Development of Dismantlement Support AR System for Nuclear Power Plants using Natural Features

Tomofumi Fujiwara*[a], Hirotsugu Minowa[a], Yoshiomi Munesawa[b]

[a] Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushimanaka, Kita-ku, Okayama-shi, Okayama 700-8530, Japan
[b] Department of Mechanical Systems Engineering, Faculty of Engineering, Hiroshima Institute of Technology, 2-1-1 Miyake, Saeki-ku, Hiroshima 731-5193, Japan
fujiwara.t@mif.sys.okayama-u.ac.jp

We developed a system that overlays a 3D model onto a target in a camera image, in other words, a system that can display a dismantled part directly onto a target. The system constructs a database called a "key frame database" in advance, which consists of multiple sets of a target's photo, a target's position and posture when a camera captured it, 2D and 3D coordinates and normal vectors of feature points in the photo, image patches around the feature points. At runtime, this system takes an image sequence of a target by using a camera and estimates the target's position and posture by matching the image with photos in the key frame database. A 3D model of the target is located on the current image based on the estimated position.
We conducted experiments for the developed system, which evaluated an accuracy of the target's position and posture estimation, limitations of distances about overlaying the 3D model, robustness for occlusions and an update rate of the estimation. As a result, we confirmed that the system had a capacity for dismantlement work supports.

1. Introduction

Nuclear power plants are generally necessary to be dismantled due to aging from 30 to 50 y after the construction. In Japan, 19 nuclear reactors have already been operated for more than 30 years at 2009 and they are needed to be dismantled or repaired (Atomic Energy Society of Japan, 2009). Demands for dismantlement works will increase as time passes for the future. Moreover, the issue of dismantlement is coming to light since the Fukushima Daiichi Nuclear Power Plant was stopped after the Tohoku Earthquake and Tsunami occurred in 2011.

On the other hand, dismantlement works are dangerous for human workers because of radiation exposures. Dismantlement tasks are generally conducted according to a following procedure: (1) to plan the dismantlement, (2) to survey the radiation dose, (3) to sectionalize targets, (4) to prepare dismantlement works, (5) to operate and record dismantlement tasks and (6) to collect radioactive wastes and clean up (Ishii et al., 2008). In the step (5), Human workers dismantle targets checking parts to be dismantled with a 3D actual environment by comparing 2D work procedures made in the preparation steps (1) to (4). Therefore, their works take much time and workers' radiation exposure increased. We are able to decrease working time and to make the workers safer if they do not have to compare work procedures with environments.

In terms of dismantle works of nuclear power plants, Iguchi et al. (2004) has developed a system that plans dismantlement works, simulates the works by virtual reality techniques and visualizes radiation doses. Ishii et al. (2008) has developed a system that displays dismantlement parts by Augmented Reality (AR) techniques with 2D barcode markers. AR is a technique that overlays 3D models onto camera images of actual environments in real time. The former system which Iguchi et al. (2004) has developed does not provide visual supports for human workers during working processes. Therefore it is not able to decrease a time to compare work procedures with actual environments. The latter system which Ishii et al. (2008) has developed has

disadvantage that it is not able to keep recognizing a dismantlement target if a part of a 2D barcode marker attached onto the target is hidden from a range of a camera. Therefore it is not able to keep displaying dismantlement parts to human workers during works.

In this study, we proposed a method that human workers were able to work without referring work procedures all working time by keeping displaying a dismantled part to human workers. We applied an AR technique with natural features which existed in the environment. Natural features mean edges or corners which are able to be clues as to how to recognize a target. The natural features have an advantage over 2D barcode markers since we can recognize a target in case that a part of a target is hidden from a camera if some of natural features on the target can be detected. We adopted a method that overlay a 3D model onto a target using the 3D model and natural features (Lepetit et al., 2003). This method makes a correspondence of multiple photos of a target to its 3D model in advance, then overlays the 3D model onto the target in the camera images at runtime. We assumed as a whole system that a camera is mounted on a human worker's helmet and images which a 3D model is overlaid are displayed onto a screen of a head mounted display.

We conducted experiments that evaluated the effectiveness of our developed system using 2 types of target objects based on the conditions described in the next section.

