Disparity as a Separate Measurement in Monocular SLAM

Talha Manzoor[1] and Abubakr Muhammad[2]

Abstract— **In this paper, we investigate the effects of including disparity as an explicit measurement (not just for initialization) in addition to the projective camera measurements, in conventional monocular/bearing-only SLAM in a 2-D world. We conduct an observability analysis for a 1.5-D scenario which theoretically shows that adding disparity measurements influences the observability rank condition, and so disparity can be regarded as a valid measurement. As disparity is dependent on the camera's location one time step earlier, we rederive the EKF equations for measurement models dependent on the previous state. It is shown that this modification results in an increase in the measurement noise covariance. We also discuss the implications of the white noise assumption in EKF being violated with the inclusion of disparity measurements. In the end, we compare the results of a simulated monocular SLAM scenario in a 2-D world with and without disparity measurements. Here we use the magnitude of the eigenvalues of the simulated state covariance matrix as a metric for observability [1]. Our results suggest that including disparity measurements in monocular SLAM improves observability both in the map and pose variables. Hence, this paper is an invitation to consider disparity measurements explicitly, in monocular SLAM.**

I. Introduction

Simultaneous Localization and Mapping with a single camera, also known as Monocular SLAM, is a well known problem in the robotics community and extensive research and development have been carried out in this area. Many monocular SLAM systems have been developed to date with the first being presented by Davison [2] who used the projection of observed features or landmarks on the image plane of the camera as measurements for the EKF. Since then, a lot of research has been done in single camera SLAM [3], [4], [5], and in many sub-areas. Another approach introduced by Paz et al [6] includes monocular information in a stereo hand held system for SLAM, using both projective and disparity measurements. In this paper, we introduce a similar idea and discuss the possibility of including stereo information in a monocular SLAM system. Thus, we consider both projective and disparity measurements. However, the stereo in this case is done in time. This type of stereo vision is commonly known as motion-stereo [7] or wide-baseline stereo [8].

One major concern while deploying a single camera as a sensor for any mapping problem is the inherent handicap in trying to reconstruct the environment from projected images. This problem of *observability* of monocular (or bearing-only) SLAM has been investigated extensively in the past. Bonnifait and Garcia [9] carry out an observability analysis of a planar robot trying to localise itself, given the bearing to 3 known beacons. They conclude that the system is observable except for the degenerate case when the robot is located on the circle defined by the 3 beacons and is stationary. In a similar paper, Calleja et al [10] have considered the observability of a planar robot, performing SLAM with a bearing-only sensor. Instead of expressing the SLAM system in its regular form, they do so in the error form presented in [11]. This allows them to approximate the system as piecewise constant, enabling them to use linear observability techniques as given in [12]. They prove that for absolute coordinates SLAM, the process is unobservable over 2 time segments when the robot is either stationary, or moving directly towards a landmark. Introducing a known landmark (beacon) renders the linearly approximated process fully observable. Tribou et al [13] present an observability analysis for a planar position based visual servoing system, also applicable to monocular SLAM. They focus on determining the unobservable configurations for their application. They too approximate the system with a piecewise constant one and consider a constant velocity motion model which constitutes a larger unobservable subspace than the vehicle specific kinematic model in [10]. The observability analysis we present over here however, is not for the purpose of solving the observability problem for 2-D monocular SLAM in general. Indeed it has already been solved by Belo et al in [14], in which they show that 3 known landmarks are enough to guarantee the observability of monocular SLAM. Our analysis over here however, is for theoretically demonstrating that the observability of monocular SLAM is indeed influenced by the inclusion of disparity measurements. We do so by evaluating the observability rank criterion for non-linear systems as presented in [15]. Due to the complexity of the analytical calculation for 2-D, we assume a 1.5-D scenario in which the robot is constrained to move along a straight line in a 2-D world and conjecture that if disparity measurements are useful in 1.5D, they must also be so in 2-D (which we show through simulations).

