I Introduction

The research and development of agricultural machinery in Japan have contributed to the reduction in the number of working hours and introduction of mechanized systems in agricultural production, especially in fields and rice paddies (Tamaki, 2006). The main purpose of the research has been to make agricultural work briefer and lighter.

In the current situation of Japan’s agriculture, fewer workers are engaged in agriculture and they are aging rapidly (Ministry of Agriculture, Forestry and Fisheries, 2010). Moreover, it has become inevitable for the country to open up to trade every year. Japan’s government is considering partaking in talks on the Trans-Pacific Partnership, which is a huge free-trade area around the Pacific Rim. Therefore, it is thought that Japan’s agriculture cannot risk facing a labor shortage, as it will be exposed to harsher global competition.

To maintain the sustainability of Japan’s agricultural industry under such difficult circumstances, following strategies are required: 1. accumulation of farmland, 2. development of automated agricultural systems, and 3. training of new farmers. To realize a more competitive agricultural industry, it is necessary to accumulate agricultural land and to cultivate the field with automated machinery. This type of agriculture will allow us to manage larger fields with a smaller workforce, therefore making agriculture more profitable and less exhausting. Therefore, the development of robotic agricultural machinery in Japan has become much more important (Noguchi and Terao, 1997; Ishida et al., 1998; Nagasaka et al., 2004).

In addition, it is important to automate efficient approaches to agricultural work, which sometimes rely on a farmer’s skill and experience developed over time, in order to simplify the entry of potential and young farmers into agriculture.

Therefore, we are attempting to develop a robotic combine harvester, which is an automated head-feeding combine harvester. A prototype of robotic harvester is also being developed in the National Agricultural Research Center (Ibaraki, Japan) (Sato et al., 1996); however, the automation of the unloading process has not yet been performed.

When the grain tank in a combine harvester is full, grains are unloaded into the container attached to a pickup truck and transported to local rice-processing facilities. In general, two workers are engaged in this process: one as the harvester operator and the other as the truck driver. The operator controls the harvester’s auger and unloads grains, and then continues harvesting, while the driver transports them to rice-processing facilities that are usually located near the paddy. If this process is performed by one worker, he must interrupt the harvesting process to transport the crop to the facilities. A robotic combine harvester can enable this task to be performed by only one worker, taking lesser time than that required to perform the task using two workers.
In addition, the operator’s skill affects the speed with which he can position the auger at an appropriate unloading point. Therefore, the automation of the unloading process is an important function for not only the development of robotic combine harvesters but also the reduction of agricultural workload, as well as proper positioning, which is independent of the operator’s skill and experience.

The problem lies in the auger finding the container (or the truck) and positioning itself at an appropriate point. We try to address this challenge using image processing techniques. This is because commercialized medium-sized head-feeding harvesters are equipped with a color camera, and almost all of them are controlled by microcomputers. Hence automation using image processing requires little hardware modification, and thus is less expensive than other methods.

In various fields within computer science, such as image-based sensing and virtual reality (VR), augmented reality (AR) has recently attracted increasing attention (Sato and Yokoya, 2008). VR is a technology that constructs a virtual domain using computer graphics, while AR is an expansion of VR, which overlays virtual objects onto real images captured by cameras. Persons who watch the monitors projecting such overlaid images observe the VR-supplemented world (Billinghurst et al., 2001).

The greatest challenge in realizing AR applications involves the precise registration of virtual images with their real images, and at the same time, sufficiently fast image processing is essential for the continuous and natural perception of the AR domain (Kato and Billinghurst, 1999). ARToolKit was developed by Kato and Billinghurst (1999) to support the creation of AR applications having sufficient accuracy and speed. It uses a square board as a marker, detects the marker, and calculates the relationship between a camera and the marker for rendering virtual images (Kato and Billinghurst, 1999; Kato et al., 2000; Farbiz et al., 2005). This function can be utilized for the fast and accurate detection of the grain container and the proper positioning of the spout.

This study aims to first present a method for the automated positioning of the auger using image processing; then to measure the positioning accuracy; and finally, to study the effectiveness and potential offered by this method.

II Materials and Methods

1. Experimental Devices and Development Environment

The experimental devices used in this study include (1) a head-feeding combine harvester, VY50 CLAM (Mitsubishi Agricultural Machinery Co., Ltd, Shimane, Japan); (2) a USB camera, UCAM-DLA200H (ELECOM Co., Ltd. Osaka, Japan); (3) a total station, SET4100s (Sokkia Co., Ltd. Tokyo, Japan); (4) a Marker (a plywood and an aluminum board).