2. Required conditions for displaying dismantled parts by using AR

We defined following conditions for displaying dismantled parts by using AR: (1) an accuracy of the estimation, (2) limitations of distances about overlaying the 3D model, (3) robustness for occlusions and (4) an update rate of a target's position and posture estimation.

The condition (1) was taken into account because an error of a displayed position of a 3D model led to an error of a dismantled part to be shown to workers. Ishii et al. (2007) reported dismantlement tasks required an accuracy level from several millimeters to several centimeters. We defined 2 [cm] as the accuracy condition with a tolerance for an error which was a width of a human's finger. We defined the condition (2) because we need to overlay a 3D model onto camera images within a human's working range. The maximum working range is defined as 1 [m] from a dismantled target based on a length of an extended human's arm. The minimum working range is defined as 20 [cm] from a dismantled target, which is not too close to a human worker for safety's sake. In terms of the condition (3), we defined it because occlusions can be happened by human's arm during works. We defined a condition that a 3D model is able to be overlaid in case that a half of a dismantled target is seen from a camera by assuming it can be happened by human's arm. We considered the condition (4) because a delay of an imaging update affected works. This condition was determined as 8 [fps], which is required for general animations not to annoy people watching them.

In the experiments, we evaluated our developed system under the 4 conditions above, which are an accuracy of a target's position and posture estimation, limitations of distances about overlaying the 3D model and robustness for occlusions and an update rate of the estimation.

3. 3D model overlaying method by using AR

Procedures of our method are divided into 2 phases, a preparation phase and an execution phase. The preparation phase is described as below.

(1) Create a 3D model of a target

Create a 3D model of a target by modelling software. If 3D CAD data of a target are available, they can be used.

(2) Take pictures of a target

Take multiple pictures of a target before works from some viewpoints which are assumed in dismantlement works. Those pictures are taken when a preparation and a survey of dismantlement works are conducted. Those pictures are also taken by a camera used in works because our system uses camera's intrinsic parameters such as a focal length.

(3) Derive positions and postures of a target in the taken photos

Make correspondences of 2D coordinates on the photos to 3D coordinates on the 3D model by hand, and calculate each position and posture of the target.

(4) Obtain 3D coordinates, normal vectors and image patches of feature points on the target

Extract feature points from the photos by using Harris operator (Harris and Stephens, 1988) and obtain their 3D coordinates and normal vectors, and image patches around the feature points. The image patches are used for estimating 3D coordinates of the target. The number of feature points detected affects an accuracy of the position and posture estimation and update rates of the estimation at runtime.We equalize histograms of images before extracting feature points in order to increase the number of extracted feature points and an accuracy of the estimation.
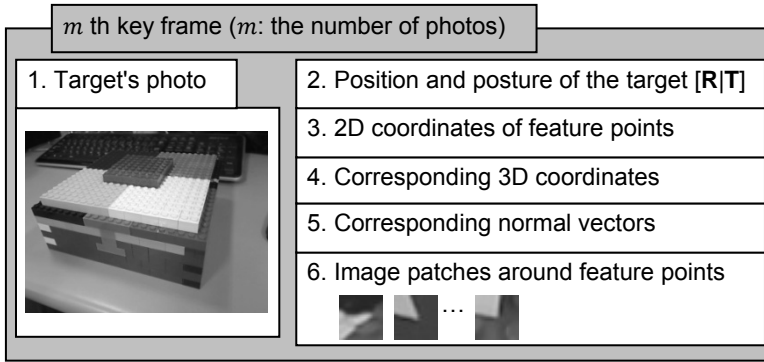
Figure 1: Key frame database

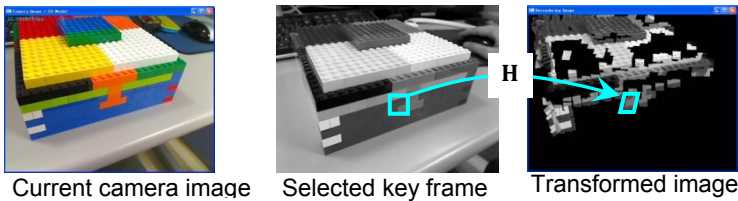Current camera image   Selected key frame   Transformed image