Our monocular SLAM algorithm is based on the Extended Kalman Filter. As mentioned before, we use both projective and disparity measurements. As our disparity measurements are based on motion-stereo, this measurement is dependent on the current *and* previous states. Similar systems using motion stereo have been developed in the past like [16], [17]. However, these approaches apply corrections separately to the image sequence, and then use the conventional filter for depth estimation. In our work, we have incorporated

[1]Talha Manzoor is a PhD student in Electrical Engineering at Lahore University of Management Sciences, Lahore, Pakistan 13060023 at lums.edu.pk

[2]Abubakr Muhammad is with the faculty of Electrical Engineering, Lahore University of Management Sciences, Lahore, Pakistan abubakr at lums.edu.pk

disparity measurements per se in the EKF by rederiving the EKF equations through the Bayesian approach given in [18]. We use the law of total probability twice, to establish the conditional dependence of the belief both on the current and previous states. After the proof is complete, we see that the affect of this dependence is simply an increase in measurement noise. As intuition suggests, this increase is dependent both on the process model and on the process noise covariance.

Ham et al [1] have shown that the eigenvalues and eigenvectors of the state covariance matrix of the Kalman Filter can provide useful insight into the observability of the system. They show that the magnitude of the eigenvalues can provide a sense of the degree of observability where the eigenvector with the smallest eigenvalue gives the most observable linear combination of the states and the eigenvector with the largest eigenvalue giving otherwise. Hence we use the trace of the covariance matrix, to analyze the performance of a simulated monocular SLAM scenario both with and without disparity measurements.

The organization of the paper is as follows. Section II gives an observability analysis, comparing two 1.5D monocular SLAM systems with and without disparity measurements. The equations for the generalized EKF for measurement models dependent on the previous state are derived in Section III. The measurement equations and discussion about the coloredness of the measurement noise are given in Section IV. Results are presented in Section V with the conclusion in Section VI.

II. Observability Analysis for a 1.5-D Monocular SLAM Scenario

We show here that including disparity as a separate measurement does indeed effect the performance of monocular SLAM. One of the necessary conditions for the estimation error of the discrete time Kalman Filter to remain bounded is that the system must satisfy the non-linear observability rank condition [19]. We first state this condition for general non-linear systems.

A. The Non-linear Observability Rank Condition

Consider the following discrete-time non-linear system

$$x_{k+1} = \mathrm{f}(x_k, u_k); \quad y_{k+1} = h(x_{k+1}), \tag{1}$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^p$ and $y \in \mathbb{R}^m$. This system is observable if for *any* initial state x_0 and *some* final time k the initial state x_0 can be uniquely determined by knowledge of the input u_i and output y_i for all $i \in [0, k]$.
Checking the observability of a system amounts to checking the rank of a certain *observability matrix* [15], [20]. If the system given by Equation 1 is in *control affine* form such that $\mathrm{f}(x_k, u_k) = g^0(x_k) + \sum g^i(x)u_{ik}$, then it is observable if and only if the rank of the following observability matrix is full

$$\mathcal{O} = \begin{bmatrix} \left(\mathbf{d}\mathcal{L}_\mathrm{f}^0(h)\right)^T \left(\mathbf{d}\mathcal{L}_\mathrm{f}^1(h)\right)^T & \cdots & \left(\mathbf{d}\mathcal{L}_\mathrm{f}^{n-1}(h)\right)^T \end{bmatrix}^T \tag{2}$$

where $\mathcal{L}_\mathrm{f}^i(h)$ is the ith order Lie Derivative [21] of the scalar field of the measurement h with respect to the vector field f and $\mathbf{d}$ denotes the gradient operator with respect to x.

B. Observability and Beacons for Monocular SLAM

The first ever monocular SLAM system presented in [2] was initialized by a certain number of known landmarks or *beacons*. This empirically suggests that the observability of monocular SLAM is related to the number of beacons present in the environment. Since then, work has been done on investigating the requirement on the minimum number of beacons for monocular/ bearing-only SLAM [10], [13]. However this is based on the approximated picewise constant systems observability theory as presented in [12]. For 2-D monocular SLAM, the rank of the observability matrix given by Equation 2 becomes too complex to compute even for a computer algebraic package. Thus, to show that disparity measurements do indeed affect the observability of monocular SLAM, we carry out an observability analysis for a simple case of the general problem.