The USB camera is epoxied onto an aluminum plate, which is screwed onto the auger so that it does not move or oscillate. The total station is used for measuring the positions of the spout and marker. The marker is placed on the roof of the truck. Figure 1 shows the unloading auger, the USB camera and the truck with the grain container. Figure 2 shows the marker, on which the letter “A” is printed.

![Fig. 1 Combine harvester and truck with container](image1)

![Fig. 2 Marker (arrows indicate 400 mm × 400 mm)](image2)

The software program used in this study is developed for image processing and for calculating the correct attitude angles. The software program utilizes the image processing libraries of OpenCV (Intel Corp.) and ARToolKit (Kato and Billinghurst, 1999), and conveys the VR information to an operator’ display using OpenGL (Silicon Graphics International Corp.) with the programming language C++ (Microsoft Corp. Visual studio 2008).

2. Calculation of Attitude Angles

The unloading auger of the harvester used in this study is modeled as a two-degree-of-freedom (2-DoF) manipulator (Figure 3). Joint 1 rotates at an angle of $-110^\circ < \theta_1 < 90^\circ$. However, the unloading range in this study is limited to $0^\circ < \theta_1 < 90^\circ$ because the actual unloading task takes place within this range. Joint 2 rotates at an angle of $0^\circ < \theta_2 < 50^\circ$.

Figure 4 shows the link structure and three coordinate systems $\Sigma_0$, $\Sigma_1$ and $\Sigma_2$, which are all right-handed systems.
Table 1 Link parameters of the auger

<table>
<thead>
<tr>
<th>Link</th>
<th>Length [mm]</th>
<th>Twist Angle [degree]</th>
<th>Distance between the links [mm]</th>
<th>Attitude Angle [degree]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( \theta_1 )</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>(-90)</td>
<td>( l_a )</td>
<td>(-\theta_2)</td>
</tr>
</tbody>
</table>

\( l_a \) is the offset distance between the origins of \( \Sigma_1 \) and \( \Sigma_2 \) along the \( Y_1 \) axis. \( l_1 \) and \( l_2 \) represent the length from the origin of \( \Sigma_2 \) to the spout and to the camera, respectively. \( l_1 \) and \( l_2 \) are the distance from the \( X_2 \) axis to the camera in the \(-Z_2\) direction and \( Y_2 \) direction, respectively.

Table 1 shows the link parameters of the auger. \( Z_0 \) is the initial coordinate system \((X_0, Y_0, Z_0)\) before rotating the auger; \( \Sigma_1 \) is the coordinate system \((X_1, Y_1, Z_1)\) rotated by the angle \( \theta_1 \) around the \( Z_0 \) axis in the system \( Z_0 \) using the homogeneous transformation matrix \( ^0T_1 \), expressed by equation (1); and \( \Sigma_2 \) is the coordinate system \((X_2, Y_2, Z_2)\) rotated by the angle \(-90^\circ \) around the \( X_1 \) axis, \(-\theta_2 \) around the \( Z_2 \) axis, and then translated by \( l_a \) in the \( Y_1 \) direction using \( ^1T_2 \) which is expressed by equation (2). \( \Sigma_{Camera} (X_C, Y_C, Z_C) \) is the coordinate system expressed by the transformation matrix \( ^2T_{Camera} \). In the transformation from \( \Sigma_2 \) to \( \Sigma_{Camera} \), \( \Sigma_2 \) is translated by \( l_a \) in the \( X_2 \) axis, \( l_1 \) in the \( Y_2 \) axis, and \(-l_2 \) in the \( Z_2 \) axis. It then rotates by \( 90^\circ \) about the \( Y_2 \) axis, \(-90^\circ \) about the latest \( X_2 \) axis, and \( \alpha \) about the latest \( X_2 \) axis.

\[
^0T_1 = \begin{bmatrix}
\cos \theta_1 & -\sin \theta_1 & 0 & 0 \\
\sin \theta_1 & \cos \theta_1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] (1)

\[
^1T_2 = \begin{bmatrix}
\cos \theta_2 & -\sin \theta_2 & 0 & 0 \\
0 & 1 & 0 & l_a \\
-\sin \theta_2 & \cos \theta_2 & 0 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] (2)

\[
^2T_{Camera} = \begin{bmatrix}
0 & -\sin \alpha & \cos \alpha & l_c \\
0 & \cos \alpha & \sin \alpha & 0 \\
-1 & 0 & 0 & -l_d \\
0 & 0 & 0 & 1
\end{bmatrix}
\] (3)

Camera \( ^0T_{Marker} \) is obtained by a function of ARToolKit. Therefore, \( ^0T_{Marker} \) can be derived as follows.

\[
^0T_{Marker} = ^0T_1 \cdot ^1T_2 \cdot ^2T_{Camera} \cdot T_{Camera} \cdot T_{Marker}
\] (4)

If we abbreviate the vector of the target point on \( \Sigma_{Marker} \) as \( ^0r_{Target} \), then \( ^0r_{Target} \) is described with \( ^0T_{Marker} \) as,

\[
^0r_{Target} = ^0T_{Marker} \cdot ^0r_{Target}
\] (5)

where \( ^0T_1 \) and \( ^1T_2 \) are given by \( \theta_1 \) and \( \theta_2 \), respectively, corresponding to the present auger’s position. \( ^2T_{Camera} \) is determined by the camera’s clamping angle \( \alpha \).

