EAEF Engineering in Agriculture, Environment and Food
published by AABEA
Research Paper

EAEF 5(3) : 96-101, 2012

Development of a Tractor Navigation System Using Augmented Reality*

Yutaka KAIZU*[1], Jongmin CHOI*[2]

Abstract

We developed an intuitive tractor navigation system using augmented reality (AR) by superimposing a computer-generated virtual three-dimensional (3D) image on a camera image. The 3D image was generated using the tractor position and direction determined by two real-time kinematic global positioning systems and an inertial measurement unit. The positioning accuracy of the AR navigation system was examined experimentally by changing the roll, pitch, and yaw angles of a tractor at rest on an actual field. The positioning errors in the world coordinate system were less than 3 cm within 3 m from the front of the tractor, and less than 3 pixels in the image coordinate system. The refresh frequency of the AR image was 30 Hz and the time taken from image capturing to displaying was within 10 ms.

[Keywords] tractor navigation system, augmented reality, safety, digital model, three-dimensional image, nighttime farming

I Introduction

Recently, the farming population has been decreasing in Japan. Many farmlands have been abandoned, and the total number of farmlands is decreasing. As a result, Japan's food self-supporting ratio has reached below 40%. After World War II, many agricultural machines such as tractors, combine harvesters, and rice planters were introduced, which lead to a dramatic improvement in productivity. The horsepower of a Japanese domestic tractor is more than 100 HP, and further up-scaling of machines might not contribute in saving labor or increasing efficiency. To maintain farming production levels in Japan, it is essential to increase the income from farming compared to other industries to attract young people. At present, since no drastic improvement can be expected with machine efficiency itself, one solution toward utilizing the land effectively is to work during the time not normally considered as work time: nighttime. Using this time efficiently will enable farmers to enhance their farmland productivity.

Nighttime farming has various merits other than increasing the machine operating time. At night, lower temperature and absence of sunlight can reduce fatigue in farmers and can also minimize pesticide drift owing to less wind. In addition, it is reported that nighttime plowing decreases weed germination (Buhler, 1997) (Hartmann and Nezadal, 1990).

Although it is not a regular practice, farmers do work at night during very busy periods of the year. However, farming at night is not recommended. This is because the field of view during nighttime is limited as compared to that during daytime, which can lead to unexpected accidents such as collisions, falling, or being run over. Even a familiar field looks unfamiliar at night. Under artificial lighting, everything looks flat and an operator's range of vision might be limited to a short distance. Thus, operators might get disoriented and lose their way in the field. In other industries, working at night is very common. However, unlike a factory or an office, where sufficient amount of light is available, an agricultural field is very dark during the nighttime. As Japanese agriculture faces many critical issues, restricting agricultural operations only to the daytime must be altered. Hence, new technologies should be developed to ensure the safety and efficiency for farmers working during the nighttime.

Even during daytime, driving a machine straight and with equal spacing for hundreds of meters in a large field is not a simple task. Various navigation systems have been developed and are available in the market to assist machinery operators (Li et al., 2009). Major agricultural machinery companies and global positioning system (GPS) companies have been selling navigation systems. These navigation systems use GPS and an inertial measurement unit (IMU) to locate the position and orientation of a tractor. A computer terminal calculates displacement from a formally planned path, and guides the operator or steers the wheel automatically. These systems aid nighttime operations. However, at present, the navigation

* Partly presented at Fourth IFAC International Workshop on Bio-Robotics, Information Technology, and Intelligent Control for Bioproduction Systems in September 2009
*1 JSAM Member, Corresponding author, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo, 113-8657, Japan ; yuta_kaizu@yahoo.co.jp
*2 Graduate School of Agriculture, Hokkaido University, Kita-9, Nishi-9, Kita-ku, Sapporo, Hokkaido, 060-8589, Japan

image displayed on the terminal is an unrealistic computer-generated image (CGI) and does not project the surroundings realistically. In an actual environment, encountering unexpected obstacles is quite possible. Hence, if an actual image is overlapped on the navigation image, it will be more intuitive and easier to perceive.