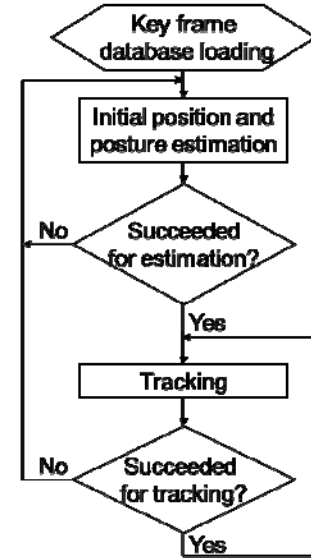Figure 3: Generation of a transformed key frame image

Figure 2: Process flow in the execution phase

The procedure (3) and (4) are repeated for every taken photo. Finally, we get a "key frame database" in this preparation phase as shown in Figure 1.

The execution phase is described as below (see Figure 2).

(1) Load a key frame database

(2) Estimate an initial position and posture of a target

Extract feature points from camera images using Harris operator and obtain image patches around the feature points. After that, compare those image patches with image patches stored in the key frame database and make correspondences of 3D coordinates in a key frame which has the most corresponding image patches. Estimate a position and posture of the target by using the correspondences of 2D coordinates with 3D coordinates by M-estimator (Arya et al., 2007), which is able to estimate parameters with noisy data including outliers, in order to improve an accuracy of the estimation.

(3) Track the target

If the procedure (2) succeeds, track the feature points between two neighbouring frames by optical flows (Lucas and Kanade, 1981) in order to keep correspondences of 2D coordinates with 3D coordinates and to increase update rates of computations. Moreover, transform a photo of a selected key frame according to a current position and posture of the target, where image patches around feature points are transformed by using a homography matrix so that a position and posture of the key frame matches a current target's position and posture as shown in Figure 3. When a feature point $\mathbf{m}_0$ which is a centre of an image patch is on a plane described by $\vec{n}^T[x\ y\ z]^T + d = 0$ in 3D space, a homography matrix $\mathbf{H}$ can be estimated by the plane as below. By defining $\mathbf{A}$ as a camera's intrinsic parameter, $\mathbf{R}_K$ and $\mathbf{R}_P$ as rotation matrices of the key frame and of the previous frame, $\mathbf{T}_K$ and $\mathbf{T}_P$ as translation vectors of the key frame and of the previous frame, a position and posture of a key frame as $\mathbf{P}_K = \mathbf{A}[\mathbf{R}_K|\mathbf{T}_K]$ and a position and posture estimated in a previous frame as $\mathbf{P}_P = \mathbf{A}[\mathbf{R}_P|\mathbf{T}_P]$, we can get a homography matrix

$$\mathbf{H} = \mathbf{A}(\delta\mathbf{R} - \delta\mathbf{T}\vec{n}'^T / d')\mathbf{A}^{-1}, \text{ where } \delta\mathbf{R} = \mathbf{R}_P\mathbf{R}_K^T, \delta\mathbf{T} = -\mathbf{R}_P\mathbf{R}_K^T\mathbf{T}_K + \mathbf{T}_P, \vec{n}' = \mathbf{R}_K\vec{n}, d' = d - \mathbf{T}_K^T(\mathbf{R}_K\vec{n}). \tag{1}$$

Now we can get transformed 2D coordinates of the surrounding points $\mathbf{m}_i$ ($i = 1, \cdots, N$) where $N = 15\times15$ as below:

$$\mathbf{m}_i' \cong \mathbf{H}\mathbf{m}_0 + \mathbf{J}(\mathbf{m}_0)(\mathbf{m}_i - \mathbf{m}_0) \quad (i = 0, \cdots, N), \text{ where } \mathbf{J}(\mathbf{m}_0) = \begin{bmatrix} \partial f_x(\mathbf{m}_0)/\partial x & \partial f_x(\mathbf{m}_0)/\partial y \\ \partial f_y(\mathbf{m}_0)/\partial x & \partial f_y(\mathbf{m}_0)/\partial y \end{bmatrix}. \tag{3}$$

