Greenhouse Robot Navigation Using KLT Feature Tracking for Visual Odometry

P. J. Younse and T. F. Burks

Dept. of Agricultural and Biological Engineering, University of Florida, Gainesville, FL, USA.

E-mail: TFBurks@ifas.ufl.edu

ABSTRACT

A visual odometer was developed for an autonomous greenhouse sprayer to estimate vehicle translation and rotation relative to the world coordinate system during navigation. Digital images were taken from a CCD camera mounted on the robot. 7 x 7 pixel features were selected in the image using the KLT algorithm (Csetverikov, 2004). Features were tracked from image to image by finding the best 7 x 7 pixel match of the feature within a 25 x 25 pixel search box. By analyzing the movement of these features, vehicle rotation and translation were estimated. Five features were tracked with the odometer. Tests were run to verify the visual odometer's accuracy during translation, rotation, and on various surfaces. The visual odometer ran at an average of 10 Hz during experimentation. Translation tests of the odometer in a lab environment gave an average error of 4.85 cm for a 30.5 cm forward translation and 12.4 cm average error for a 305 cm translation. Rotation tests of the odometer in a lab environment gave an average error of 1° for a 45° rotation and an 8° error for a 180° rotation about the vehicle z-axis. Tests completed on concrete, sand, and gravel demonstrated adaptability of the odometer on different ground surfaces that are common in greenhouses. The visual odometer was successfully integrated into a visual navigation system for intersection navigation of an autonomous greenhouse sprayer.

Keywords: Feature tracking, KLT algorithm, machine vision, vehicle navigation, visual odometry

1. INTRODUCTION

Vehicle automation is a growing interest among the agricultural community. Many feasibility studies were made on autonomous agricultural vehicles. Have et al. (2002) investigated the development of autonomous weeders for Christmas tree plantations. Design requirements for such a vehicle were explored, specific behaviors for navigation and operation defined, and a system architecture proposed. Hellström (2002) performed a similar study for autonomous forest machines as part of a project to develop unmanned vehicles that transport timber. Another application of autonomous agricultural vehicles is autonomous greenhouse spraying. Benefits of an autonomous greenhouse sprayer include increased accuracy and precision with spraying, which would contribute to more efficient use of resources, capability to operate 24-hours a day, and decreased health risks associated with human exposure to dangerous chemicals.

Singh (2004) designed an 81 x 41 cm autonomous greenhouse sprayer. The vehicle was designed to navigate through 46, 51, and 61 cm aisles. Two 560 W DC motors with 20:1 gear reducers powered two separate three-wheel drive trains. Turning was accomplished with differential steering. Vehicle control down the center of test paths was carried out independently using ultrasonic range sensors, ladar, and machine vision with a fuzzy PD controller (Singh, 2004;

Singh and Subramanian, 2004). The navigation system developed in Younse (2005) performed intersection detection and navigation using machine vision. Visual odometry was utilized in this navigation system during the intersection navigation algorithm.

Visual odometry allows estimation of vehicle translation and rotation relative to the world coordinate system during navigation solely based on images from a camera. The advantage of visual odometry over conventional odometry techniques is that it bases vehicle movement on the movement of the path relative to the vehicle and is not affected by wheel slippage or turning. Also, if the vehicle is already equipped with a camera for navigation, no additional instruments are required, reducing vehicle cost and complexity. Visual odometry has been accomplished by tracking feature points from one image to the next and then computing the vehicle translation and rotation from these points. Nistér et al. (2004) detected Harris corners, tracked these features between image frames, and used a combination of the 5-point algorithm, 3-point algorithm, and triangulation to estimate 3D camera pose, and in turn, vehicle movement. Pollefeys (2004) also discussed and compared techniques to calculate 3D information and camera motion by tracking sets of feature points through image sequences using structure and motion.

The visual odometry technique discussed in this paper differs from traditional visual odometry and structure and motion approaches in that it takes advantage of a flat ground plane assumption made for a vehicle with a fixed camera traveling on relatively flat ground. From this assumption, the visual odometry described only requires tracking a minimum two feature points (though five points were used for testing to make the system more robust and help eliminate error). Also, algorithms used to calculate vehicle rotation and translation between frames in a 3D-environment, such as the 8-point algorithm (Chojnacki et al., 2003), are not required, which allows faster processing and the ability to more easily run navigation software with the visual odometry on a single camera system in real-time.

2. SYSTEM COMPONENTS

A Sony FCB-EX7805 CCD camera was mounted onto the robotic sprayer developed by Singh (2004), as shown in figure 1.

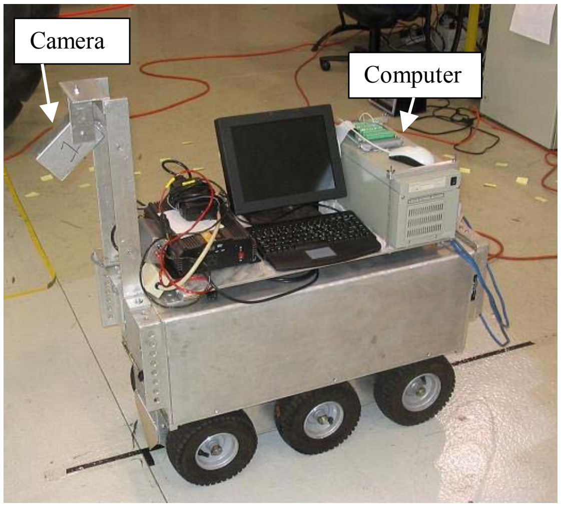


Figure 1. Camera-mounted robotic sprayer.

A single camera was chosen as opposed to a multiple camera setup or stereovision to reduce costs that come with additional equipment and reduce complexity associated with calibrating multiple cameras. An Integral Technologies Flashbus MV Pro frame grabber was used to capture 640 x 480 pixel color images from the camera. A PC with a 2.4 GHz processor acquired images from the camera and performed the visual odometry routine for each frame. All programming was implemented in C++.

3. CAMERA MODEL

Sets of intrinsic and extrinsic camera parameters were found to describe the camera. Intrinsic parameters account for the focal length, principal point, skew, and lens distortion. Extrinsic parameters account for the rotation and translation of the camera coordinate system relative to the vehicle coordinate system. Using the intrinsic and extrinsic camera parameters, ground points in the vehicle coordinate system could be transformed into pixel coordinates and vice versa, as shown in figure 2.

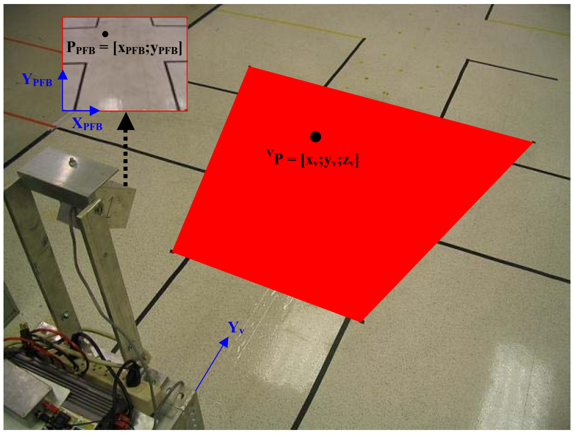


Figure 2. Relationship between a ground point in the vehicle system and image plane.

Intrinsic parameters were calculated using the Camera Calibration Toolbox for Matlab developed by Bouguet (2004). Extrinsic parameters were calculated by measuring the rotation and translation of the camera coordinate system relative to the vehicle coordinate system shown in figure 3.

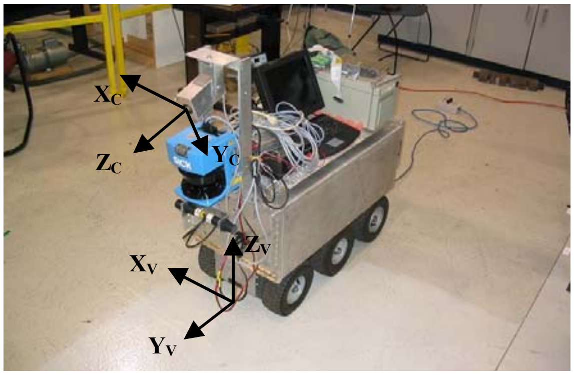


Figure 3. Camera (C) and vehicle (V) coordinate systems.