The technique that blends computer graphics (virtual reality) with the real-world environment in real time is called augmented reality (AR) or mixed reality (MR) (Furht, 2011). In this technique, text or objects are drawn on an actual image to help people understand reality. A common example of AR occurs in sports programs on television. A computer-generated topographical mesh is drawn on the putting green in a golf program, or virtual advertisements are displayed on an actual wall of a baseball or football stadium. In addition, AR is used in many other areas such as military, aerospace, video games, surgery (Liao et al., 2010), and machine maintenance. Recently, AR has been applied to a parking assist system for automobiles.

The objective of our study is to develop an intuitive tractor navigation system using AR. To precisely blend a video image with computer graphics, the position and orientation should be known in real time. The intrinsic parameters (such as focal length and format size) of the camera should also be known precisely. This paper describes a developed system, as well as its calibration and static accuracy.

II Materials and Methods

1. Tractor navigation system

Fig. 1 shows a schematic of the tractor navigation system. First, a three-dimensional (3D) model was developed by surveillance using a real-time kinematic GPS (RTK-GPS). The 3D coordinates of the corners of the field, edges of the road, and other objects were located. This allows any landmark to be included in the model. Next, paths were generated at regular intervals. These processes were completed before actual operation of the tractor on the field. Two RTK-GPSes (Trimble MS 750) were used to measure the yaw angle and position of the tractor. An IMU determined the roll and pitch angles. From these data, the position and orientation of the actual camera were estimated. A virtual camera was positioned at exactly the same coordinates as those of the actual camera. From that viewpoint, the points and lines in a 3D model projected on an image plane should be matched with the actual landmarks in an image taken by the actual camera. Since the image taken by the actual camera was distorted due to the characteristics of the lens, it was rectified and then mixed with the virtual perspective image. The resulting image is called the AR image. In this image, a path and other objects are drawn on the real field image. The operator watches the display while driving the tractor.

Fig. 2 shows an example of the AR image created in this study. A 3D model shows the left and right sides of the farm road. This model is recorded as a sequence of 3D coordinates. Looking from the point indicated by a yellow arrow in Fig. 2(a), the 3D model appears as a virtual perspective image in Fig. 2(b), while the actual image appears in Fig. 2(c). The synthesized image (Fig. 2(d)) shows the edge of the road and an unseen bridge.

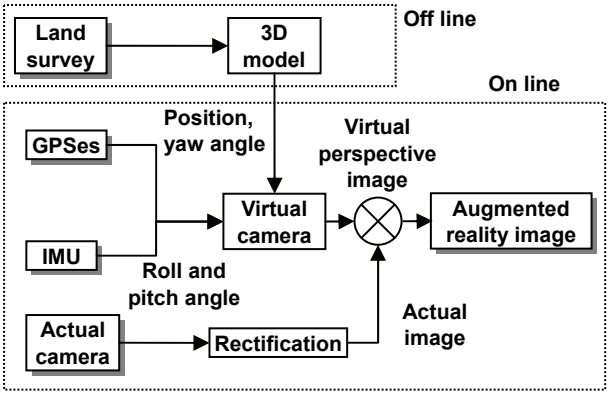

Fig. 1 Schematic of the tractor navigation system

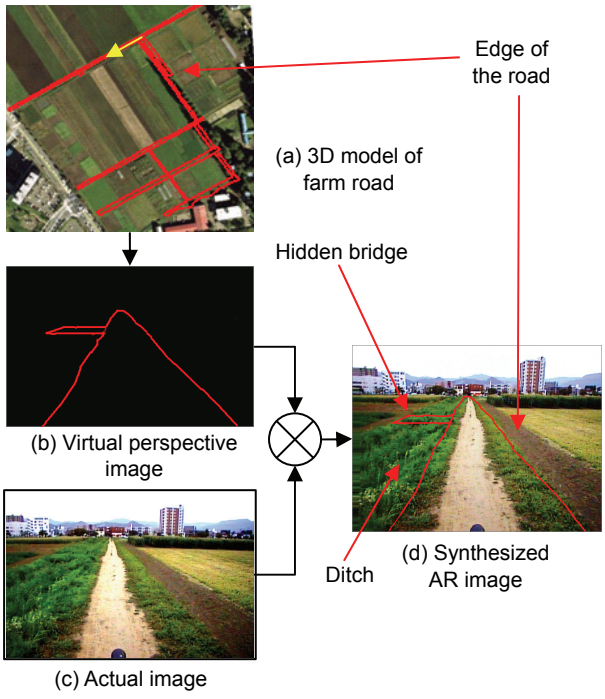

Fig. 2 Example of the AR image

Fig. 3 shows a tractor equipped with the AR navigation system, which includes GPS antennae, an IMU, a PC, and an IEEE 1394 digital color camera (SONY DFW-VL500). They were mounted on an aluminum frame on the tractor roof. The camera was inclined forward. A wide conversion lens was