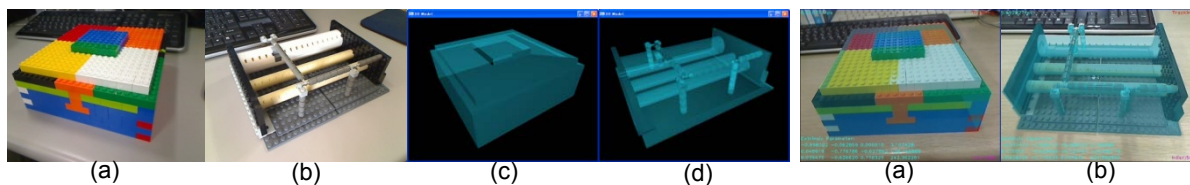(a)   (b)   (c)   (d)   (a)   (b)

Figure 4: (a) Target 1, (b) Target 2, (c) 3D model of the target 1, (d) 3D model of the target 2

Figure 5: Appearance of overlaying 3D models of (a) the target 1 and (b) the target 2

This transformation generates an image shown in the right of Figure 3, which can interpolate frames between key frames. Make correspondences of image patches around feature points extracted from this photo with them from a current camera image and increase correspondences of 2D coordinates with 3D coordinates. A key frame which has a posture with the highest similarity to a current target's posture is selected. A position and posture is estimated by using these correspondences. New correspondences are added when they are detected in order to improve an accuracy of the position and posture estimation and to be able to overlay 3D models as far as possible, which makes the system robust for occlusions.

(4) Detect a failure of the tracking

Z elements in translation vectors are taken into account to detect a failure of the tracking. Z elements mean distances from a camera to a target. If this element is minus, it means the target is behind the camera. Therefore, we can detect a failure of the position and posture estimation by monitoring this element. A failure can also be detected by monitoring if estimated position and posture is out of range of drawing of the 3D model or the number of corresponding points is under a threshold. In case that a failure is detected based on the criteria above, the procedure will be started from (2) again.

4. Evaluation experiments

We conducted below experiments on 2 types of target objects as shown in Figure 4(a) and (b). The target 1 is a colourful, simple and boxy object, while the target 2 is a colourless and complex object which imitates pipes in a power plant. The target 1 is 176 [mm] x 176 [mm] x 77 [mm] large and the target 2 160 [mm] x 197 [mm] x 55 [mm] large. Their 3D models are shown in Figure 4(c) and (d), where both are made as transparent to confirm if they are overlaid onto the camera images properly. 16 photos (640 x 480 pixels) were taken from 16 directions which can be assumed in the actual works around the target in the preparation phase. The specification of the used computer was AMD Athlon DualCore Processor 5050e 2.6GHz CPU and 4GB RAM. We used Logicool Qcam Pro 9000 shown in Figure 10 as a camera. Our developed system was executed in a single thread.

4.1 Accuracy of a target's position and posture estimation

We evaluated an accuracy of estimation by an estimated position of a target for every axis shown in Figure 10. The appearances of both overlaid 3D models are shown in Figure 5. We used 50 frames and calculated actual, true positions (ground truths) every 5 frames. The camera was moved around a target. The results of this experiment are shown in Figure 6 and 7. The box plot of these results is shown in Figure 8. The average error for the target 1 was 1.9 [mm] in X axis, 2.6 [mm] in Y axis and 4.8 [mm] in Z axis. The average error for the target 2 was 0.5 [mm] in X axis, 0.4 [mm] in Y axis and 3.9 [mm] in Z axis. The largest outlier was 13 [mm] in Z axis for the target 1.

4.2 Limitations of distances about overlaying the 3D model

We measured maximum and minimum distances which 3D models were able to be overlaid during the tracking. We did not evaluate distance limitations for the initial position and posture estimation because it depended on the positions where key frame photos have been taken. We measured a distance from a camera to a centre of a target. The result of this experiment is shown in Table 1.

