Vision-only Outdoor Localization of Two Wheel Tractor for Autonomous Operation in Agricultural Fields

Lasitha Piyathilaka∗, Rohan Munasinghe†,
∗† Department of Electronics and Telecommunication Engineering
University of Moratuwa,
Email:* lasitha@ieee.org ,†rohan@ent.mrt.ac.lk

Abstract—This paper outlines the problem of outdoor au-
tonomous robot localization for agricultural operations. Vision
based outdoor localization system was developed entirely of off-
the-shelf components that can accurately guide a field robot in
small agricultural fields. Visual odometry using downward faced
camera is proposed as a high resolution relative localization
system. A stereo vision based range measurement system is also
developed and field tested as an absolute localization system that
can bound the incremental error caused by the visual odometry
system. The extended Kalman filter with measurement gaiting
is implemented using both visual odometry and stereo range
measurement data. Final Results demonstrated proposed system
can be used for outdoor localization of two wheel tractor at a very
low cost with acceptable accuracy in small agricultural fields.

I. INTRODUCTION

Introducing autonomous systems for agriculture is the new
trend in automation where numerous research are done. It is
also called as "Precision Agriculture" which has a high com-
mercial potential. By introducing automation to the agriculture
number of benefits can be yielded. Automation in agriculture
reduces arduousness of the driving task, allow driver to devote
his time for monitoring and control the work ,exact placement
of field inputs (seeds, fertilizers, pesticides, etc.) can be
achieved which can reduce the cost. Many types of agricultural
vehicles are commercially available. Most common is four
wheel tractor which is used for various applications such as
tilling ,ploughing and seeds/fertilizer placement. But in south
and south east asia power tiller or two wheel tractor is used
heavily for farming because of its low operating cost. Also it
can be utilized for field operations in small agricultural fields
easily. But still operating these machineries requires extensive
man power. In this context we believe by developing a re-
motely operable and autonomous farm vehicle latest technolo-
gies can be introduce to the traditional farming community.
Automatic guidance of agricultural vehicles requiers a very
accurate localization system. Kinematic GPS sensors provide
real-time position information with a centimeter accuracy and
used in many commercial autonomous agricultural vehicles.
[1]. But these available systems are expensive and even cost
more than the price of a power-tiller. On the other-hand unlike
autonomous military or mining vehicles autonomous agricul-
tural vehicles will become marketable only if there are really
feasible in the cost benefit analysis. Therefore autonomous
navigation system should cost low while maintaining required
accuracy. In this research we focused of developing a vision
based localization system for the two wheel tractor that enables
autonomous operation. The proposed system provides the
reasonable accuracy required while maintaining the cost of
the system as low as possible.

Use of vision based sensors for in agriculture is not new. The
most commonly used machine vision methods are detecting
crops patterns, soil tillage, and the edges along harvested
crops. [2]- [3]. But these systems only works if the crop is
arranged in systematic manner and work for only specific

type of crops and environments. Wheel odometry and inertial
sensors can be used as low cost localization systems but
requires accurate kinematic model which is hard to achieve
when the wheel slip become significant. GPS and machine
vision is also used [2] in outdoor localization systems to
improve the accuracy but sensor integration can be expensive.
Proposed vision based localization system uses only off-the-
shelf web cameras as the localization system and use open
source image processing softwares. Visual odometry based
relative localization system provides high resolution incre-
mental update of position while stereo vision based absolute
localization system bounds the accumulated error caused by
the visual odometric system. Extended Kalman Filter (EKF)
with measurement gaiting is proposed as the sensor fusion
technique.

The rest of the paper is organized as follows. Section II de-
scribes the experimental setup. Section III describes the visual
odometry algorithm which is used as the relative localization
system .Sections IV describes absolute localization system
that use stereo vision for range measurement. It consists of
stereo depth measurement and a tracking algorithm that track
a color coded tag attached to the moving tractor . Section
V presents use of Extended Kalman Filter(EKF) as a sensor
fusion system that combined absolute range measurement from
stereo vision system and relative measurement from visual
odometry. Section VI details the experimental results. Sections
VII and VIII consist of discussion and conclusion.

II. EXPERIMENTAL SETUP

Proposed vision based outdoor localization system is con-
sisted of two main components, namely relative localization
with visual odometry and absolute localization using stereo vision. These systems were built using low cost web cameras as the only sensor and open source image processing softwares were used for algorithm development.