attached on the camera to broaden the field of view. The camera interface was IEEE 1394, which allowed the camera settings to be controlled programmatically from a computer. National Instruments LabVIEW 8.5 was used for programming. For the subroutines that required much calculation, Microsoft Visual Studio 2008 was used to make a dynamic link library (DLL). The resolution of the image was 640 pixels × 480 pixels. The refresh frequency of the AR image was 30 Hz. This frequency was restricted by the camera. A monitor was placed in front of the steering wheel. The time taken from image capturing to displaying was within 10 ms.

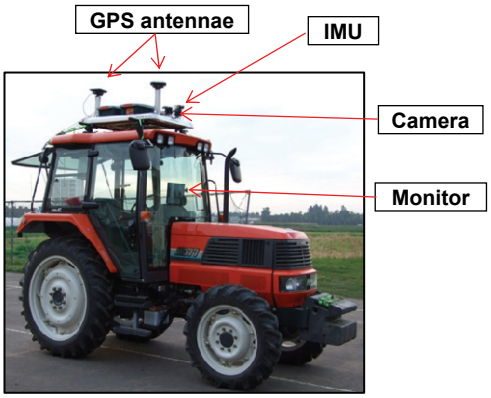

Fig. 3 AR tractor navigation system

2. Definition of frames of reference for the tractor navigation system

In this system, frames of reference were defined as shown in Fig. 4. The universal transverse mercator (UTM) system was used for the world coordinates $X_w - Y_w - Z_w$. Its X_w, Y_w, and Z_w axes are the easting, northing, and altitude of the UTM, respectively. The origin of the reference frame of the tractor was set at the point right under the center of the rear axle. The positions of the tractor $(X_{wt}, Y_{wt}, Z_{wt})^T$ and the camera $(X_{wc}, Y_{wc}, Z_{wc})^T$ in the world reference frame were defined as follows:

$$\begin{bmatrix} X_{wt} \\ Y_{wt} \\ Z_{wt} \end{bmatrix} = \begin{bmatrix} X_{wGPS1} \\ Y_{wGPS1} \\ Z_{wGPS1} \end{bmatrix} + E_{x180} E_{z\psi} E_{wt} \begin{bmatrix} x_{tGPS1} \\ y_{tGPS1} \\ z_{tGPS1} \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} X_{wc} \\ Y_{wc} \\ Z_{wc} \end{bmatrix} = \begin{bmatrix} X_{wGPS1} \\ Y_{wGPS1} \\ Z_{wGPS1} \end{bmatrix} + E_{x180} E_{z\psi} E_{wt} \begin{bmatrix} x_{cGPS1} \\ y_{cGPS1} \\ z_{cGPS1} \end{bmatrix} \quad (2)$$

$$E_{x180} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 180° & -\sin 180° \\ 0 & \sin 180° & \cos 180° \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (3)$$

$$E_{z\psi} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$E_{wt} = E_{IMU} \left(E_{IMUini} \right)^{-1} \quad (5)$$

$$E_{IMU} = \begin{bmatrix} \cos\rho & 0 & \sin\rho \\ 0 & 1 & 0 \\ -\sin\rho & 0 & \cos\rho \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \quad (6)$$

$$E_{IMUini} = \begin{bmatrix} \cos\rho_{ini} & 0 & \sin\rho_{ini} \\ 0 & 1 & 0 \\ -\sin\rho_{ini} & 0 & \cos\rho_{ini} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{ini} & -\sin\theta_{ini} \\ 0 & \sin\theta_{ini} & \cos\theta_{ini} \end{bmatrix} \quad (7)$$