To transform pixel coordinates to points in the vehicle system, the intrinsic and extrinsic parameters were used with the assumption that points in the image exist on a fixed ground plane.

4. VISUAL ODOMETER

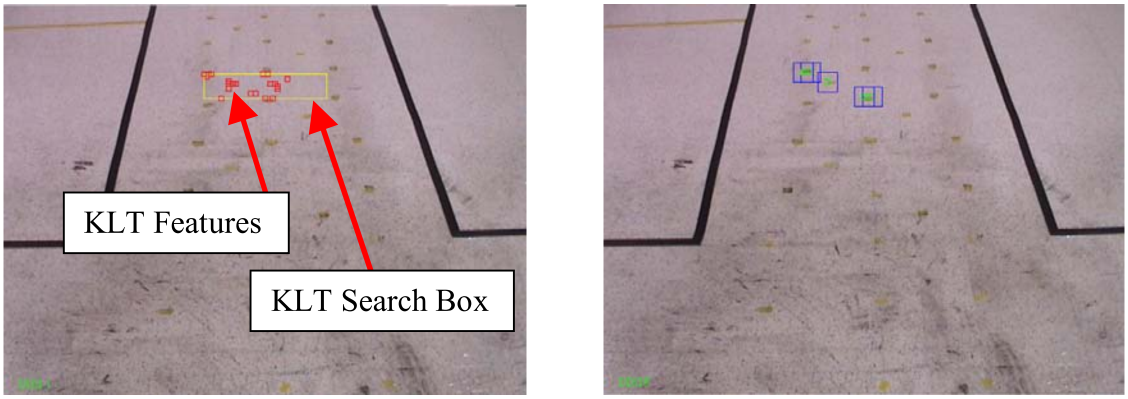
The visual odometer was developed to estimate vehicle position and orientation relative to the world coordinate system over time. The odometer determines vehicle movement by tracking features on the ground. A total of four steps are carried out in this process: odometer initialization, initial search for features, tracking of features, and the determination of position and orientation change.

4.1 Initialize Odometer

When the odometer is initialized, the world coordinate system is defined relative to the vehicle coordinate system. All translations and rotations in subsequent images will be relative to this world system.

4.2 Initial Search for Features

The odometer relies on tracking features through a set of image frames. When starting the odometer, an initial set of features must be found for tracking. The large, rectangular box in the upper-center area of the image in figure 4A shows where initial features are found. The KLT search box is 151 pixels wide, 31 pixels high, and located at pixel row 380 and column 320. These specifications were chosen to allow pixels to be tracked a large distance when the vehicle is moving forward. The KLT search box size was specified because searching the entire image for features would take too much computation time.



A)

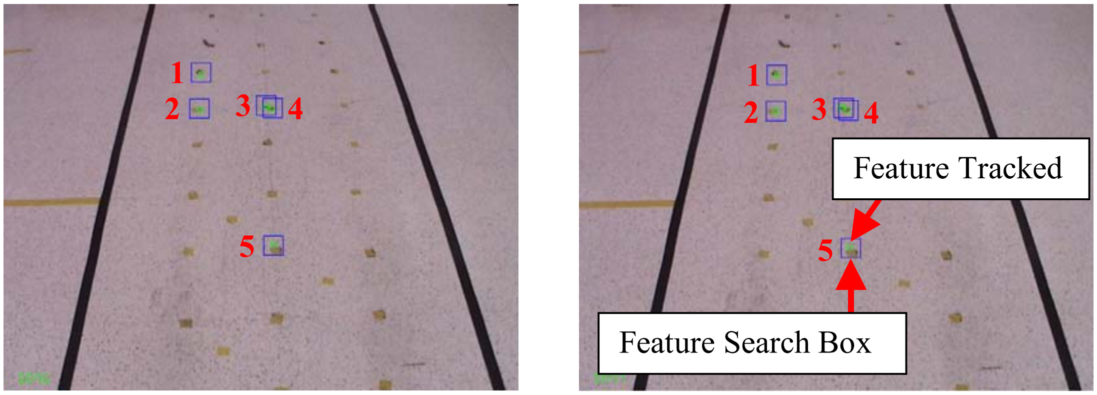
B)

Figure 4. Initial search for features. A) Good features found in search box. B) Five strongest KLT features selected for tracking.

Within the search box, good features are defined as corners found using the KLT algorithm (Csetverikov, 2004). 7x7 pixel sizes were chosen as feature sizes, because they are small enough for fast searching and large enough to uniquely define the feature. The strongest corner features found by the KLT algorithm are shown as smaller boxes within the search box in figure 4A. Because tracking features takes significant computation time, the number used in the tracking process is limited to a minimum. A minimum of two features is needed to define the vehicle motion relative to the ground. The five strongest features found with the KLT algorithm, represented by the smaller boxes in figure 4B, are used for tracking.

4.3 Tracking of Features

Figure 5 demonstrates feature tracking in two subsequent images. The small boxes mark the five 7x7 pixel features being tracked. Features are tracked by saving the 7x7 pixel feature in the first image and looking for a 7x7 pixel area that best matches it in the second image. Only the area within the larger 25x25 pixel box is searched. The size of the search box was based on the velocity of the vehicle and selected to cover the possible movement of features in the image plane from frame to frame. The location of the feature search box is based on the movement of that feature found in the last frame.



A) B)

Figure 5. Tracking features. A) First frame. B) Second frame.

To locate the features defined in image I1 (fig. 5A) in image I2 (fig. 5B), the sum-of-squared difference (SSD) is calculated for each possible 7x7 pixel area in the search box using the following equation (Barron et al., 1993):

$$SSD_{1,2}(x;d) = \sum_{j=-n}^{n} \sum_{i=-n}^{n} W(i,j)\left[I_1(x+(i,j)) - I_2(x+d+(i,j))\right]^2$$

where W is a discrete 2-d window function, x is the center pixel coordinates of the feature defined in frame 1, n = *floor(feature pixel width/2)*, and d is the pixel row and column shift used to move the 7x7 pixel area within the search box in image I_2. The best match gives the least sum-of-squared difference. Figure 6 shows a larger view of the five features tracked from the first and second frame in figure 5. The outer feature pixels are not visible due to the box drawn around the feature.

The restraint made on the features for the visual odometer is that they must be fixed on the ground plane. This restraint can be made because most greenhouse floors are relatively flat, and any small variations in the surface will be averaged out by the odometer over the course of the drive. However, certain conditions that must be avoided when tracking features are:

1. The feature tracked is significantly above or below the ground plane (due to features found in a gulley, in holes, or on mounds).
2. The feature moves independent of the ground plane (due to features from moving objects or changing shadows on the ground).
3. The feature is lost or not correctly tracked (due to a change in scenery over the feature area or errors from the feature matching process).

To prevent these conditions from occurring during feature tracking, a geometric relationship amongst the features is defined from frame to frame. The distance from each feature to all other features is measured and recorded.

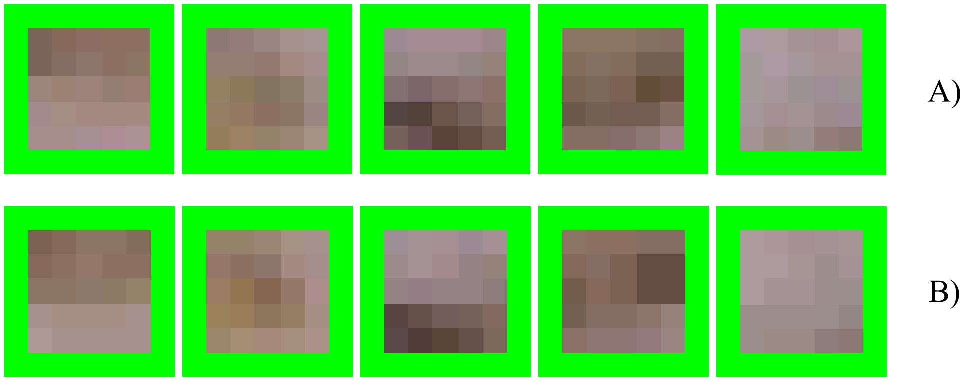


Figure 6. Large view of features matched in figure 5. A) Features 1 through 5 in figure 5A. B) Features 1 through 5 in figure 5B.