A. Visual Odometry Setup

The experimental setup is shown in fig.1 A web camera with 640X480 pixels resolution is fixed facing downward at front of the tractor from about 75 cm from the wheel base. This camera is connected to a laptop pc that runs visual odometry algorithm in C++ and OpenCv.

B. Stereo Vision Setup

Stereo camera system used in the experiment is consisted of two low cost web cameras having resolution of 640X480 pixels. They were fixed on a platform with baseline distance of 40 cm. Stereo range measurement will work only if the object lie on the common area of two cameras. Therefore when the target object (Robot) moves on the field, it should be tracked and stereo camera system should adjust accordingly. This was achieved by fixing two cameras on a platform that can be rotated by a servo motor. When the object is detected by cameras, it position is estimated. If the object starts to move servo motor starts to rotate according to array the position information received by the image processing software. Image processing was done on a Intel Core Duo computer using OpenCv image processing libraries. Servo controller was connected to the computer via a RS 232 serial link which sends position information to the servo. Figure 2 shows the rotatable stereo camera system mounted on a tripod.

III. VISUAL ODOMETRY-RELATIVE LOCALIZATION

One of the most challenging problems of field robots is self-localization. Relative localization systems involve incremental update of position while in motion and used when high accuracy and resolution is needed in short-run. Wheel odometry and use of Inertial Measurement Unit (IMU) are the mostly used relative localization systems. But they rely on the accuracy and integrity of the kinematic model and wheel odometry fails when wheel slip is significant. As an alternative visual odometry can be used which does not rely on a kinematic model of the vehicle. This involve use of one or more cameras to acquire images from the surrounding which are used to estimate relative change in pose of the robot. In our previous research [5] we studied visual odometry estimation on feature tracking. Therefore same system was used as the relative localization system. In this approach comparatively distinguishable features are identified from the downward faced camera and tracked across video frames in order to estimate the motion of the robot. This motion updates overtime and calculate the robot position in the world fixed coordinate system. Shi/Tomasi algorithm [10] is used for feature tracking and Lucas Kande algorithm [6] is used for optical flow field estimation. Following fig.3 shows optical flow estimated from two consecutive video frames captured by the downward faced camera attached to the moving tractor. Heading change of the robot was calculated by using optical flow patterns. In our previous study [5] we proposed optical flow field based heading angle calculation for a skid-steered vehicle. It is noted that optical flow field creates an Instantaneous Center of Rotation.
(ICR) when robot is rotating. If the ICR is estimated then the robot heading changes can be calculated. It is determined by finding the intersection point of each flow vectors bisecting line for consequent frames as shown in the Fig.4.

IV. STEREO VISION BASED -RELATIVE LOCALIZATION

Inherit problem of any odometric model is that incremental error accumulation. Therefore even a slight change of angle caused by measurement errors results considerable deviation to the final position of the robot. Because of these constrains, we look for some means to sense the robots pose absolutely, with reference to a fixed point by using stereo vision. Stereo vision is used over the years to extract depth measurement from two image pairs obtained by two cameras placed with some distance. As the range measurement from stereo camera system is intended to bound the incremental error caused by the visual odometry, ±1m accuracy is believed to be sufficient for stereo range measurement unit. Other advantage is same hardware used by the visual odometry system can be used by the stereo range measurement system without much added cost which reduce overall cost of the system. Stereo measurement system was capable of measuring range distance to objects less than 30 m with acceptable accuracy. For this experiment two low cost web cameras were used as stereo pairs and Intel’s OpenCv image processing library was used to track and calculate the distance to the object. Fundamental problem of range measurement using stereo pair is the object should lie on the common area of the two cameras in order to measure the disparity. This reduce the area that the object (robot) can lie. Therefore as the object move camera’s fixed coordinate system also should rotate with the object. This was achieved by fixing the stereo camera pair on a platform that is rotated by a servo motor. By assuming pin-hole camera model distance to a object can be calculating from the followig equation 1 where, B is the baseline distance of the two cameras, f is the focal length and \( x_l - x_r \) is the disparity of two corresponding point from right and left cameras.

\[
z = \frac{2Bf}{x_l - x_r} \tag{1}
\]

A. Object tracking and stereo image correspondence process

The major step in the stereo depth perception is the computation of correspondent between points in the two stereo image pair. This step is known as the correspondence process which involves matching feature points in two stereo images. In this experiment our scope was to measure the range distance to a moving object in order to localize it in the world coordinate frame. To reduce the complexity of the correspondence process colored spherical tag is attached to the robot. By tracking the position of the color tag position of the moving robot was tracked. Main steps of the correspondence matching and tracking algorithm is shown below. It was assumed that there aren’t any objects with the same color as color tag in the environment.