Here $(X_{wGPS1}, Y_{wGPS1}, Z_{wGPS1})^T$ is the coordinate of GPS1 in the world reference frame. $(x_{tGPS1}, y_{tGPS1}, z_{tGPS1})^T$ and $(x_{cGPS1}, y_{cGPS1}, z_{cGPS1})^T$ are the initial coordinates of the tractor origin and the camera origin, respectively, as viewed from GPS1. E_{x180} is a rotation matrix around X axis for 180°. $E_{z\psi}$ is a rotation matrix around Z axis for an angle ψ. E_{wt}, E_{IMU}, and E_{IMUini} are the frames of reference of the tractor, the IMU, and the initial condition of the IMU, respectively. ψ is the yaw angle calculated from the relative positions of GPS1 and GPS2 (Alkan and Baykal, 2001). θ and ρ are the roll and pitch angles measured by the IMU, respectively. θ_{ini} and ρ_{ini} are the roll and pitch angles of the IMU when the tractor is put on a flat surface. The initial parameters of the IMU were measured simultaneously with the extrinsic parameters of the camera.

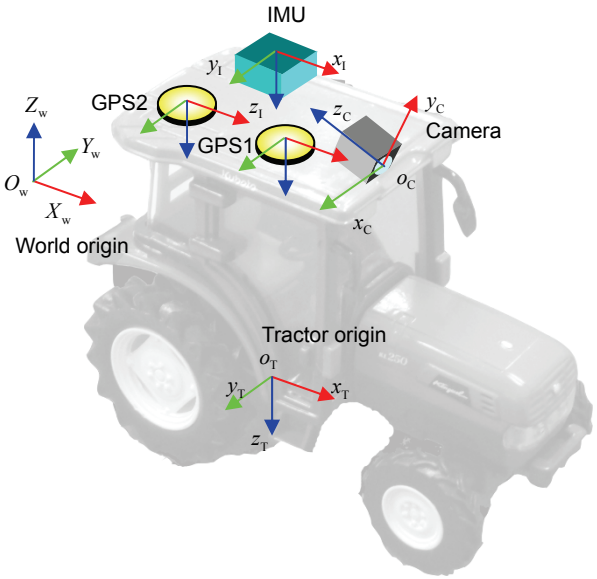

Fig. 4 Frames of reference of the system

3. Camera calibration

To match the virtual image with the actual image, the

intrinsic parameters (focal length, image center, and lens distortion coefficients) and the extrinsic parameters (position and orientation of the camera) were estimated using the camera calibration method.

(1) Estimation of the camera's intrinsic parameters

Unlike the image taken by an ideal virtual camera, the actual image includes distortion and the offset of principal points. The intrinsic parameters (focal length, principal point position, and lens distortion) of the camera must be known. To estimate the intrinsic parameters, commercial photogrammetric software PhotoModeler 6.0, from Eos Systems, was used. The images of a company-distributed test grid pattern were taken from 12 different directions according to the instructions of the software. The parameters were calculated automatically. As shown in Fig. 5, defining the coordinates in distorted image $(x_d, y_d)^T$, rectified image $(x_r, y_r)^T$ and defining principal point $(p_x, p_y)^T$, rectification was performed with the following equations (EOS Systems Inc., 2007):

$$\begin{bmatrix} x_r \\ y_r \end{bmatrix} = (1 + d_r) \begin{bmatrix} x_d - p_x \\ y_d - p_y \end{bmatrix} \tag{8}$$

$$d_r = K_1 r^2 + K_2 r^4 + K_3 r^6 \tag{9}$$

$$r^2 = (x_d - p_x)^2 + (y_d - p_y)^2 \tag{10}$$

where K_1, K_2, and K_3 are distortion constants defined from calibration. In the developed program, to shorten the processing time, the distorted image was transformed on the basis of predefined lookup table images in which the destination coordinates of the *x* and *y* axis were stored.

