Occlusion Handling in Augmented Reality System for Human-Assisted Assembly Task

Chi Xu[1], Shiqi Li[1], Junfeng Wang[1], Tao Peng[1], and Ming Xie[2]

[1] School of Mechanical Science & Engineering, Huazhong University of Science & Technology, Wuhan, Hubei Province, 430074, China
[2] Nanyang Technological University, School of Mechanical & Aerospace Engineering, Block N3, Nanyang Avenue, 639798 Singapore
tomxu2008@yahoo.com.cn, sqli@mail.hust.edu.cn, wangjf@mail.hust.edu.cn, pntao@smail.hust.edu.cn, mmxie@ntu.edu.sg

Abstract. Augmented reality (AR) based human-machine interaction (HMI) provided a seamless interface between user and application environment, but occlusion handing remained as a tough problem in AR applications. Recently most AR occlusion handling algorithms aimed at general environment, more special research on AR occlusion handling for application in the context of industrial assembly, such as assembly training and assembly task guidance, is required. An occlusion handling method aimed at video see-through AR-based assembly system is presented based on the analysis of the occlusion between virtual part, visual assembly feature, navigation information, physical part and physical assembly environment in the context of virtual/physical assembly working space. The method was implemented in prototype AR-assembly system and was proved to be efficient in handling the virtual/physical occlusion in augmented assembly scene.

Keywords: Augmented reality; Human-machine Interaction; Virtual assembly; Occlusion; Depth image.

1 Introduction

In recent years, augmented reality (AR) and human-machine interaction (HMI) have been becoming more and more tightly interrelated, with the help of AR technology physical working space can be enhanced with various useful digital content [1]. In AR system for assembly task application, AR displays can enhance the user's perception of the physical assembly working space and assembly constraint feature between virtual/physical parts, providing access to information the user cannot directly perceive when unaided. AR can be used in a wide range of areas in the industrial assembly field, and potential applications can be concluded into following 5 aspects [2]: AR-based assembly feature design [3], AR-based assembly sequence evaluation [4], AR-based assembly training, AR-based assembly task guidance, and AR-based assembly inspection.

Two basic choices are available to accomplish the combining of virtual and physical scene: optical and video see-through HMD. Compare with optical see-through

C. Xiong et al. (Eds.): ICIRA 2008, Part II, LNAI 5315, pp. 121–130, 2008.

HMDs, video see-through HMD is widely used in AR applications because of its flexibility in combining strategies, easier to match the brightness of virtual and physical objects, and relatively cheaper. Video see-through HMD is used in Prototype AR system presented in this paper to achieve a virtual/physical mixed assembly environment (see **Fig. 8**).

Most of the AR application directly overlay the computer generated graphic of virtual parts on physical scene video captured by camera, therefore the image of physical parts is occluded by the graphic of virtual parts all the time (see **Fig. 1**). Many researches aimed at solving the incorrect occlusion problem: Fuhrmann [5] selected a portion of the physical scene and aligned 3d geometry models to them, and the 3d geometry information of models was used to manipulate occlusion. Yokoya [6] calculated depth information of physical objects by stereo vision theory, and the method cost large amount of calculation as it deals with the whole scene to calculate the depth information of physical environment, but stereo vision method can not be used in real-time AR system. Ohta [7] used client/server method to accelerate calculation speed, but vision field of client and server suffered from inconsistent problem. Lepetit [8] accelerated calculation speed by calculating approximate depth information of physical objects through their 2d contours in the video, however the method could not achieve the accurate depth information of complex 3d geometry object such as assembly part.

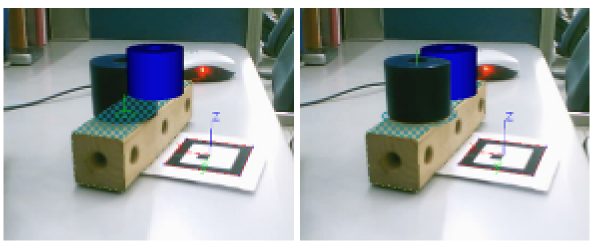

Fig. 1. Virtual/physical occlusion. (Left) Incorrect occlusion. (Right) correct occlusion.

Occlusion handling methods mentioned above are general methods, however there is seldom special research on occlusion between virtual/physical parts in the AR-based assembly applications. It is noted that accurate depth information of virtual and physical parts are required to manipulate occlusions by AR-based assembly system. However both accuracy and calculation speed of stereo vision theory based method need to be improved. It is considered that CAD model of physical assembly parts are always available in industrial assembly application, therefore occlusion handling method presented in this paper used 3d CAD model of physical parts to calculate depth information for virtual/physical occlusion handling.