- Capture two images from the two web cameras.
- Use a colour filter to extract image points having the same color as the tag.
- Smooth the image in small windows to filter out noise.
- Find the point \( P_{max} \) with maximum intensity of the smoothed image on the right image.
- Extract a 20X20 pixel image template around the point \( P_{max} \).
- Match the template with left image. If a matching window is found its center position coordinates are extracted.
- Verify whether matched windows center is lied on the same scan line as the template image’s original position . (If the cameras are perfectly aligned, then correspondence points should lie on the same scan line). If the two point are not in the same scan line repeat the correspondence process from step 1.
- If the above verification is succeeded calculate the disparity of the corresponding points.

B. Camera Coordinate Transformation

Since the moving object’s position is measured with reference to the camera’s coordinate system it is required to transform the measured position to world fixed coordinate system. Suppose that the rotating stereo camera system is fixed on the origin of the world coordinate system (XYZ) and stereo camera’s orthogonal reference frames are \( X_c \) and \( Z_c \) respectively. If the stereo camera is rotated by an angle of \( \theta \) with reference to the world coordinate system in order to track the moving object P and measured position of the object P by the stereo camera system is \( (X'_c, Z'_c) \) then using the equation 2 it can be transformed to the world’s fixed coordinate system.

\[
\begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix} =
\begin{bmatrix}
\cos\theta & -\sin\theta & 0 \\
\sin\theta & \cos\theta & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
x'_c \\
y'_c \\
z'_c \\
\end{bmatrix}
\tag{2}
\]

V. SENSOR FUSION WITH EKF FOR LOCALIZATION

Kalman filter [9] has been frequently applied to the problem of robot localization. It works recursively, and so does not require a history of the robots previous states to be kept. In this experiment we fused absolute range measurements from the stereo measurement system with the high resolution dead reckoning visual odometry data using Extended Kalman Filter(
EKF. EKF has to be used due to the non linearity of the range measurements. This chapter describes formulation of EKF for the proposed vision based localization system.

### A. Process Model

If the tractor pose at time $k$ is represented by the state vector $q_k = [x_k, y_k, \theta_k]^T$, then the following set of equations can be used to model the kineticle of the tractor.

$$q_{k+1} = \begin{bmatrix} x_k + \Delta D_k \cos(\theta_k) \\ y_k + \Delta D_k \sin(\theta_k) \\ \theta_k + \Delta \theta_k \end{bmatrix} = f(q_k, u_k) + v_k \quad (3)$$

Here $v_k$ is a noise vector and $D_k$ is the distance traveled by the tractor between subsequent frames which is obtained by the visual odometry using the downward faced camera. The incremental orientation change $\Delta \theta_k$ is obtained also by the angle measurement technique described in the chapter 4. These measurements are used as the the control input vector $u_k = [\Delta D_k, \theta_k]^T$.

### B. Measurement Model

Absolute measurement information obtained from the stereo measurement system is modeled in the measurement model. The stereo measurement system located at the origin measures the angle between the moving robot and the ground (x axis) and the distance between the robot and the stereo system. These measurement at time $k+1$ is given by the following equation.

$$h_{k+1} = \begin{bmatrix} \theta \\ r \end{bmatrix} = \begin{bmatrix} atan(\frac{y_{k+1}}{x_{k+1}}) + v_1 \\ \sqrt{x_{k+1}^2 + y_{k+1}^2} + v_2 \end{bmatrix} \quad (4)$$

where $v_1$ and $v_2$ are the random variables which represent the process noise. Then the $m$ by $n$ jacobian matrix of partial derivatives $H$ can be obtained as follows

$$H_{k+1} = \frac{\partial h_i}{\partial q_i[k]} \quad (5)$$

### C. Measurement Gating

Measurement validation is required to filter erroneous measurement which can degrade robot’s state estimations. If the uncertainty of the current measurement is great it can be caused by a erroneous measurement and if the confident of the current measurement is great, gate should assign a higher criteria for the incoming measurement. The first step of the gating process is calculating the normalized innovation squared which is proposed in [7].

$$E_v(k) = v(k)S(k)^{-1}v(k)^T \quad (6)$$

Here $(k)$ is the innovation (i.e., the difference between the measurement and its expected value according to the model) and $S$ is the matrix $S_k = H_k P_k H_k^T + R_k$.The matrix has a chi-square distribution, so the gate’s bounding values can be read from the chi-square table, and a measurement will be discarded if its $E_v(k)$ value is outside the bounding values.