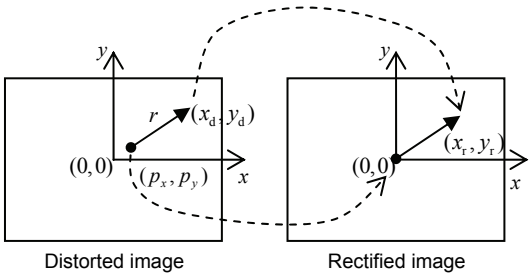

Fig. 5 Image rectification

(2) Estimation of the camera's extrinsic parameters

The extrinsic parameters of a camera include its XYZ coordinates and the roll, pitch, and yaw angles. They can be estimated from captured images and known landmarks. A marker was placed in front of the camera and moved to different positions, and these positions were measured by a total station (TS). An accurate position of the GPS antennae, front and rear center points of the tractor, and right and left center points of the rear axle were also measured using the TS. The initial inclination of the IMU was recorded at the same time. The extrinsic parameters were calculated by comparing the coordinates of the center of the markers in an image and their real coordinates measured by the TS. Fig. 6 is a composite photograph of 13 images of a single marker placed in different positions. In this image, the poles that supported the markers were deleted to make it easier to view. Fig. 7 shows a 3D view of the estimation result.

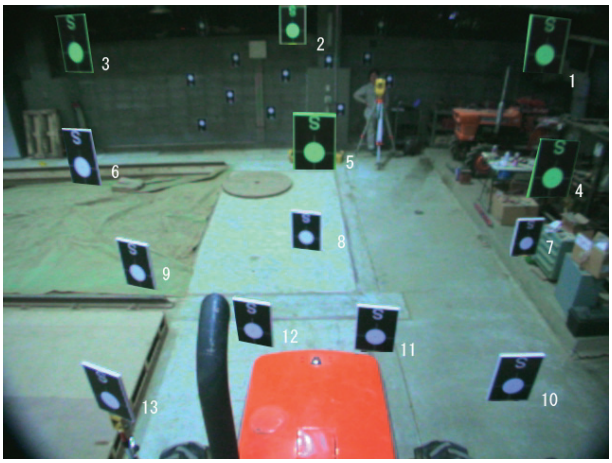

Fig. 6 Image of the markers used to estimate the extrinsic parameters

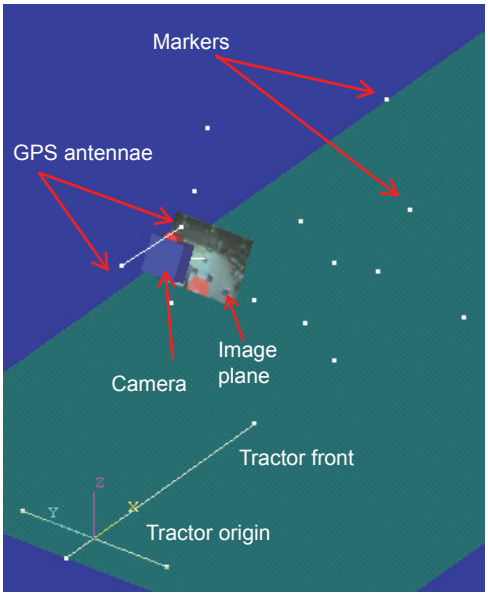

Fig. 7 Three-dimensional view of the estimation result

III Results and Discussion

To verify the accuracy of the match between the virtual and actual images, the following experiment was conducted. As shown in Fig. 8, twenty-five markers were placed on corners of a 4 × 4 grid with a 2 m pitch in front of the tractor. Marker

C was located at the center of the front row of the grid. The tractor was moved to different places and inclined at various attitudes. Table 1 shows a combination of distance, roll angle, pitch angle, and yaw angle in the experiments.

The blue and red lines drawn over the image form the virtual grid (Fig. 8). This virtual grid was generated on the ongoing GPS and IMU data. If the position and orientation of the camera were correctly determined, the intersections of the grid would appear exactly at the center of each marker. The accuracy of camera estimation was verified by comparing the virtual grid and the markers.