Figure 7 shows the distance calculated between features 1 and 5. ${}^{1}F_1$ and ${}^{1}F_5$ represent the coordinates of features 1 and 5 in the vehicle coordinate system in image frame 1. Note that these coordinates assume the features exist on the ground plane as defined in the camera model. ${}^{1}d_{1,5}$ represents the distance between features 1 and 5 in image frame 1.

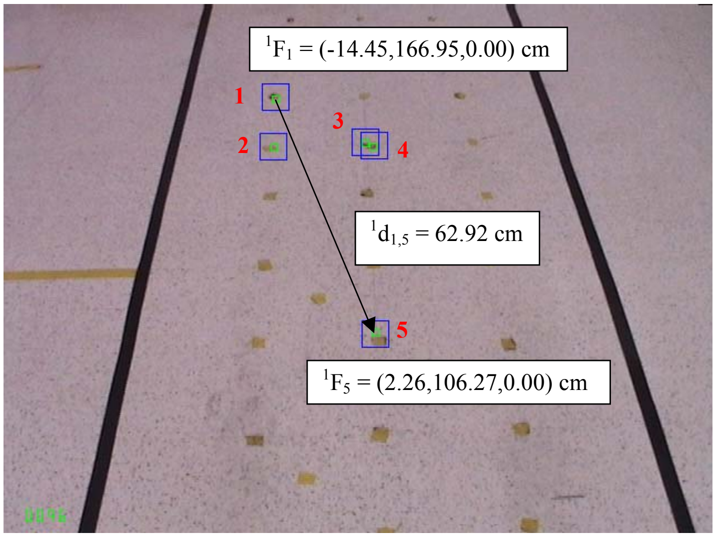

${}^{1}F_1$ = (-14.45,166.95,0.00) cm
1
2
3
4
${}^{1}d_{1,5}$ = 62.92 cm
5
${}^{1}F_5$ = (2.26,106.27,0.00) cm

Figure 7. Distance between features 1 and 5 in image frame 1.

Table 1 shows the complete distance measurements between features from frame 1 in figure 5A. When the features from the first frame are found in the second frame, the distance relationship calculations are repeated. Table 2 shows the complete distance measurements between features from frame 2 in figure 5B. Next, the difference between the distance measurements from frame 1 to frame 2 are calculated and shown in table 3. Note that this difference in distances between points is not dependent on how far the features are from one another, but based on how accurate the feature matching algorithm can estimate the position of each feature at the particular distance it is from the camera (closer features have higher pixel resolution and therefore their position can be calculated more accurately).

Table 1. Distances between features in frame 1 shown in figure 5A in cm.

Feature	1	2	3	4	5
1	${}^{1}d_{1,1} = 0.00$	${}^{1}d_{1,2} = 16.36$	${}^{1}d_{1,3} = 21.82$	${}^{1}d_{1,4} = 23.44$	${}^{1}d_{1,5} = 62.92$
2	${}^{1}d_{2,1} = 16.36$	${}^{1}d_{2,2} = 0.00$	${}^{1}d_{2,3} = 14.91$	${}^{1}d_{2,4} = 16.13$	${}^{1}d_{2,5} = 47.04$
3	${}^{1}d_{3,1} = 21.82$	${}^{1}d_{3,2} = 14.91$	${}^{1}d_{3,3} = 0.00$	${}^{1}d_{3,4} = 1.63$	${}^{1}d_{3,5} = 45.62$
4	${}^{1}d_{4,1} = 23.44$	${}^{1}d_{4,2} = 16.13$	${}^{1}d_{4,3} = 1.63$	${}^{1}d_{4,4} = 0.00$	${}^{1}d_{4,5} = 44.58$
5	${}^{1}d_{5,1} = 62.92$	${}^{1}d_{5,2} = 47.04$	${}^{1}d_{5,3} = 45.62$	${}^{1}d_{5,4} = 44.58$	${}^{1}d_{5,5} = 0.00$

Table 2. Distances between features in frame 2 shown in figure 5B in cm.

Feature	1	2	3	4	5
1	${}^{2}d_{1,1} = 0.00$	${}^{2}d_{1,2} = 16.59$	${}^{2}d_{1,3} = 21.69$	${}^{2}d_{1,4} = 23.55$	${}^{2}d_{1,5} = 62.92$
2	${}^{2}d_{2,1} = 16.59$	${}^{2}d_{2,2} = 0.00$	${}^{2}d_{2,3} = 14.86$	${}^{2}d_{2,4} = 16.03$	${}^{2}d_{2,5} = 46.84$
3	${}^{2}d_{3,1} = 21.69$	${}^{2}d_{3,2} = 14.86$	${}^{2}d_{3,3} = 0.00$	${}^{2}d_{3,4} = 1.85$	${}^{2}d_{3,5} = 45.75$
4	${}^{2}d_{4,1} = 23.55$	${}^{2}d_{4,2} = 16.03$	${}^{2}d_{4,3} = 1.85$	${}^{2}d_{4,4} = 0.00$	${}^{2}d_{4,5} = 44.37$
5	${}^{2}d_{5,1} = 62.92$	${}^{2}d_{5,2} = 46.84$	${}^{2}d_{5,3} = 45.75$	${}^{2}d_{5,4} = 44.37$	${}^{2}d_{5,5} = 0.00$

Table 3. Difference in feature distance in frame 1 (tab. 1) and frame 2 (tab. 2) in cm.

Feature	1	2	3	4	5
1	0.00	0.23	0.13	0.10	0.00
2	0.23	0.00	0.05	0.10	0.20
3	0.13	0.05	0.00	0.23	0.13
4	0.10	0.10	0.23	0.00	0.20
5	0.00	0.20	0.13	0.20	0.00

If all features were fixed to the ground and tracked accurately, the distance relationships between each of the features should remain the same. Analyzing the change in distances between the two frames compared in table 3 show little change, proving all five of the tracked features are valid.

If any of the features being tracked falls into the three conditions described above (they existed above or below the ground plane, were on a moving object, or not correctly matched), they can be identified by looking at the difference in feature distances between frames as described above.

Figure 8 demonstrates a case when a feature was not tracked correctly between two consecutive frames. A close-up of the five features being tracked is shown in figure 8C and figure 8D. In figure 8B, all features were successfully tracked in frame 2 except for feature 1. Feature 1 was determined to be non-valid by performing the distance relationship analysis between features in each of the two frames.

