Title	実空間での位置合わせを可能とするAR コンテンツオーサリング手法の提案
Author(s)	碓﨑，敬祐
Citation	
Issue Date	2019-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/15809
Rights	
Description	Supervisor:宮田　一乗，先端科学技術研究科，修士（知識科学）

A Proposal of AR Content Authoring Method Enabling Alignment in Real Space

In 2016 it is called the first year of VR and the technology of 3DCG (three dimensional computer graphics) becomes common, and it is expected that it will become more popular from now. With the development of the game engine 3DCG contents became possible to produce high quality works individually. And with the development of HMD (Head Mounted Display), AR (Augmented Reality) is compatible with smartphones equipped with various sensors and cameras, making VR (Virtual Reality) more easily hands-on experience. With regard to AR, AR frameworks such as ARKit for Apple iOS, ARCore for Google android, vforia of Unity are remarkable development. Estimation of the position of the device by motion tracking, desk and ground by detecting the horizontal and vertical surfaces, object placement using the wall surface, expression of light and darkness of the object by estimation of the illumination state of the environment, etc., enabling a more realistic experience Has become.

On the other hand, even if you compare graphs of investment amounts, it is hard to say that AR development differs from VR development. The reason is that development and utilization of AR content are more difficult than VR content. The biggest advantage of AR not in VR is that real space can be used. However, when actually developing, 3DCG space is created by basically arranging objects in a space other than a marker. When IDE IDE (Integrated Development Environment) is used, GUI-based support is not done in the first place since it develops with code. Since it is unknown what kind of environment the contents are used in this situation, it is difficult to produce complicated contents acting on the real space. Also, in real space there are obstacles such as restrictions on size and furniture, and it is difficult to experience large-scale content like VR. For the above reasons, simple content such as placing and appreciating objects is becoming mainstream.

By recognizing feature points of the space, 3D objects can be displayed without using patterned markers. By using image feature points, images and objects as well as patterned figures can be used as markers. Because you can use what is in the environment, you can experience AR more naturally. However, as with the marker expression, it is necessary to fit the feature points within the angle of view, so it has similar disadvantages.

Instead of the conventional markerless type AR, in recent years, instead of the conventional markerless type AR, in recent years the AR case using the position estimation of the camera using SLAM, motion capture, etc. is called a markerless type AR Has become. By recognizing image feature points, camera pose estimation can be performed by self position estimation of spatial recognition, so compatibility with AR is good. In this method, it is unnecessary to prepare a marker by a specific image or solid, and it is not necessary to hold a marker in the angle of view. Therefore, it is possible to display even a large 3D object that does not fit in the angle of view in the case of a marker expression . However, it is necessary for the experiencer to decide the origin of displaying the 3D object, and the 3D object can not be placed at the place intended by the creator.

In the study of Naoki Ikeda, we visualized numerical simulation results of dynamic phenomena

using AR visualization system based on SLAM technology. A marker is used for initial alignment of alignment between the simulation result model prepared in advance and the real space. Since this method can acquire an accurate initial position, a more accurate AR can be implemented. On the other hand, if the installation position of the marker shifts, the initial position of the SLAM shifts, so that the object has a disadvantage that the object is shifted and displayed.

In the AR system using SLAM, we mainly use camera pose estimation and we have not been able to grasp the environment of the space using the created environmental map.

In this research, we propose authoring method to facilitate the creation of content using real space in specified space specified by content creator.

In this method, we first create an environmental map using ORB - SLAM 2. The ORB-SLAM 2 is a real-time SLAM library using a monocular camera, a stereo camera, RGB-D (depth information can be acquired in addition to a color image) camera. An environmental map is used for grasping the space during 3DCG object placement and for estimating the position and orientation of the camera during AR. Next, the 3D CG object is placed at the intended position by using the environment map. The user can perform parallel movement to move the object in the x, y, z axis direction, rotation movement to rotate the object in the x, y, z axis direction and expansion / contraction performed around the origin of the model coordinate system of the object I will make adjustments by repeating.

In indoor cases, it is possible to place 3DCG objects in the actual intended place. In the outdoor case, it was confirmed that it is possible to arrange the 3DCG object by using this system without preparing the environmental map by visiting the site by preparing the image input such as the data set etc in the outdoor case.

In order to actually experience the content, it is desirable to have a portable device such as a smartphone that can be carried easily, but since this system can not be executed in real time unless it is a powerful CPU, it is installed on a desktop PC.

When creating an environmental map, this system can be used with only a monocular camera, but in order to obtain a more accurate environmental map it is not a monocular camera but an indoor case RGB-D camera, in outdoor cases a stereo camera Data set was used. These are not general equipment, but the use conditions of this system are strict.

Since the environmental map obtained by ORB - SLAM 2 is a point cloud, it is impossible to grasp environmental materials etc on its own without checking the input image and AR viewer. Accurate placement is difficult with the environmental map alone. As a future prospect, since occlusion expression is not currently available, there is a problem that even if it can be placed in the intended place, it can be seen from a place that it can not be seen. Since we can grasp the obstacles in the environment such as the wall by the environmental map, we think that occlusion expression is possible using it.