2 Visual Elements in AR-based Assembly Scene

In the context of AR-based assembly application, visual elements can be classified into 2 categories: virtual and physical elements.

2.1 Virtual Elements

(1) Virtual Part. Virtual part does not exist in the physical assembly working space, and computer generates the graphic of virtual parts and overlays it on the video captured by camera mounted on the HMD. In the designing stage of an assembly project, AR system gets 3d geometry model of virtual parts from CAD software and merges it into virtual/physical environment to accomplish the pre-assembly procedure, and deficiency of the assembly design can be found and modified. In the training stage, virtual part was transformed to target place through defaulting path and then constrained in the computer driven mixed assembly scene, so that user can learn the assembly procedure intuitively.

(2) Visual Assembly Feature. Visual assembly feature means visual geometric elements representing assembly relations and constraints between assembly parts. AR-based assembly system help the user identify the constraints between assembly parts intuitively with visual assembly feature while manipulating the part to the target place. For instance, part A and part B have a face-mating relationship, AR system displays the faces to be mated on A and part B while part A approaching to part B, so that face-mating constraint of the parts can be augmented by visual assembly feature and be identified intuitively (see **Fig. 2**). Elements of assembly feature include:

- Line: including beeline, curve and circle, these assembly features are used to represent assembly constraints such as hole-shaft fitting, edge alignment and so on.
- Surface: including plane, curved and sphere surface, mainly used to represent constrains such as face-mating, face-parallel and face-alignment and so on.

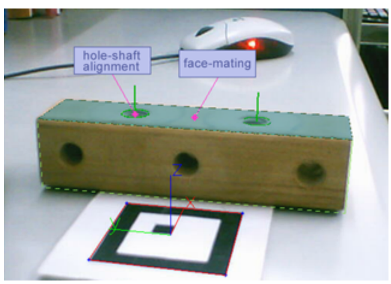

Fig. 2. Visual assembly feature with line and surface style

(3) Navigation Information. Navigation information includes product name, part name and id, guidance for assembly procedure in assembly handbook, prompt message (see **Fig. 2**), and also graphic elements such as line, rectangle, circle, arrow and so on, these elements are used to let the user pay attention to the assembly information such as active region, assembly direction, assembly path and so on (see **Fig. 7**).

2.2 Physical Elements

(1) Physical Part. Physical part means the assembly part exists in the physical assembly working space, and its feature is presented in the video capture by the camera mounted on the HMD.

It is possible that physical part's feature is not clear enough for user to identify in the physical scene video because of the insufficient light condition or out of focus, therefore physical part's outline contour need to be augmented. AR-based assembly system achieves 3d model data from CAD software, align the model to corresponding physical part's region in physical scene video captured by camera, and then enhanced the user's perception by overlaid physical part's location and geometry information.

(2) Physical Assembly Environment. Physical assembly environment means the physical working space in which assembly procedure takes place, including assembly platform, brace, clamp, and workshop etc. Physical assembly environment is dealt as background in the AR scene by default. Prototype AR-based assembly system presented in this paper dealt the physical assembly environment as bounding box and constrained assembly elements within certain range of physical space by collecting detection technology.

3 Occlusion between Virtual/Physical Elements

Virtual elements in the AR-based assembly scene include virtual part, visual assembly feature and navigation information; Physical elements include physical part and working environment. Occlusion between virtual and physical elements in AR assembly application is analyzed in this section.

3.1 Occlusion between Virtual Part and Physical Part

The most important portion in occlusion handling method in AR-based assembly application is occlusion between virtual and physical parts, because users always pay more attention to assembly relational region than other of the video image. Accurate depth information of both virtual and physical parts is needed for occlusion handling, for inaccurate depth information will lead to incorrect occlusion between virtual/physical parts when virtual part is attached to physical part through assembly constraint and tightly mated together. In this paper we calculate the accurate depth information of virtual and physical parts based on the 3d CAD model of assembly parts while it is always available in industrial assembly application. Occlusion of virtual and physical parts can be classified into 4 categories:

- Virtual occlude physical: When the location of virtual part is closer to the observer than physical part, system directly overlaid virtual graphic generated by computer up on the physical scene captured by camera mounted on the observer's head.
- Physical occlude virtual: On the other hand, if physical part's location is closer to the observer than virtual part, system remained the region of physical part in physical scene video unchanged and neglected virtual part's graphic in the same region.
- Virtual/physical mutual occlusion: In this case, virtual and physical parts occluded mutually (see Fig. 3) and we can not simply directly overlaid graphic of one part on to another. For the region of virtual part's graphic generated by computer and physical scene video captured from camera can be divided into pixels, and AR-based assembly system deals with virtual/physical occlusion region pixel by pixel to solve the mutual occlusion problem.