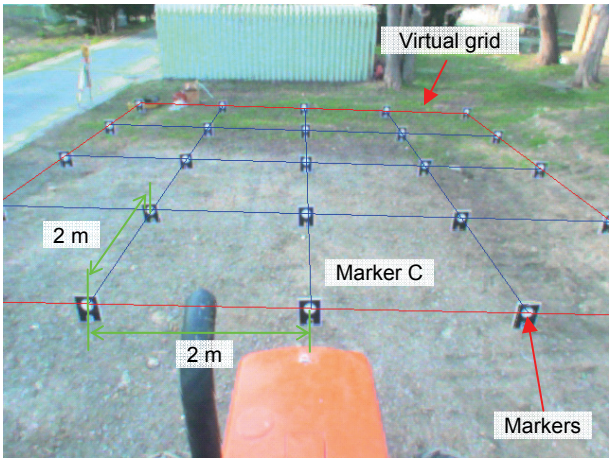

Fig. 8 Accuracy verification of image matching

Table 1 Position and attitude of the tractor

Changing parameters	Number	Horizontal distance between the camera and marker C [m]	Roll [°]	Pitch [°]	Yaw [°]
Distance	1	3.4	−1.7	−1.6	−0.5
	2	5.8	−0.4	−1.7	−0.1
	3	7.3	−0.4	−1.4	0.0
	4	7.7	−0.8	−1.5	−1.1
	5	10.9	−0.1	−0.6	−1.0
Roll	6	5.4	9.8	−1.2	0.7
	7	5.4	4.2	−1.3	−0.5
	8	5.1	−6.6	−1.6	−0.1
	9	5.6	−11.1	−1.4	−0.7
Pitch	10	5.3	−1.5	8.0	−1.4
	11	5.7	−1.0	3.8	−1.2
	12	4.9	−1.7	−7.1	−1.4
	13	5.3	−1.6	−10.6	−1.5
Yaw	14	6.6	0.5	−1.3	42.5
	15	6.9	1.2	−1.4	40.3
	16	5.5	0.5	0.1	28.9
	17	5.9	0.0	−0.5	26.0
	18	6.4	−2.0	−0.6	−26.6
	19	6.6	0.0	−0.5	-−46.4

As the positions of the markers were already known, the photogrammetric software could estimate the positions and orientations of the camera in the same way as the estimation of the extrinsic parameters. The estimated positions and orientations of the camera were compared with those estimated by the GPSes and the IMU. Table 2 shows the results.

Table 2 Accuracy of the camera position and orientation estimation

	X [cm]	Y [cm]	Z [cm]	Roll [°]	Pitch [°]	Yaw [°]
Mean	2.49	0.45	2.89	−0.33	0.35	−0.12
RMS	2.60	1.26	3.25	0.34	0.36	0.34
STDV	0.78	1.21	1.54	0.095	0.088	0.32
Max	4.03	4.10	5.60	0.59	0.51	0.91

Considering that the horizontal and vertical accuracies of the GPSes are 2 cm and 3 cm, respectively, the errors in the X, Y, and Z axes were acceptable. The root mean square (RMS) errors in roll and pitch angles were 0.34° and 0.36°, respectively. Since the nominal accuracy of the IMU is 0.2° in roll and pitch, errors were larger than expected. However, the standard deviations of the errors were less than 0.1°. This implies that these measured angles were biased. The IMU used in this experiment was a fiber-optic gyroscope. Basically, this type of IMU consists of a drift error. If we can compensate for this error, the accuracy of the position and orientation estimation will improve. The accuracy of the yaw angle was better than expected because considering that the baseline length between the antennae was 94 cm, the 0.3° error meant that the GPS positioning error was less than 0.5 cm. By increasing the baseline length, higher accuracy can be expected.

Fig. 9 shows the distances from the camera to the markers plotted against the horizontal error distance in the world coordinates system between the markers and the reversely projected intersections of the virtual grid. The line indicates the average error within each one meter of horizontal distance. As the distance increased, the projection error increased simultaneously. This is because the area equivalent to 1 pixel in an image increases with an increase in distance. The error in front of the tractor was less than 5 m. In fact, assuming that an AR is used for a tractor navigation system, high accuracy is required within an area less than 10 m from the tractor. Therefore, the accuracy obtained was suitable for navigation.

Fig. 10 shows the distances from the camera to the markers plotted against the error distance between the markers and the cross sections in the images, measured in units of pixels. The

error did not increase with distance. The projection errors in images were caused solely by the error in estimating the camera's position and orientation. The averaged error was acceptable.