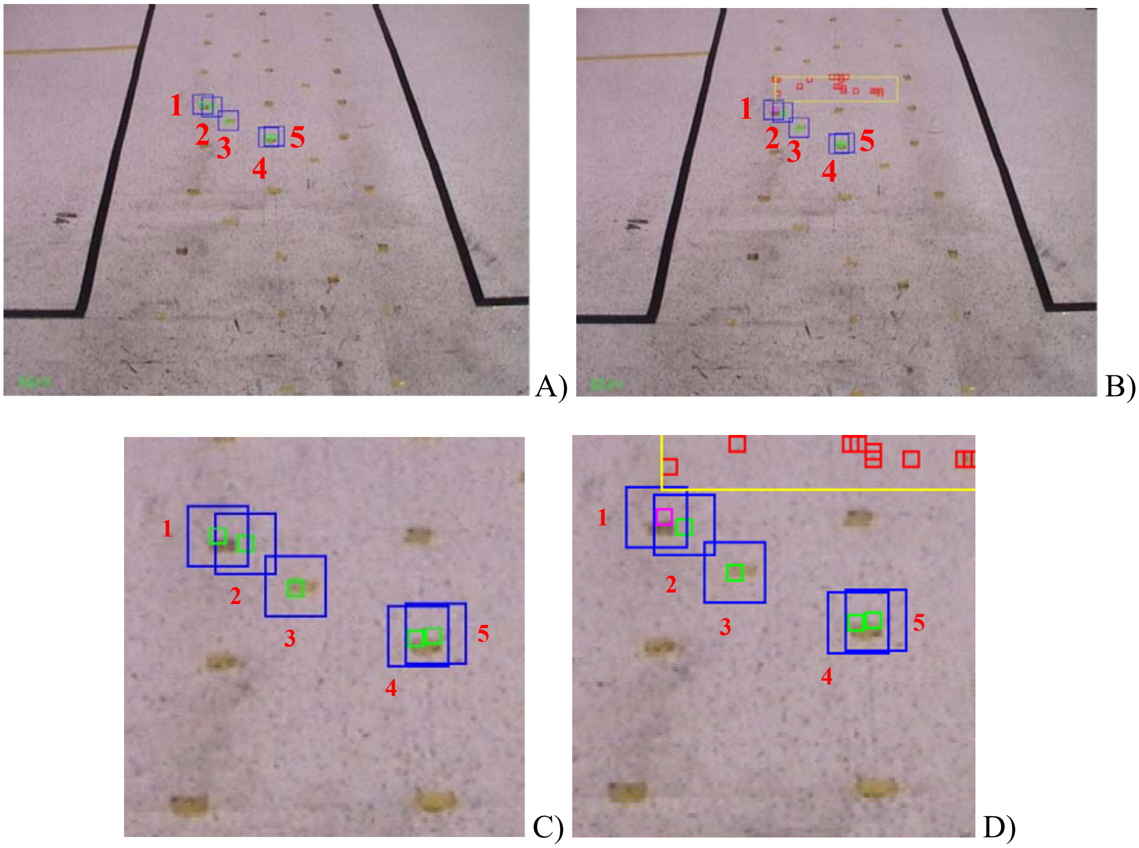


Figure 8. Mistracked feature. A) First frame. B) Second frame. C) Close-up of features from first frame. D) Close-up of features from second frame.

Table 4 and table 5 show the complete distance measurements between features from frame 1 in figure 8A and frame 2 in figure 8B. The difference between the distance measurements from frame 1 to frame 2 are calculated and shown in table 6.

Table 4. Distances between features in frame 1 shown in figure 8A in cm.

Feature	1	2	3	4	5
1	${}^1d_{1,1} = 0.00$	${}^1d_{1,2} = 2.34$	${}^1d_{1,3} = 9.32$	${}^1d_{1,4} = 19.71$	${}^1d_{1,5} = 20.42$
2	${}^1d_{2,1} = 2.34$	${}^1d_{2,2} = 0.00$	${}^1d_{2,3} = 7.21$	${}^1d_{2,4} = 17.48$	${}^1d_{2,5} = 18.14$
3	${}^1d_{3,1} = 9.32$	${}^1d_{3,2} = 7.21$	${}^1d_{3,3} = 0.00$	${}^1d_{3,4} = 10.54$	${}^1d_{3,5} = 11.35$
4	${}^1d_{4,1} = 19.71$	${}^1d_{4,2} = 17.48$	${}^1d_{4,3} = 10.54$	${}^1d_{4,4} = 0.00$	${}^1d_{4,5} = 1.24$
5	${}^1d_{5,1} = 20.42$	${}^1d_{5,2} = 18.14$	${}^1d_{5,3} = 11.35$	${}^1d_{5,4} = 1.24$	${}^1d_{5,5} = 0.00$

Table 5. Distances between features in frame 2 shown in figure 8B in cm.

Feature	1	2	3	4	5
1	$^{2}d_{1,1} = 0.00$	$^{2}d_{1,2} = 2.08$	$^{2}d_{1,3} = 9.17$	$^{2}d_{1,4} = 19.43$	$^{2}d_{1,5} = 20.09$
2	$^{2}d_{2,1} = 2.08$	$^{2}d_{2,2} = 0.00$	$^{2}d_{2,3} = 7.14$	$^{2}d_{2,4} = 17.32$	$^{2}d_{2,5} = 17.98$
3	$^{2}d_{3,1} = 9.17$	$^{2}d_{3,2} = 7.14$	$^{2}d_{3,3} = 0.00$	$^{2}d_{3,4} = 10.46$	$^{2}d_{3,5} = 11.28$
4	$^{2}d_{4,1} = 19.43$	$^{2}d_{4,2} = 17.32$	$^{2}d_{4,3} = 10.46$	$^{2}d_{4,4} = 0.00$	$^{2}d_{4,5} = 1.22$
5	$^{2}d_{5,1} = 20.09$	$^{2}d_{5,2} = 17.98$	$^{2}d_{5,3} = 11.28$	$^{2}d_{5,4} = 1.22$	$^{2}d_{5,5} = 0.00$

Table 6. Difference in feature distance in frame 1 (tab. 4) and frame 2 (tab. 5) in cm.

Feature	1	2	3	4	5
1	0.00	**0.25**	0.15	**0.28**	**0.33**
2	**0.25**	0.00	0.08	0.15	0.15
3	0.15	0.08	0.00	0.08	0.08
4	**0.28**	0.15	0.08	0.00	0.03
5	**0.33**	0.15	0.08	0.03	0.00

As seen in table 6, many of the changes of distances measured from feature 1 (the highlighted row) changed significantly more than those for the other four features. A threshold value of 0.25 cm was set to pick out change in feature distances that indicated significant change. Any change of distance $\geq$ 0.25 cm is labeled as a significant change. These are marked in bold in table 6. If the number of high changes is $\geq$ (n/2), where n is the number of features, then the feature is labeled as invalid. For the case in table 6, $n/2 = 5/2 = 2.5$. The number of significant changes for feature 1 is 3. Since $3 > 2.5$, the feature is classified as invalid. When a feature is classified as invalid, it is removed from the tracking list and not used in the visual odometer calculations, and a new search is run on the current image to find a new feature to replace it (seen by the new rectangular KLT search box in figure 8B).

All features whose search area has moved outside the image are also removed from the tracking list. A new feature search is calculated to find new features to replace them. Figure 9 demonstrates a three-frame sequence in which a feature is lost, replaced, and tracked.

4.4 Determination of Position and Orientation Change

Based on the movement of the tracked features, the change in vehicle position and orientation is calculated between frames. Given two features tracked successfully from frame 1 to frame 2, the translation of the vehicle and the rotation of the vehicle about the world z-axis are found. The first calculation performed is the rotation of the features in frame 2 relative to the features in frame 1. Since the features are fixed in the world coordinate system, this is equivalent to finding the rotation for the world system in frame 2 relative to the world system in frame 1. Figure 10 shows the calculations of two consecutive frames using the coordinates of two features. $^{1}F_1$ and $^{1}F_2$ represents the coordinates of features 1 and 2 from frame 1 in the vehicle coordinate system. $^{2}F_1$ and $^{2}F_2$ represents the coordinates of features 1 and 2 from frame 2 in the vehicle coordinate system.