Fig. 3. (Left) mutual occlusion of virtual/physical parts: virtual part antenna is partial occluded by physical part shell, and physical part is also partial occluded by virtual part. (Right) Incorrect occlusion.

- Translucent occlusion: In this case, visual property of virtual part is translucent. In other words, physical part's feature behind the virtual part can be seen simultaneously by alpha channel blending method. In AR-based assembly application, if translucent occlusion mode is on, user could always perceive physical assembly part through video see-through HMD even when it is occluded by translucent virtual ones.

3.2 Occlusion between Visual Assembly Feature and Physical Part

Assembly constraints among assembly parts, such as face-mating, alignment, hole-shaft mating, can be augmented in the users' vision field intuitively by visual assembly feature in AR-based assembly system. Visual assembly feature is a kind of virtual 3d geometry object whose elements include line and surface as mentioned above. Similar to the virtual/physical parts' occlusion, the occlusion between visual assembly feature and physical part includes: virtual occlude physical, physical occlude virtual, virtual/physical mutual occlusion and translucent occlusion.

It is recommended that the property of surface element in visual assembly feature can be set to translucent style by defaulting, so the observer can always perceive the geometry and location of virtual and physical parts even when they are occluded by translucent assembly feature (see **Fig. 2**).

3.3 Occlusion between Navigation Information and Physical Part

2d navigation information is always visible and cannot be occluded by physical part or other visual elements, it occluded or translucently occluded other elements in mixed assembly environment. 3d navigation information such as arrow, rectangle region, and circle are virtual 3d geometry elements, Occlusion between 3d navigation information and physical part is similar to virtual and physical parts' occlusion described above, it includes: virtual occlude physical , physical occlude virtual , virtual/physical mutual occlusion and translucent occlusion.

3.4 Occlusion between Virtual Elements and Physical Assembly Environment

In AR-based assembly application, physical assembly environment includes assembly platform, brace, workshop, etc. While user is performing assembly task in virtual/physical mixed assembly working space, physical environment is shown as background in the scene, and computer generates virtual elements are directly overlaid on to the physical assembly environment in the video captured by camera, since physical environment was always further to the observer than virtual elements.

Bounding box is used to constrain virtual elements be in a certain region of working space and make sure that virtual elements to be always closer to observer than physical environment.

4 Occlusion Handling Method for AR-based Assembly Application

4.1 Placing Virtual Part in Contact with Physical Part in AR Assembly Scene

3d geometry models of virtual/physical parts are achieved from CAD software interface, they are placed in the AR assembly scene in contact with each other automatically by registration procedure. First parts' transform matrices from physical world coordinate to camera are estimated, and then parts' CAD models are transformed and projected onto image plane aligning with physical assembly scene. Great deal of researches can be found on this subject. Our registration method is similar to Kato's method, and the details can be found in [9] and [2].

It is noted that lens distortion of camera mounted on users' head would reduce precision of virtual/physical parts' alignment. The distortion can be approximated using the following expression:

$$P_d = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) P_n + dp. \tag{1}$$

where k1, k2, k3 are coefficients for radial distortion, dp is tangential distortion vector, and P_d is distorted position in image coordinate. Distortion of physical scene video is verified in our system by calibration procedure, details can be found in [2].

4.2 Virtual/Physical Depth Image

Virtual depth image records the depth information of virtual elements such as virtual part, visual assembly feature and navigation information within rending area, it is a gray scale image with the same size as computer generates graphic. The value of virtual depth image indicates the distance from observer to virtual element's surface being rendered at current pixel. Virtual depth image can be acquired from 3d graphic engine such as OpenGL and OpenInventor while rendering procedure.

Physical depth image records depth information of the video captured by camera mounted on observer's header pixel by pixel. AR system aligns physical part's CAD model to its corresponding region in physical scene video, by rending the physical model, its depth image can be easily achieved from rendering engine. The range of gray scale value in depth image is from 0 to 1, it is recommended that depth information of physical assembly environment be set to the maximal value 1, so that it will be directly overlaid by other elements in the mixed scene by default (see Fig. 4).