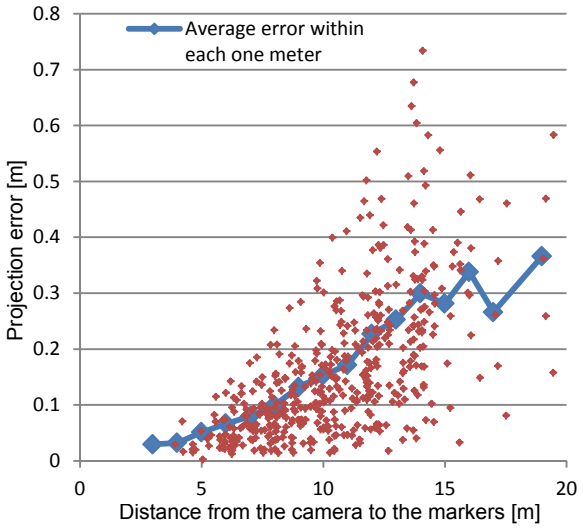

Fig. 9 Distance from the camera to the markers plotted against the projection error distance in the world coordinates

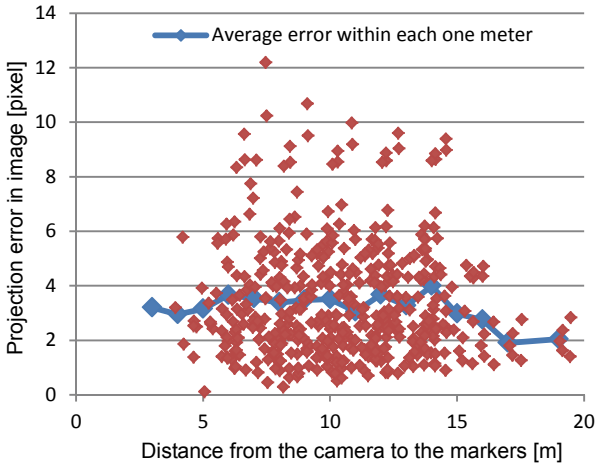

Fig. 10 Distance from the camera to the markers plotted against the error distance between the markers and the cross sections in the image

IV Conclusions

The following conclusions were derived from the results and discussion.

1. The developed AR system was capable of mixing a computer-generated 3D model with an actual video image in real time. The refresh frequency of the AR image was 30 Hz and the delay from image capturing to displaying was within 10 ms.

2. The position and orientation of the camera were accurately estimated using data from the two RTK-GPSes and the IMU. The estimated roll and pitch angles appeared to include bias errors.

3. The average error in distance between the markers and the intersections of the virtual grid was 3 cm, where the distance from the camera to the markers was about 3 m and 40 cm at approximately 19 m. This value was accurate enough for navigation.

4. The average error in distance between the markers and the intersections of the virtual grid in the image was 3 pixels, where the distance from the camera to the markers was about 3 m and 2 pixels at approximately 19 m. The error did not increase with distance.

Acknowledgment

We gratefully acknowledge Mr. Munetaka Iio for helping us with the experiment. This study was supported by Grant-in-Aid for Young Scientists (B) 19780193.

References

Alkan, R. M. and O. Baykal. 2001. Survey boat attitude determination with GPS/IMU systems. Journal of Navigation 54:135-144.

Buhler, D. D. 1997. Effects of tillage and light environment on emergence of 13 annual weeds. Weed Technology 11:496-501.

EOS Systems Inc. 2007. Radial lens distortion; Photomodeler 6 help file.

Furht, B. 2011. Handbook of Augmented Reality. New York, Springer.

Hartmann, K. M. and W. Nezadal. 1990. Photocontrol of weeds without herbicides. Naturwissenschaften 77:158-163.

Li, M., K. Imou, K. Wakabayashi and S. Yokoyama. 2009. Review of research on agricultural vehicle autonomous guidance. International Journal of Agricultural and Biological Engineering 2:1-16.

Liao, H. E., T. Inomata, I. Sakuma and T. Dohi. 2010. 3-D augmented reality for MRI-guided surgery using integral videography autostereoscopic image overlay. IEEE Transactions on Biomedical Engineering 57:1476-1486.

(Received : 24. October. 2011, Accepted : 19. January. 2012)