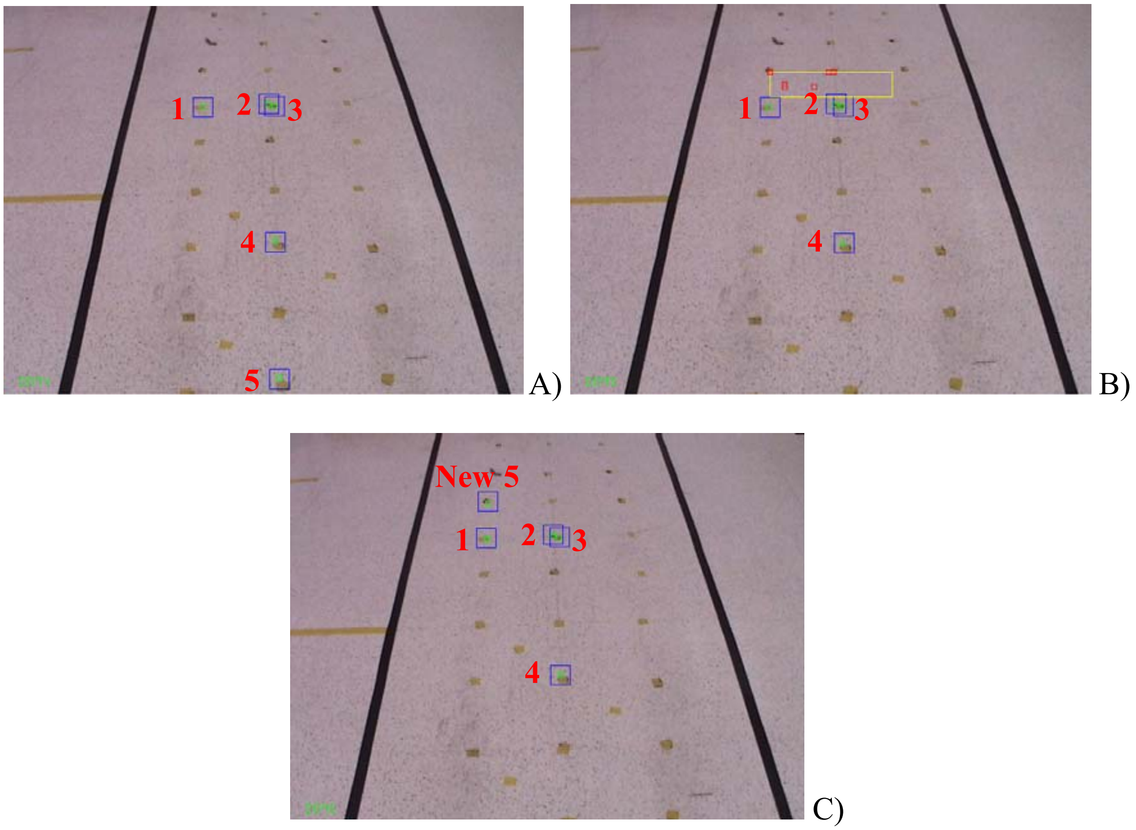


Figure 9. Feature leaving image and replaced. A) Frame 1: Five features tracked. B) Frame 2: feature 5 moved off image and replacement feature search. C) Fame 3: New feature 5 tracked.

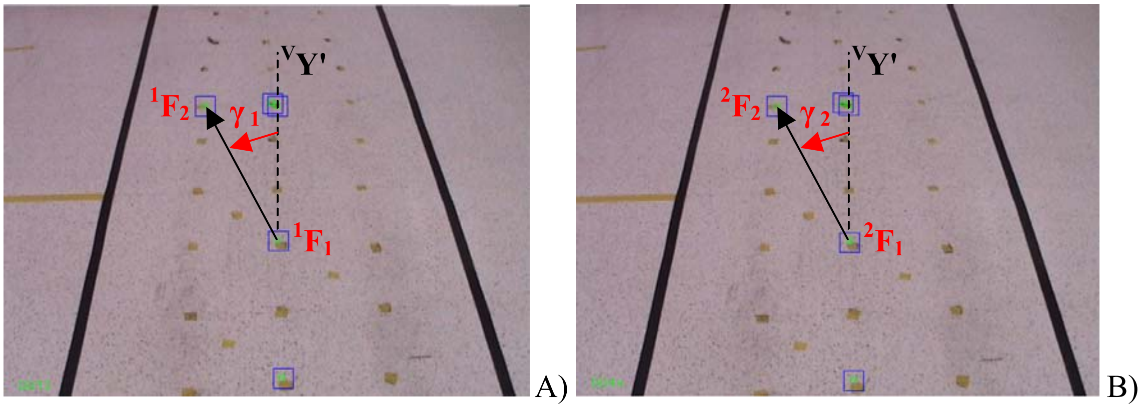


Figure 10. Calculation of rotation between two consecutive frames. A) Frame 1. B) Frame 2.

${}^{V}Y'$ represents a line drawn through feature 1 parallel to the y-axis of the vehicle coordinate system. γ_1 and γ_2 represent the angles between ${}^{V}Y'$ and a ray drawn from feature 1 to feature 2 in the two frames. The rotation of the world system in frame 2 relative to the world system in frame 1, γ, is calculated:

$$\gamma = \gamma_2 - \gamma_1$$

Given the vehicle system coordinates of feature 1 in frame 1, $^{1}F_{1} = (^{1}F_{1,x}, {}^{1}F_{1,y}, {}^{1}F_{1,z})$, and feature 1 in frame 2, $^{2}F_{1} = (^{2}F_{1,x}, {}^{2}F_{1,y}, {}^{2}F_{1,z})$, the translation of the world system in frame 2 relative to the world system frame 1 is calculated next. Let $^{1}T_{2}$ represent the transformation matrix relating the world system in frame 2 to the world system frame 1. The relationship between the coordinates of feature 1 in frame 1 and frame 2 is given by:

$$^{1}F_{1} = {}^{1}T_{2} \cdot {}^{2}F_{1}$$

Expanding this matrix and assuming vehicle rotation only occurs about its z-axis (which is sufficient on a level floor),

$$\begin{bmatrix} ^{1}F_{1,x} \\ ^{1}F_{1,y} \\ ^{1}F_{1,z} \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 & T_{x} \\ \sin(\gamma) & \cos(\gamma) & 0 & T_{y} \\ 0 & 0 & 1 & T_{z} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} ^{2}F_{1,x} \\ ^{2}F_{1,y} \\ ^{2}F_{1,z} \\ 1 \end{bmatrix}$$

With the assumption that all points are on the ground ($^{1}F_{1,z}, {}^{2}F_{1,z} = 0$), the translation of the world frame in the vehicle coordinate system in the x, y, and z directions, T_x, T_y and T_z, are solved:

$$T_{x} = {}^{1}F_{1,x} - \left(\cos(\gamma) \cdot {}^{2}F_{1,x} - \sin(\gamma) \cdot {}^{2}F_{1,y}\right)$$

$$T_{y} = {}^{1}F_{1,y} - \left(\sin(\gamma) \cdot {}^{2}F_{1,x} + \cos(\gamma) \cdot {}^{2}F_{1,y}\right)$$

$$T_{z} = 0$$

To find the best estimates for the rotation, γ, and translation, T_x, T_y, and T_z, between frames, these values are calculated for all combinations of the valid tracked features. For each set of rotation and translation values calculated, $^{1}T_{2}$ is composed and multiplied with the coordinates of each of the features, $^{2}F_{i}$, to calculate $^{1}F'_{i}$.

$$^{1}F_{i}' = {}^{1}T_{2} \cdot {}^{2}F_{i}$$

An error for the estimated set of rotation and translation values is

$$error = \sum_{i=1}^{n} (^{1}F_{i}' - {}^{1}F_{i} \ )^{2}$$

where n is the number of valid tracked features available. The set of rotation and translation values yielding the least error is used as the best estimate for γ, T_x, T_y and T_z. As demonstrated before, this set of values is used to develop $^{1}T_{2}$, which represents the transformation matrix relating the world system in frame 2 to the world system frame 1:

$$
{}^{1}T_{2} = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 & T_x \\ \sin(\gamma) & \cos(\gamma) & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

If ${}^{W}T_{V,old}$ represents the last calculated transformation matrix relating the vehicle coordinate system to the world system, a new ${}^{W}T_{V}$ is calculated using the latest ${}^{1}T_{2}$:

$$
{}^{W}T_{V} = {}^{W}T_{V,old} \cdot {}^{1}T_{2}
$$

The final location of the vehicle in the world coordinate system is ${}^{W}P_{V} = ({}^{W}x_{V}, {}^{W}y_{V}, {}^{W}z_{V})$. The final orientation of the vehicle about the world z-axis is ${}^{W}\gamma_{V}$. Both are extracted from ${}^{W}T_{V}$:

$$
{}^{W}T_{V} = \begin{bmatrix} \cos({}^{W}\gamma_{V}) & -\sin({}^{W}\gamma_{V}) & 0 & {}^{W}x_{V} \\ \sin({}^{W}\gamma_{V}) & \cos({}^{W}\gamma_{V}) & 0 & {}^{W}y_{V} \\ 0 & 0 & 1 & {}^{W}z_{V} \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$
{}^{W}\gamma_{V} = a\tan 2(\sin({}^{W}\gamma_{V}), \cos({}^{W}\gamma_{V}))
$$

${}^{V}T_{W}$, representing the transformation matrix relating the world system to the vehicle coordinate system, can also be calculated from ${}^{W}T_{V}$ (Crane and Duffy, 1998):

$$
{}^{W}T_{V} = \begin{bmatrix} \cos({}^{W}\gamma_{V}) & -\sin({}^{W}\gamma_{V}) & 0 & {}^{W}x_{V} \\ \sin({}^{W}\gamma_{V}) & \cos({}^{W}\gamma_{V}) & 0 & {}^{W}y_{V} \\ 0 & 0 & 1 & {}^{W}z_{V} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} & R_W & & T_W \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$
{}^{V}T_{W} = \begin{bmatrix} & R_W{}^{T} & & -R_W{}^{T} \cdot T_W \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

5. EXPERIMENTAL SETUP

Three experiments were run to test the visual odometer accuracy and its utilization on various level surfaces: translation test, rotation test and verification tests on various surfaces.