\( ^0r_{Spout} \), which is the positional vector of the spout on \( \Sigma_0 \), is written with \( ^0r_{Spout} \) on \( \Sigma_2 \) as follows:

\[
^0r_{Spout} = ^0T_1 \cdot ^1T_2 \cdot ^2T_{Camera} \cdot T_{Camera} \cdot T_{Marker} \cdot ^0r_{Spout}
\] (6)

Figure 5 shows the geometric relations among the spout, target point, and calibration marker, as well as the two coordinate systems \( \Sigma_0 \) and \( \Sigma_{Marker} \). The coordinate system \( \Sigma_{Marker} \) is a calibration marker based coordinate system generated by \((X_M, Y_M, Z_M)\).
Considering the fact that $r_{\text{Spout}}$ is expressed in equation (7),

$$
2r_{\text{Spout}} = [I_b \ 0 \ 0 \ 1]^T
$$

equation (6) gives

$$
r_{\text{Spout}} = 
\begin{bmatrix}
I_b \cos \theta_1 \cos \theta_2 - l_a \sin \theta_1 \\
I_b \sin \theta_1 \cos \theta_2 + I_a \cos \theta_1 \\
-I_b \sin \theta_2 \\
1
\end{bmatrix}
$$

If we denote the vector $0r_{\text{Target}}$ by $[x_0 \ y_0 \ 0 \ 1]^T$, which is calculated according to equation (5), then $\theta_1$ and $\theta_2$ are determined as follows:

$$
\theta_1 = \tan^{-1} \frac{x_0}{y_0} - \sin^{-1} \left( \frac{l_a}{\sqrt{x_0^2 + y_0^2}} \right)
$$

$$
\theta_2 = \cos^{-1} \left( \frac{x_0^2 + y_0^2 - l_a^2}{I_b} \right)
$$

which are the desired attitude angles.

Again, a target point is initially set on $\Sigma_{\text{Marker}}$. The vector of the target point on $\Sigma_00^0r_{\text{Target}}$ is calculated on the basis of $0^1T_1$, $1^2T_2$, and $2^3T_3$ which are determined by the link structure and its parameters (Figs. 3 and 4), and $3^4T_4$, which is obtained from the camera images using ARToolKit. The positional vector of the spout on $\Sigma_00^0r_{\text{Spout}}$ is calculated in equation (6). Finally, by solving the equations $0r_{\text{Spout}} = 0r_{\text{Target}}$ (about $x$ and $y$ coordinates), the attitude angles are determined such that the spout moves to the target point.

### 3. ARToolKit

ARToolKit is a computer program library for C/C++, developed by Kato and Billinghurst (1999), to enable the simple creation of AR applications. It detects markers and gives homogeneous transform matrices from the markers’ coordinates to the camera’s coordinates. In image processing, the black rectangular region in which a white subregion is included is first extracted (Fig. 2). The region is normalized and the image in the subregion is compared by template matching with patterns that are initially provided to the algorithm to identify a specific marker. Therefore, it is preferred to use a simple and longitudinally asymmetric design as the used image. Thus, in this study, we used the letter “A” as the design to be printed on the marker.

Our method utilizes ARToolKit to detect a marker and to obtain the homogeneous transformation matrix $3^4T_4$. It allows us to determine the marker’s position and orientation, and then calculate the optimal attitude angles for positioning.

### 4. Prerequisites and Experimental Procedures

Our experiment requires certain assumptions. First, the harvester is parked in a paddy alongside the truck that is located on a farm road along the paddy. Second, the marker lies within the line of sight of the camera, and the spout can physically reach the grain container.

A combine control program is used to send commands to the combine. In addition, the software program developed in this study is used for image processing and calculating the proper attitude angles. The experiment is performed according to the following steps: (1) start the software program when the auger is at a state in which the camera can catch the entire marker; (2) manually input the attitude angles into our combine control program and control the auger; (3) manually measure the position of the spout and the marker with the total station. For the four types of the harvester-truck parking patterns shown in Fig. 6, the spout and marker positions are measured five times for each pattern.

### 5. Accuracy Specification

In the unloading process, it is important to eliminate any loss. To unload grains without any loss, there needs to be sufficient room between the grain stream and both side edges of the container. Figure 7 shows the dimensions of the grain container and the maximum width of the grain stream.