Table 1: Distance which 3D models are able to be overlaid

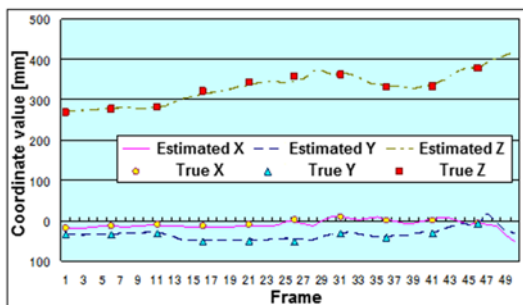|	Target 1	Target 2
Minimum distance [mm]	120	150
Maximum distance [mm]	1,240	1,100

Figure 6: Transition of the centre of the target 1

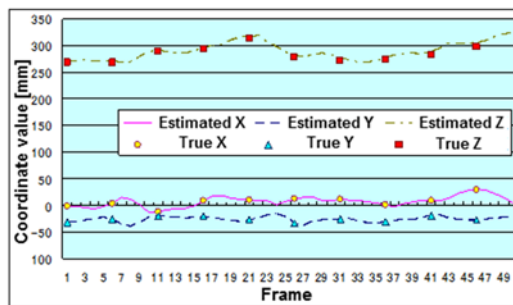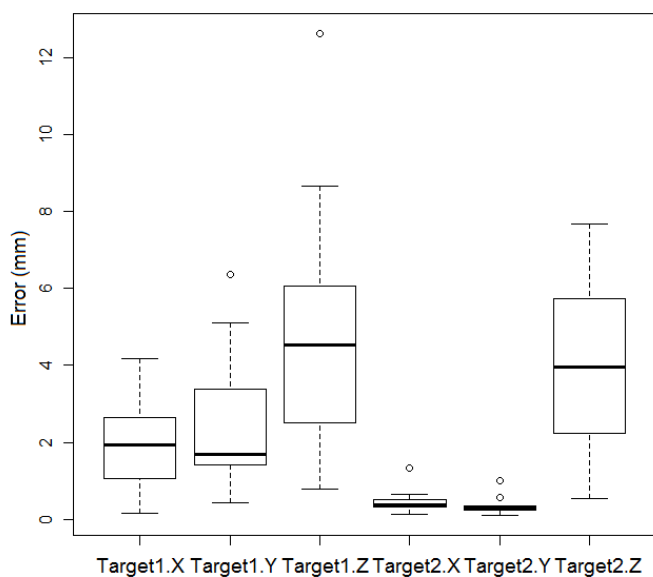Figure 7: Transition of the centre of the target 2

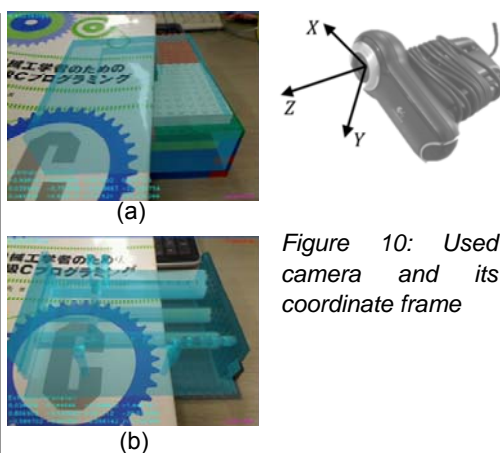Figure 8: Box plot of the errors of the targets' positions

(a)

(b)

Figure 9: Appearance of the experiment on occlusions

Figure 10: Used camera and its coordinate frame

4.3 Robustness for occlusions

In this experiment, we gradually hid the targets until the 3D models were not able to be overlaid and measured an area of an occlusion for the targets. We assumed the view point face to the front side of the target as 0 degree (the view point of Figure 5), and rotated the camera about an axis vertical to the ground by 90, 180 and 270 degrees 5 times, respectively. The averages of the results of this experiment are shown in Table 2, and the appearances are shown in Figure 9.

Table 2: Occlusion which 3D models are able to be overlaid

	Target 1				Target 2			
View point [deg]	0	90	180	270	0	90	180	270
Occlusion of the target [%]	61.8	65.2	65.7	61.2	66.0	75.3	71.1	64.4

4.4 Update rate of a target's position and posture estimation

We measured update rates of the initial position and posture estimation and tracking for 30 frames, respectively. The results of this experiment are shown in Table 3. The results show that the average of the update rate of the initial position and posture estimation is 4 to 5 [fps] and that of tracking is 9 to 10 [fps] on average for 30 frames, respectively.