5.1 Translation Test

Vehicle translation estimation is used during intersection navigation to determine when the vehicle has traveled far enough into the intersection to start its turn. The position must be accurate to ensure proper alignment of the vehicle with the next path after turning. Since the vehicle is commanded to drive a specified distance forward along a straight line, the odometer translation test is run on a straight line along the vehicle y-axis. Small squares of tape were placed 15.24 cm apart along the vehicle x- and y-coordinate plane on the ground along the path to provide adequate features for the visual odometer to track (fig. 11). Figure 12 shows the tape marks as seen from the camera image.

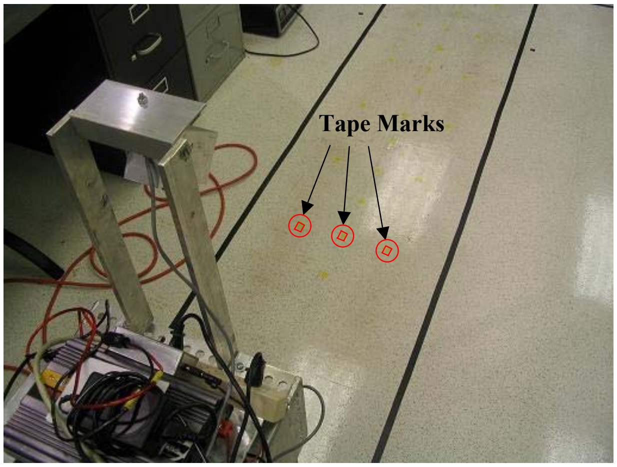


Figure 11. Yellow tape marks used for features during odometer test.

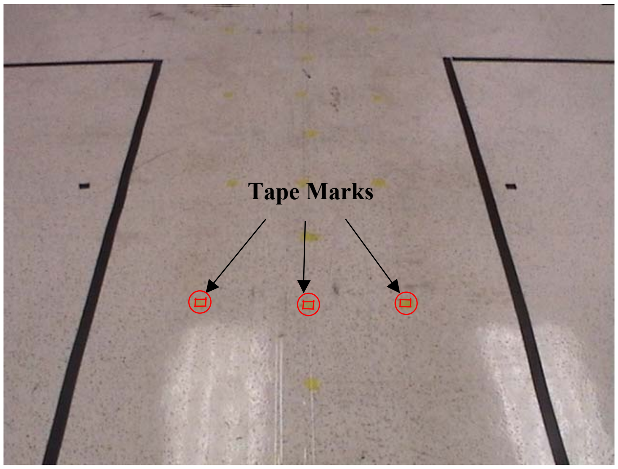


Figure 12. Camera view of path with tape marks.

For each test, the vehicle was lined up with a reference line and set up at a starting position, and the world coordinate system was initialized as the starting vehicle coordinate system. The vehicle was instructed to drive straight for a specified distance along the world y-axis direction and stop when the visual odometer read a distance greater than the specified distance. Distances of 15.24 cm to 304.8 cm in 15.24 cm increments were tested. The largest distance, 304.8 cm, was selected as the maximum range a vehicle would have to travel to reach the end of a 152.4 cm wide intersection if the visual odometer was started when the beginning of the intersection was 114.3 cm from the vehicle. Three runs were performed for each distance. The vehicle was driven forward at approximately 11 cm per s. The x and y translations from the visual odometer were recorded at the end of each run. The origin of the vehicle coordinate system at the end was marked and its translation in the x- and y-directions was measured relative to the world coordinate system at the start. Measurements were made to the nearest 0.3 cm using a ruler. The measured vehicle translation was compared to the odometer estimation and an error was calculated.

5.2 Rotation Test

A similar test was conducted for rotation. Instead of tape marks on the floor, paper marks were used. The marks were placed in a circle approximately 150 cm in front of the vehicle so they only appear in the portion of the image where new features are searched for by the visual odometer (fig. 13).

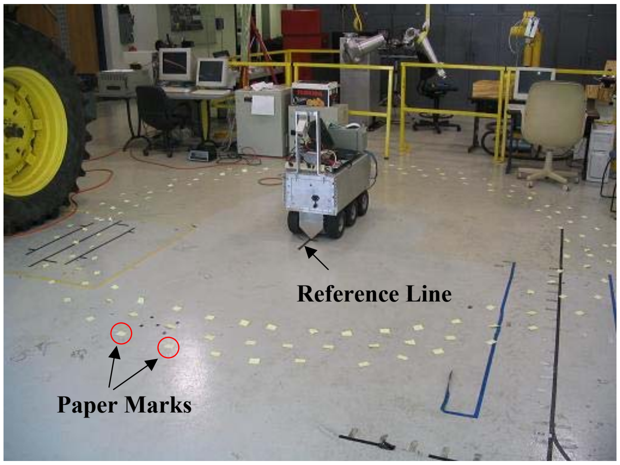


Figure 13. Experimental setup for visual odometer rotation test.

For each test, the vehicle was lined up with a reference line, as shown in figure 13, and set up at a starting position. The world coordinate system was initialized as the starting vehicle coordinate system. The vehicle was instructed to rotate a specified angle clockwise about the vehicle z-axis and stop when the visual odometer read a distance equal to or greater than that angle. Angles of 45° to 180° in 45° increments were used. Three runs were performed for each angle. The vehicle was rotated clockwise at approximately 3° per second.

The angle of rotation from the visual odometer was recorded at the end of each run. The front and back of the vehicle were marked on the ground, and a line was drawn between the two to measure vehicle rotation relative to the starting reference line. The angle was measured to the nearest degree. The measured vehicle rotation was compared to the odometer estimation and an error was calculated.

5.3 Verification Tests on Various Surfaces

To prove that the visual odometer can work in a greenhouse environment, tests were also run on various surfaces that may be found in a greenhouse. These included concrete, sand, and gravel.

All three tests were performed outdoors. The experimental setups for each surface and the view of the surface as seen by the camera are shown in figure 14, figure 15, and figure 16. Sand and gravel were placed along the driving path with a width large enough to cover the area of the image where feature tracking occurs. The material was placed on a tarp for ease of cleanup and did not affect results.

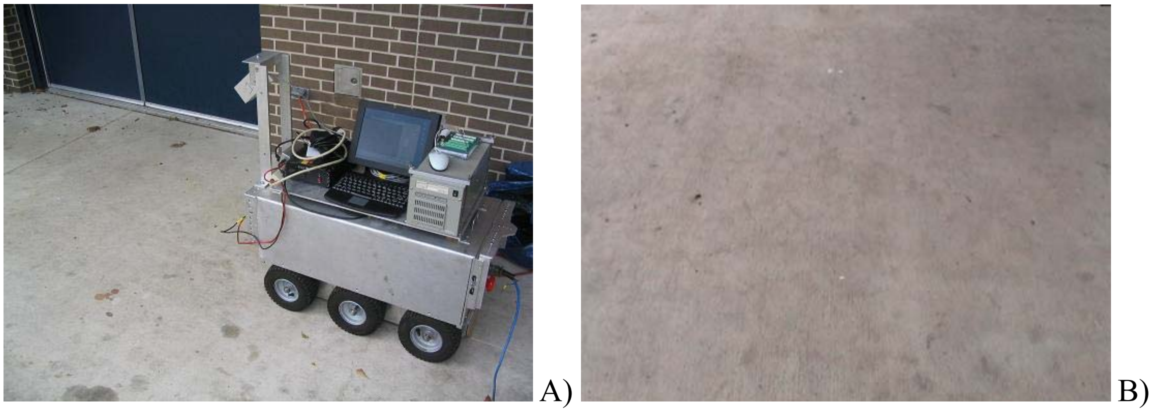


Figure 14. Concrete test setup. A) Concrete surface. B) Camera view.

