_j \left( x_i \right) \right) \times R_{scale}} & \text{if } \frac{1}{N \sum_{i} \min_{j} \text{dist} \left( I \left( x_i \right), B_j \left( x_i \right) \right) \times R_{scale}} > R \left( x_i \right) / R_{scale}^2 \\
\frac{1}{N \sum_{i} \min_{j} \text{dist} \left( I \left( x_i \right), B_j \left( x_i \right) \right) \times R_{scale}} & \text{else}
\end{cases}
\]

(3)
where $R(x_i)$ is the adaptive threshold; $d(x_i)$ is the background complexity; $I(x_i)$ is the current pixel value, $B_r(x_i)$ is the background model; $dist(*)$ is the minimum Euclidean distance between $I(x_i)$ and $B_k(x_i)$; $R_{scale}$ is the current background complexity threshold adjustment control; $\Delta t=R(x_i) \cdot R_{inc/dec}$ is the fixed threshold transformation, which is determined by the initial threshold and the threshold change $R_{inc/dec}$.

An improved shadow removal method has been studied, which including chrominance, brightness and brightness change rate, to solve the problem of wrong shadow detection and void phenomenon when the color of object and shadow are too closely which will affect the background contour extraction. As shown in the formula (5)

$$
\begin{align*}
&\left| y_i^j - y_j^r \right| \leq T, \quad \forall z \in \{r, g\} \\
&\varphi \leq \frac{y_i^j}{y_j^r} \leq \delta, \quad z \in \{I\} \\
&\left| \frac{y_i^j - y_j^r}{x_i^j} \right| \leq \lambda, \quad z \in \{I\}
\end{align*}
$$

Where $y_i$ is the current frame image; $y_j$ is the image of the established background model; $z$ representing any value of $r$, $g$, $I$, and $T, \varphi, \delta, \lambda$ is the judgment threshold of the shadow and non-shadow.

**Judgment and Recovery of Virtual and Real Occlusion Relationship**

The high accuracy depth information recovery method and the fast pixel traversal method is combined to ensure the handling speed and accuracy simultaneously to improve the judgment efficiency of the virtual-real occlusion relationship. The contour depth lines is used to equally divide the object background contour area from bottom to top in the depth recovery method. The depth value is shown in the formula (6).

$$
D(l_j) = 128 + \frac{255}{m} \times j \times k
$$

Where, $l_j$ is the contour depth line $j$, $m$ is the number of contour depth lines, $255/m$ is the amount of contour depth change, $k$ is the angle tilt of camera shooting, $D(l_j)$ is the depth value of the contour depth line $j$.

The stencil buffer technology of OpenGL is used to recover the occlusion relationship, and the virtual and real occlusion recovery is completed by the stencil test and the depth buffer in the stencil buffer. The stencil test technology is used to test the pixel values in stencil buffer. If the stencil test condition is met, that is, the pixel template value is 1, the pixel is drawn. Otherwise, it is not drawn. The depth buffer first sets a larger initial pixel depth value, and compares the current pixel depth value with the initial pixel depth value. If the current pixel depth value is smaller than the initial pixel depth value, the current pixel is drawn. Otherwise, the current pixel is not drawn.

**Results and Discussion**

Figure 2 shows the virtual and real occlusion handling effect of the virtual and real occlusion handling method proposed in this paper when the human hand performs a small range of movement before the virtual differential shell which registered in the real scene based on marker. The experimental result shows that the background contour of the moving object can be accurately detected, the occlusion handling and recovery have higher accuracy.
Figure 2. Experimental effect when occlusion object moves.

Figure 3 shows the experimental results of the same frame for the proposed method and dynamic difference method. Compared Figure 3(a) with Figure 3(b), when the human hand performs a small range of movement before the virtual differential shell, the effect of the occlusion handling method proposed in this paper is better than the dynamic difference method. The dynamic difference method can also finish the virtual and real occlusion handling, but the error in the contour detection of the moving object is slightly larger than the proposed method, so that the occlusion relationship judgment and recovery are slightly worse. It shows that the proposed method has some improvement in the contour detection of moving objects, the judgment of virtual and real occlusion relationship and the recovery of virtual and real occlusion.

Summary

In this paper, background contour extraction, virtual reality occlusion relationship judgment and recovery have been studied to obtain the correct virtual and real occlusion relationship. Firstly, an improved background modeling method is obtained by adjusting the selection and update of the judgment threshold, and the brightness change rate is added to obtain an improved shadow removal method. The two improved methods have been combined to extract the background contour of moving object. Then, the depth information recovery method and the pixel traversal method are combined to achieve accurate and fast judgment of the virtual and real occlusion relationship. Finally, the experimental verification of the proposed occlusion handling method is carried out. The experimental result shows that the method can accurately complete the virtual and real occlusion handling, which has certain improvement compared with other algorithms.

Acknowledgement

This research was supported by sub-project of the new smart factory model for the lightweight of auto chassis key components, which is real axle/independent suspension assembly process planning.
system development based on VR/AR technology (Government revenue and expenditure classification 2159999, budget number 21507707999107).

References