Table 3: Update rate of the initial position and posture estimation and the tracking

	Target 1	Target 2
Initial position and posture estimation	5.0 [fps]	4.1 [fps]
Tracking	9.1 [fps]	10.2 [fps]

5. Discussions

We analysed experimental results based on the required conditions for displaying dismantled parts by using AR mentioned in Section 2. For the condition (1), the results showed the error of the estimated position was 13 [mm] at a maximum, which satisfied the condition 2 [cm]. In terms of the condition (2), limitations of distances about overlaying the 3D model, the result around 15 [cm] to 1[m] satisfied the condition 20 [cm] to 1 [m]. For the condition (3), robustness for occlusions, we confirmed that the 3D models were able to be overlaid as far as 30 to 40 [%] of the target can be seen from the camera. It satisfied the condition more than a half of the target. Our developed system satisfied the condition (4), an update rate of 8 [fps] in the tracking, but not in the initial position and posture estimation. We can solve this problem by reducing key frames by arranging their positions. However, it will not affect dismantlement works because the initial position and posture estimation takes only a few seconds. Therefore, our developed system satisfied all of the required conditions mentioned in Section 2. We also got almost the same results for different 2 types of target objects, which indicated that our developed system had a potential to dismantle targets continuously by switching database if shapes of the targets are getting changed as works proceed.

6. Conclusions

We developed a system that displays a dismantled part by AR so that human workers can work without referring work procedures in dismantlement works of nuclear power plants. The developed system recognized a dismantled target by using natural features and overlaid a 3D model of the dismantled target onto the camera image. As a result of experiments with a simple target object and a complex target object, we confirmed that the developed system satisfied the required conditions for displaying dismantled parts by using AR. Therefore, we confirmed that the system had a capacity for dismantlement work supports.

We will improve the update rate, display 3D models with more detailed dismantled parts and radiation doses and construct a whole system which integrates a head mounted display and a camera to conduct experiments in an actual nuclear power plant for the future.

References

About the examination's progress of decommissioning processes in Atomic Energy Society of Japan (in Japanese), 2009, Atomic Energy Society of Japan
<http://www.meti.go.jp/committee/materials2/downloadfiles/g90421b04j.pdf> accessed 04.03.2015.

Arya K.V., Gupta P., Kalra P.K., Mitra P., 2007, Image registration using robust M-estimators, Pattern Recognition Letters, 28(15), 1957-1968.

Harris C., Stephens M., 1988, A combined corner and edge detector, Proc. the 4th Alvey Vision Conf., Manchester, UK, September 1988, 147-151.

Iguchi Y., Kanehira Y., Tachibana M., Johnsen T., 2004, Development of Decommission Engineering Support System (DEXUS) of the Fugen Nuclear Station, Journal of Nuclear science and Technology, 41(3), 367-375.

Ishii H., Bian Z., Sekiyama T., Shimoda H., Yoshikawa H., Izumi M., Kanehira Y., Morishita Y., 2007, Development and Evaluation of Tracking Method for Augmented Reality System for Nuclear Power Plant Maintenance Support (in Japanese), Journal of Japan Society of Maintenology, 5(4), 59-68.

Ishii H., Nakai T., Bian Z., Shimoda H., Izumi M., Morishita Y., 2008, Proposal and Evaluation of Decommissioning Support Method of Nuclear Power Plants using Augmented Reality (in Japanese), Journal of the Virtual Reality Society of Japan, 13(2), 289-300.

Lepetit V., Vacchetti L., Thalmann D., Fua P., 2003, Fully Automated and Stable Registration for Augmented Reality Applications, Proc. 2nd IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR), October 2003, Tokyo, Japan, 93-102.

Lucas B. D., Kanade T., 1981, An Iterative Image Registration Technique with an Application to Stereo Vision, Proc. the Seventh International Joint Conf. on Artificial Intelligence (IJCAI), Vancouver, BC, Canada, August 1981, 674-679.