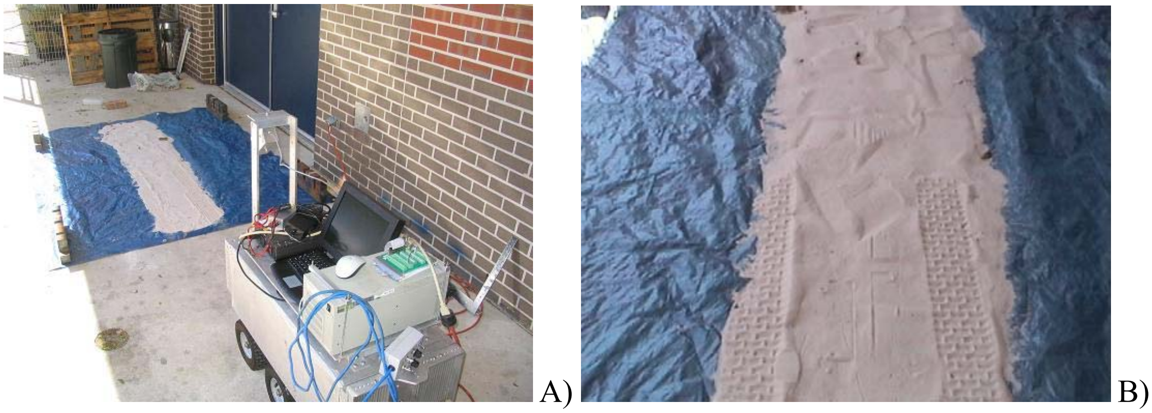


Figure 15. Sand test setup. A) Sand surface. B) Camera view.

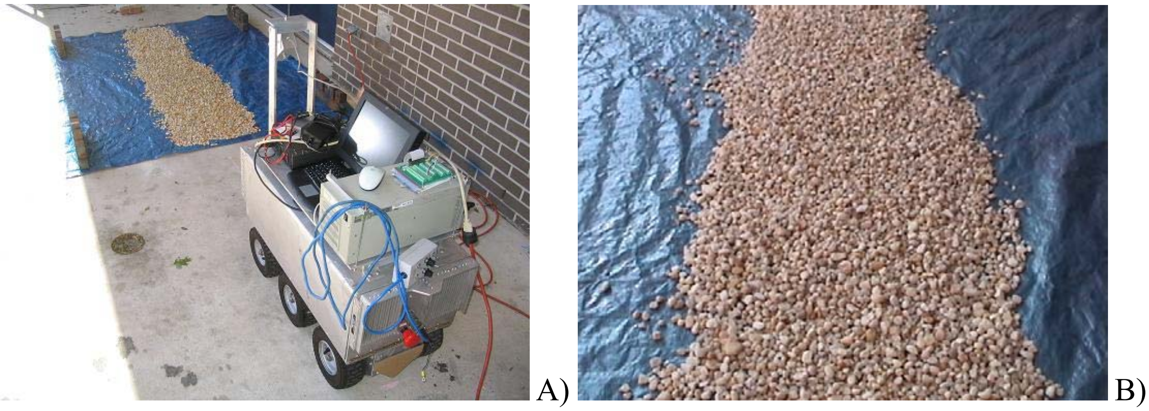


Figure 16. Gravel test setup. A) Gravel surface. B) Camera view.

For each test, the vehicle was lined up with a reference line and set up at a starting position, and the world coordinate system was initialized as the starting vehicle coordinate system. The vehicle was instructed to drive straight for a 152.4 cm distance along the world y-axis direction and stop when the visual odometer read a distance greater than that distance. Three runs were made for each surface. The vehicle was driven forward at approximately 11 cm per second, as in the earlier translation test.

The x and y translations from the visual odometer were recorded at the end of each run. The origin of the vehicle coordinate system at the end was marked, and its translation in the x- and y-directions was measured relative to the vehicle coordinate system at the start. Measurements were made to the nearest 0.3 cm using a ruler. A string was laid down the center of the path along the initial vehicle y-axis at the end of each run as a reference to measure translation. The measured vehicle translation was compared to the odometer estimation and an error was calculated.

6. RESULTS

Results were obtained for the translation tests, rotation tests, and verification tests on various surfaces. The visual odometer ran at an average of 10 Hz during experimentation.

6.1 Translation Test

Results obtained from the three translation test runs are shown in table 7, table 8, and table 9.

Table 7. Translation test Run 1.

Command Distance [cm]	Odometer x_v [cm]	Odometer y_v [cm]	Measured x_v [cm]	Measured y_v [cm]	Error [cm]
30.48	4.22	32.44	-0.64	33.02	4.88
60.96	0.66	62.89	0.00	65.10	2.29
91.44	-3.28	92.61	0.33	95.25	4.47
121.92	-10.92	123.95	0.33	124.46	11.25
152.40	-3.23	154.25	0.97	156.85	4.93
182.88	1.24	184.73	-1.60	185.12	2.87
213.36	5.54	215.62	0.00	214.63	5.61
243.84	-6.27	246.10	-4.45	243.21	3.43
274.32	12.01	276.15	-3.81	276.86	15.85
304.80	-29.49	307.14	-1.60	306.07	27.91

Table 8. Translation test Run 2.

Command Distance [cm]	Odometer x_v [cm]	Odometer y_v [cm]	Measured x_v [cm]	Measured y_v [cm]	Error [cm]
30.48	1.57	32.49	-0.33	33.66	2.24
60.96	10.41	64.29	0.33	67.31	10.52
91.44	0.74	93.68	0.64	96.52	2.84
121.92	3.58	123.49	0.00	125.73	4.22
152.40	-4.93	154.33	0.00	154.31	4.93
182.88	9.37	184.28	-1.27	184.15	10.64
213.36	-1.04	214.83	-1.91	211.46	3.48
243.84	-0.30	245.03	-3.51	244.48	3.25
274.32	-6.71	275.74	-0.64	272.11	7.09
304.80	0.18	305.94	-2.54	300.69	5.92

Table 9. Translation test Run 3.

Command Distance [cm]	Odometer x_v [cm]	Odometer y_v [cm]	Measured x_v [cm]	Measured y_v [cm]	Error [cm]
30.48	7.26	33.20	0.00	34.93	7.47
60.96	-7.21	62.97	0.00	64.77	7.44
91.44	0.84	92.74	0.33	95.25	2.57
121.92	-1.22	123.60	0.33	126.37	3.15
152.40	-7.80	154.69	0.33	153.04	8.28
182.88	3.51	185.22	0.97	182.25	3.91
213.36	-5.08	215.19	0.97	215.60	6.05
243.84	-9.50	246.35	-3.51	246.08	6.02
274.32	7.16	275.97	-2.54	274.65	9.78
304.80	-0.76	307.06	-3.18	304.80	3.30

The errors listed in the tables represent the difference in distance between the measured vehicle position and the position estimated by the visual odometer. Table 10 shows the average, minimum, maximum errors and standard deviations from the three runs. The general error range remained consistent throughout the range of distances tested. These odometer errors are the result of the camera model errors relating pixels to points in the vehicle coordinate system as discussed in the previous experiment. Even though a ground feature may have been tracked successfully from the top of the image to the bottom, its perceived coordinates by the camera model could be off by several centimeters by the time it reaches the bottom of the image. Depending on where in the image a feature was originally found and how far it was tracked down the path determined the error contribution it made.

Table 10. Average error over the three translation test runs.