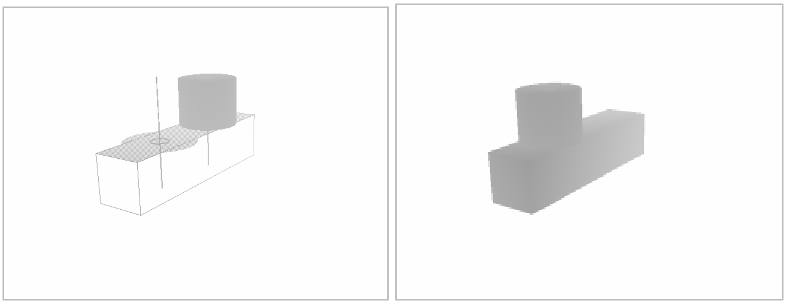

Fig. 4. (Left) depth image of virtual elements. (Right) depth image of physical elements.

4.3 Occlusion Handling Procedure

When virtual scene, virtual depth image, physical scene and physical image are available, system integrates the virtual/physical scene. Virtual scene is generated by computer, physical scene is captured by camera mounted on observer's head, and virtual/physical depth image are acquired from rendering engine mentioned above. While integrating, system compares the depth value of virtual scene with physical pixel by pixel: if physical depth value is less than virtual one, it means that virtual element's region at current pixel is further to observer and should be occluded by physical element, system fills current pixel with physical scene, and vise versa. Occlusion relation of 3d virtual elements and physical elements will be correctly dealt with through this procedure (see **Fig.5**).

After 3d geometry elements are integrated into one image, 2d navigation information will be overlaid on top of mixed video. It is noted that there are also occlusions between 2d navigation information, and the order of 2d information overlaying on the video is determined by their layer id. 2d information is drawn on the video from lest important to the most important for the sake that important information would not be occluded by less important ones. The whole procedure of occlusion handling method is illustrated below (see **Fig. 6**).

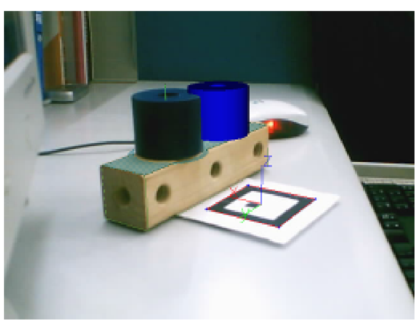

Fig. 5. Occlusion between virtual and physical 3d geometry elements

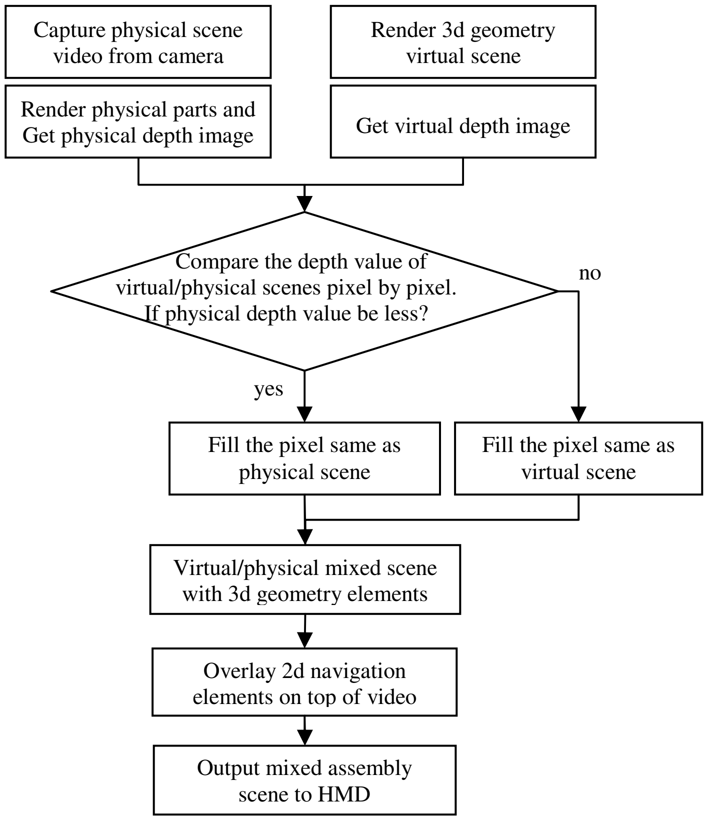

Fig. 6. Occlusion handling method

5 Prototype AR-based Assembly System

Prototype AR-based assembly system is a sub-system of smart assembly system which belongs to national defense "eleven five" foundational scientific research project committed by Huazhong university of science and technology. Prototype system presented in this paper was implicated in assembly task training for an aerospace product, and it has been proved that occlusion handling method can real-time manipulate the occlusion in AR-based assembly application (see **Fig. 7**).