### VI. RESULTS

To analyse the performances of the proposed localization system experiments were performed by manually driving a two wheel tractor in an outdoor field on a path that is similar to maneuvers used for tilling in agricultural field preparation. The trajectory of the tractor is tracked by the stereo vision based absolute localization system and visual odometry based relative localization system. A red colored marker (red ball) was mounted on the tractor as the object that can be tracked by the stereo range measurement system. This colored marker is clearly visible to the stereo range measurement unit throughout the test field. When the tractor moves in the field stereo range measurement system tracks the red colored object across the field and calculate the range and the absolute angle to the tractor. A metal rod was attached to the two wheel tractor touching the ground that marks the path travelled by the tractor on the ground. This path is later measured and recorded to use as the ground truth information.

Results of the experiments by maneuvering the tractor manually in a 10mX10m is shown in the following figure. The accuracy and precision (error standard deviation) of the measurements are obtained by comparison with the ground.

---

**Fig. 5.** Field testing. Stereo range finder in the front and the tractor with colored marker in the back

**Fig. 6.** Extended Kalman Filter Localization - Figure shows ground truth of the trajectory the tractor travelled and the path estimated by the EKF localization system.
TABLE I
PERFORMANCES OF THE KALMAN FILTER BASED LOCALIZATION SYSTEM.

<table>
<thead>
<tr>
<th></th>
<th>Average Error</th>
<th>Maximum Error</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cartesian Error</td>
<td>53.84 cm</td>
<td>152.40 cm</td>
<td>35.59 cm</td>
</tr>
</tbody>
</table>

Fig. 7. Visual Odometry Localization - Figure compares the ground truth and the path estimated by visual odometry.

Fig. 8. Absolute localization using stereo range measurements - Stereo range finder tracked the tractor across the field and the absolute position is updated at a rate of 5Hz.

VII. DISCUSSION

The analysis of the experimental data of EKF based localization showed that the proposed system is capable of localizing a moving two wheel tractor in an out-door environment with acceptable accuracy. Standard deviation of the Cartesian error was 35.6 cm. Fig. 7 - 8 show that both visual odometry and stereo vision based localization systems fail to correctly localize the tractor when run alone. In Visual odometry, error gets accumulated and the resolution of the stereo vision based range measurement system deteriorates with the increase of the distance as shown in fig. 7 and fig. 8.

Therefore these localization systems can not cater autonomous navigation. EKF data fusion showed acceptable results that can be used for autonomous operation of the tractor. EKF has bound the cumulative error caused by visual odometry using the low resolution absolute stereo vision range measurements.

A measurement validation gate is implemented to augment
the Kalman filter’s weighting of the range measurements; this is meant to disallow occasional erroneous measurements. Figure 10 shows such gate validation process is needed in order to avoid degrade of the pose estimations through EKF. Cumulative Frequency Distribution of the Cartesian Error showed in the figure 11. 70 percent of the time cartesian error is less than 70cm which is acceptable for a low cost outdoor localization system.

VIII. CONCLUSION

This paper outlines the problem of outdoor autonomous robot localization in agricultural fields. A commonly used two wheel tractor is modified and developed as an autonomous vehicle test bed. The Kalman filtering methodology is used in conjunction with the chi-square gating technique for a robust localization using vision based systems. The test results showed that the position error is acceptable for autonomous operation, but requires additional investigation and vision algorithm development. Stereo vision based absolute localization system was developed with object tracking to bound the cumulative error caused in visual odometry. It is suggested the fusion of visual odometry data and absolute stereo range measurements using Extended Kalman Filter. Field test were done using a two wheel tractor in an out door field on a path that is similar to maneuvers used for tilling in agricultural field preparation. Test results proved EKF as a reliable method that can improve the performance of the proposed vision based localization system. The advantage of the developed system is, it can be instrumented using low cost off the shelves electronics components.

ACKNOWLEDGMENT

Authors acknowledge the support of the National Science Foundation (NSF) of Sri Lanka. (Grant RG/2007/E/05).

REFERENCES