Command Distance [cm]	Average Error [cm]	Minimum Error [cm]	Maximum Error [cm]	Standard Deviation [cm]
30.48	4.85	2.24	7.47	2.62
60.96	6.76	2.29	10.52	4.16
91.44	3.30	2.57	4.47	1.02
121.92	6.22	3.15	11.25	4.40
152.40	6.05	4.93	8.28	1.94
182.88	5.82	2.87	10.64	4.22
213.36	5.05	3.48	6.05	1.37
243.84	4.22	3.25	6.02	1.55
274.32	10.90	7.09	15.85	4.48
304.80	12.40	3.30	27.91	13.52

Several runs gave very high errors above 10 cm, with the largest occurring during Run 1 at the 304.80 cm distance with a 27.91 cm error. The main contributions to these errors were poor estimation of the vehicle translation along its x-axis. Since the error remains relatively low out to 243 cm, it is possible that vision distortion on the far horizon may contribute to the significant jump in error beyond 243 cm. However, it is clear that at moderate command distances, the average error is less than 7 cm with relatively small variation. In practice, this accuracy has proven adequate for navigating corners in greenhouse simulation tests.

6.2 Rotation Test

Results obtained from the three rotation test runs are shown in table 11, table 12, and table 13.

Table 11. Rotation test Run 1.

Command Angle [°]	Odometer Angle [°]	Measured Angle [°]	Error [°]
45	45	45	0
90	90	90	0
135	136	145	9
180	182	190	8

Table 12. Rotation test Run 2.

Command Angle [°]	Odometer Angle [°]	Measured Angle [°]	Error [°]
45	46	48	2
90	90	88	2
135	136	150	14
180	182	166	16

Table 13. Rotation test Run 3.

Command Angle [°]	Odometer Angle [°]	Measured Angle [°]	Error [°]
45	46	44	2
90	90	91	1
135	135	155	20
180	181	182	1

The errors listed in the tables represent the difference in rotation between the measured vehicle orientation and the orientation estimated by the visual odometer. Table 14 shows the average, minimum, maximum errors and standard deviations from the three runs. There is a general trend that the farther the command distance, the larger the error. The maximum error during experimentation was 20°, which occurred during a 135° turn. It would be interesting to evaluate whether the turning rate and dynamics of the vehicle influenced the error at larger command angles. The accuracy of the 45° and 90° turn angles are extremely good and suggest that it may be possible to improve the performance at larger turn angles.

Table 14. Average error over the three rotation test runs.

Command Angle [°]	Average Error [°]	Minimum Error [°]	Maximum Error [°]	Standard Deviation [°]
45	1	0	2	1
90	1	0	2	1
135	14	9	20	5
180	8	1	16	8

6.3 Verification Tests on Various Surfaces

The results obtained for a commanded vehicle translation of 152.4 cm on concrete, sand, and gravel are shown in table 15, table 16, and table 17.

Table 15. Translation test for concrete.

Run	Odometer x_v [cm]	Odometer y_v [cm]	Measured x_v [cm]	Measured y_v [cm]	Error [cm]
1	-19.84	154.69	-0.33	150.50	19.96
2	-10.44	154.08	-0.33	150.50	10.74
3	-6.30	154.53	-0.33	152.10	6.45

Table 16. Translation test for sand.

Run	Odometer x_v [cm]	Odometer y_v [cm]	Measured x_v [cm]	Measured y_v [cm]	Error [cm]
1	-5.97	154.03	-3.51	150.83	4.06
2	-7.32	153.97	-2.54	148.59	7.19
3	-7.19	154.23	-4.45	152.40	2.97

Table 17. Translation test for gravel.

Run	Odometer x_v [cm]	Odometer y_v [cm]	Measured x_v [cm]	Measured y_v [cm]	Error [cm]
1	-11.30	152.83	-5.08	146.05	9.19
2	-7.09	154.36	-5.08	147.32	7.32
3	-6.78	154.00	-3.81	144.78	9.68

The errors listed in the tables represent the difference in distance between the measured vehicle position and the position estimated by the visual odometer. Table 18 shows the average, minimum, maximum errors and standard deviations from the three runs for each surface. The errors acquired from the testing on the lab floor at 152.4 cm as reported in table 10 are also shown in table 18 for comparison. The error results for sand and gravel compare favorably with that found on the lab floor. In these cases, the availability of distinct surface features provides a good basis for visual odometry. However, the concrete surface demonstrates significantly higher errors, which may be attributed to the less distinct image features.

Table 18. Average error over the three test runs for concrete, sand, and gravel.

Surface	Average Error [cm]	Minimum Error [cm]	Maximum Error [cm]	Standard Deviation [cm]
Concrete	12.40	6.45	19.96	6.91
Sand	4.75	2.97	7.19	2.19
Gravel	8.74	7.32	9.68	1.25
Lab	6.05	4.93	8.28	1.94

7. CONCLUSION

The visual odometer gave accurate estimation of vehicle translation and rotation. Translation tests of the odometer in a lab environment gave an average error of 4.85 cm for a 30.5 cm forward translation and 12.4 cm average error for a 305 cm translation. This increased error may be due to far horizon image distortion. Rotation tests of the odometer in a lab environment gave an average error of 1° for a 45° rotation about the vehicle z-axis and 8° error for a 180° rotation. These were within an acceptable range, as demonstrated in the intersection navigation tests for turning in Younse (2005), allowing the visual odometer to guide the vehicle to a position in the intersection suitable for the turn. Any positional errors in the location of the vehicle turning center were dealt with successfully by the path following algorithm after the turn, which corrected any offset error and guided the vehicle towards the center of the second path. Finally, tests completed on concrete, sand, and gravel demonstrated adaptability of the odometer on different ground surfaces that are common in greenhouses.

The visual odometer can be improved by averaging vehicle translation and rotation over multiple points from several frames, as opposed to using just two features per frame. Utilization of techniques researched by Nistér et al. (2004) can allow the visual odometer to track features that are not restricted to the ground plane.

8. ACKNOWLEDGEMENTS

This research was conducted at the University of Florida, Institute of Food and Agricultural Sciences, through funding provided by the National Foliage Foundation.

9. REFERENCES

Bouguet, J. 2004. Camera Calibration Toolbox for Matlab. Available at: http://www.vision.caltech.edu/bouguetj/calib_doc, Accessed Aug. 2004.

Chojnacki, W., M. Brooks, A. Hengel, and D. Gawley. 2003. Revisiting Hartley's Normalized Eight-Point Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 9.

Crane, C., and J. Duffy. 1998. Kinematic Analysis of Robot Manipulators. New York, New York: Cambridge University Press, 9.

Csetverikov, D. 2004. Basic Algorithms for Digital Image Analysis: a course. Budapest, Hungary: Institute of Informatics Eötvös Loránd University. Aug. 2004. Available at: http://visual.ipan.sztaki.hu, accessed Aug. 2004.

Have, H., S. Blackmore, B. Keller, H. Fountas, S. Nielsen, and F. Theilby. 2002. Autonomous Weeders for Christmas Tree Plantations- A Feasibility Study. *Pesticide Research 59*. Denmark: Danish Environmental Protection Agency, Danish Ministry of the Environment.

Hellström, T. 2002. Autonomous Navigation for Forest Machines: A Pre-Study. Umeå, Sweden: Department of Computing Science, Umeå University.

Nistér, D., O. Naroditsky, and J. Bergen. 2004. Visual Odometry. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, 652-659.

Pollefeys, M. 2004. Visual 3D Modeling from Images. Tutorial Notes. Chapel Hill, NC: University of North Carolina.

Singh, S. 2004. Autonomous Robotic Vehicle for Greenhouse Spraying. ME thesis. Gainesville, Fla.: University of Florida, Department of Agricultural and Biological Engineering.

Singh, S., and V. Subramanian. 2004. Autonomous Greenhouse Sprayer Vehicle using Machine Vision and Ladar for Steering Control. *Proceedings of the 2004 ASAE Automation Technology for Off-road Equipment*, St. Joseph, Mich., 79-90.

Younse, P. 2005. Intersection Detection and Navigation for an Autonomous Greenhouse Sprayer using Machine Vision. ME thesis. Gainesville, Fla.: University of Florida, Department of Agricultural and Biological Engineering.