Geometry model was exported to VRML file ".wrl" from CAD software pro/e; assembly scene management and graphic rendering by virtual reality SDK OpenInventor; assembly feature rendering and depth image achievement by OpenGL; capturing video from camera by DirectShow; image processing and occlusion handling by open-source library OpenCV. Development platform was Microsoft Visual C++ 6.0 and operation system was Microsoft windows XP.

Physical scene video was captured by camera mounted on the video see-through HMD "AR-vision-3D" (see **Fig. 8**), CMOS color sensor, resolution 640×480, maximal capture rate 48fps. Virtual scene is generated by pc with Intel Pentium 4 CPU 2.4G, GeForce MX 440 AGP8x 64M, memory 512M. Virtual/physical assembly scene was displayed on the HMD with solution 800×600 and rate 12fps.

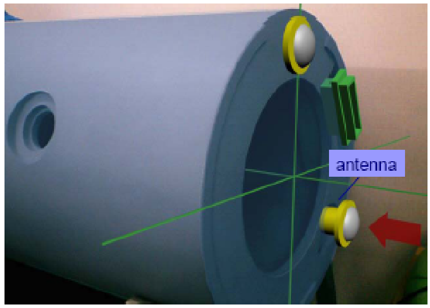

Fig. 7. Prototype AR-based assembly system

Fig. 8. Video see-through HMD "AR-vision-3D"

6 Conclusion

Augmented reality technology can be implicated in industrial assembly field to form a virtual/physical parts mixed assembly environment, and it can effectively improve the product quality, shorten period of development and reduce the cost of the product.

Virtual/physical occlusion handling is one of the main problems in AR technology, by now there is no special research that analyzed occlusion handling of virtual/physical elements in the industrial assembly field. In this paper occlusion among virtual part, visual assembly feature, navigation information, physical part and physical assembly environment is analyzed, and then occlusion handling method is presented for video see-through AR-based assembly application. The experiment of prototype AR system proved that occlusion handling method can solve the occlusion problem in industrial assembly application efficiently.

References

1. Azuma, R.T.: A survey of augmented reality. J. Presence: Teleoperators and Virtual Environments 6(4), 355–385 (1997)
2. Xu, C.: Research on the virtual/physical scene integration of augmented assembly system. M.Sc. Huazhong University of Science and Technology, WhuHan (2007)
3. Yan, P., Nee, A.Y.C., Soh Khim, O., Yuan, M., Youcef-Toumi, K.: Assembly feature design in an augmented reality environment. J. Assembly Automation 26(1), 34–43 (2006)
4. Klinker, G., Dutoit, A.H., Bauer, M., Bayer, J., Novak, V., Matzke, D.: Fata Morgana: a presentation system for product design. In: Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality, pp. 76–85. IEEE Comput. Soc., Darmstadt (2002)
5. Fuhrmann, A., Hesina, G., Faure, F., Gervautz, M.: Occlusion in collaborative augmented environments. J. Computers & Graphics 23(6), 809–819 (1999)
6. Okuma, T., Kiyokawa, K., Takemura, H., Yokoya, N.: An augmented reality system using a real-time vision based registration. In: Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170), vol. 2, pp. 1226–1229. IEEE Comput. Soc., Brisbane (1998)
7. Ohta, Y., Sugaya, Y., Igarashi, H., Ohtsuki, T., Taguchi, K.: client/server depth sensing for see-through head-mounted displays. J. Presence: Teleoperators and Virtual Environments 11(2), 176–188 (2002)
8. Vacchetti, L., Lepetit, V., Fua, P.: Combining edge and texture information for real-time accurate 3D camera tracking. In: Third IEEE and ACM International Symposium on Mixed and Augmented Reality, pp. 48–56. IEEE Comput. Soc., Arlington (2004)
9. Kato, H., Billinghurst, M.: Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In: Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR 1999), pp. 85–94. IEEE Comput. Soc., San Francisco (1999)