![Fig. 6 Relative positions between the harvester and pickup truck (grain container)](image)

![Fig. 7 Dimensions of the container and the maximum width of the grain stream](image)
From the actual unloading process, about 0.3 m of room would be adequate to meet our requirement of realizing no loss. The maximum width of the grain stream is about 0.35 m. Therefore, the spout should be positioned within a range of about 0.45 m × 0.45 m. We then set ± 0.2 m (i.e., 0.40 m × 0.40 m) as an acceptable error for our autostopositioning system.

Figure 8 illustrates the overhead view of the marker and grain container. We set a target point at \((x, y) = (1.5, 0)\) on the coordinate system shown in the figure. The point is initially assigned in the software program.

![Fig. 8 Marker and grain container](image)

**III Results and Discussion**

Figure 9 shows the positioning accuracy, which shows the difference between the spout calculated by input angles and that calculated by the actual angles after positioning.

![Fig. 9 Positioning accuracy](image)

Table 2 Positioning errors of the spout

<table>
<thead>
<tr>
<th>x [m]</th>
<th>y [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>0.180</td>
</tr>
<tr>
<td>Mean</td>
<td>0.082</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.101</td>
</tr>
</tbody>
</table>

Some possible sources of these errors may be the resolution of the camera, positioning accuracy, flexure of marker, and measurement precision of the camera’s clamping angle and the lengths of the link parameters. In particular, the camera’s resolution and the positioning accuracy are believed to be critical. Although the marker flexure and measurement precision may be sources of errors, they are not as important when compared to resolution and positioning accuracy,

Figure 10 shows the experimental results for all trials. Positions 1 to 4 correspond to the relative positions depicted in Fig. 6. Each value indicates the position of the spout after positioning, which is measured by the total station, and is represented on the coordinate system shown in Fig. 8. In Fig. 10, the square red point denotes the target point (1.5, 0) (Fig. 8) that is set in the software program. The dashed-line-square represents the acceptable error range, \((1.5 \pm 0.2, 0 \pm 0.2)\).

RMSE is about 0.1 m in both x and y direction. The maximum error (0.180, 0.168) is also less than 0.2 m (the acceptable error range) in both directions.
because we have tried two types of marker, namely a plywood marker (relatively flexible) and an aluminum board (relatively flat), neither of which make a significant difference. We have measured the camera’s clamping angle and the lengths $l_a$, $l_b$, $l_c$, $l_d$ six times. However, their values are almost the same. In addition, the measurement precision values are a few centimeters and few degrees, which geometrically result in an error of only a few centimeters at the spout. With respect to positioning accuracy, this can result in a maximum error of about 5-6 cm. The harvester cannot perfectly position its auger at the input angles (Fig. 9), which is mainly due to the inertia of the auger’s movement. Between the relative positions in Fig. 10, the measured values are scattered in the x directions for position 1 and 3. We assume that the inertia may be the cause of the scattering (Fig. 6). It is inevitable for the proposed system to include this type of error.

In this study, ARToolKit is used to obtain the homogeneous transformation matrix $T_{\text{Camera}}$. The auger’s attitude angles are calculated at about 20 frames per second, and the latest average value of the angles for every 30 frames is used for positioning. To confirm the magnitude by which the calculated angles vary, we keep the auger static under the condition where the marker is caught by the camera and obtain the calculated attitude angles. The standard deviations of the angles calculated with 350 frames are 0.01° and 0.01° for $\theta_1$ and $\theta_2$, respectively. Therefore, the errors of the attitude angles calculated using ARToolKit are sufficiently small, and therefore, ARToolKit enables a sufficiently reliable positioning of the auger.

The error sources mentioned above are believed to cause the errors shown in Table 2. However, the spout is positioned within the acceptable error range in all trials.

**IV Conclusion**

In this study, we proposed a new method for detecting a grain container and positioning the spout of a harvester at an appropriate point. For this purpose, a software program was developed that detects a marker from images obtained with a USB camera attached to the auger. The auger was modeled as a 2-DoF manipulator. The software program also calculated the correct attitude angles for positioning using the images. In the experiments performed, the attitude angles were determined using this method, and the coordinates of the spout were measured with a total station after positioning.

The results showed that the method can position the auger with sufficient precision. The RMSE was about 0.1 m and the maximum error was within 0.2 m in both x and y directions. These values fulfill the accuracy specification, and there is therefore sufficient room between the grain stream and container edges. The main sources of error are considered to be camera resolution and positioning accuracy. Therefore, the positioning of the auger could be more accurate using a higher-resolution camera.

**References**


(Received : 22. February. 2012, Accepted : 14. May. 2012)