C. Observability for 1.5-D Monocular SLAM

Consider the problem posed in Figure 1. A robot constrained to move on a single line, attempts to perform SLAM in a 2-D world containing a single beacon ($L1$), and a single landmark ($L2$). It carries a single camera and receives as measurements, $u1_k$ and $u2_k$ which are the respective projections of $L1$ and $L2$ on the image plane at time k. The robot can only move along the x-axis and its location at time k is given by c_k. The location of a landmark is given by $(x_L, \frac{1}{d_L})$ where $\frac{1}{d_L}$ is the inverse depth of the landmark. We

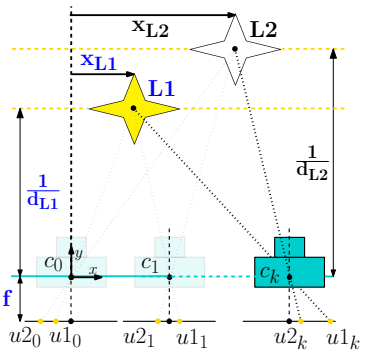

Fig. 1. The robot performing SLAM in a 2-D world with 1 beacon. Here, f is the focal length of the camera. Quantities highlighted in blue are known, whereas those in black are to be estimated.

now calculate the observability matrix for a constant velocity motion model. The state vector is given by

$$X_k = \begin{bmatrix} c_k & \dot{c}_k & x_{L2_k} & d_{L2_k} \end{bmatrix}^T. \tag{3}$$

The constant velocity model is given by the following vector-valued function

$$\mathrm{f} = \begin{bmatrix} c_k + \Delta t \dot{c}_k & \dot{c}_k & x_{L2_k} & d_{L2_k} \end{bmatrix}^T. \tag{4}$$

The measurement vector is given by

$$Y_k = \begin{bmatrix} u1_k - x_{L1} f d_{L1} \\ u2_k \end{bmatrix} = \begin{bmatrix} -c_k f d_{L1} \\ x_{L2_k} f d_{L2_k} - c_k f d_{L2_k} \end{bmatrix}.$$

Note that the first measurement is not exactly equal to the projective measurement. A bias term, consisting of all known quantities has been subtracted. Applying the definition of Equation 2, the observability matrix of this system turns out to be full rank, hence the process is observable. Also, it can be checked that the rank of O will not be full if we exclude the beacon. Thus at least 1 beacon is required for monocular SLAM in this setting to be observable. The exact form of the observability matrix for both cases can be found in [22].

D. 1.5-D Monocular SLAM with disparity measurements

We now see the effect of including disparity as a measurement on observability of the system in Figure 1. Here we neglect the measurements of $L1$ and assume that there is no beacon. The process and state are the same as given in Equations 3 and 4. We now have a disparity measurement given by

$$y_{3_k} = u2_k - u2_{k-1} = -f c_k d_{L2_k} + f c_{k-1} d_{L2_k}. \tag{5}$$

Note that this measurement is dependent on the previous state. Section III discusses the effect this dependence would have on the EKF. The observability matrix in this case turns out to have full rank [22]. Thus we see that including disparity as a measurement has eliminated the requirement of the beacon in order to keep SLAM observable. This shows that disparity is indeed a valid measurement and that it also improves the observability of monocular SLAM in 1.5-D.

III. EKF for Measurement Models dependent on the previous State

The measurement model shown in Equation 5 is dependent on the previous and currrent states. However, the conventional Kalman filter assumes that the measurement is dependent on the current state only. We derive the Kalman filter for this case from the basic Bayes Filter following the method given in [18]. Before presenting our equations, we first describe the motion and measurement models. After presenting the equations, we discuss the mathematical proof.

A. Functions and variables

1) Motion Model: The motion or process model is given by $X_k = f(X_{k-1}, \mu_k, \omega_k)$ where ω_k is the process noise at step k. Its covariance is given by Q_k. After linearization we get the following form

$$X_k = f(\hat{X}_{k-1}, \mu_k, 0) + F_{X_{k-1}}(X_{k-1} - \hat{X}_{k-1}) + F_{\omega_k}\omega_k,$$

where $F_{X_{k-1}}$ is the Jacobian of f w.r.t X_{k-1}, and F_{ω_k} is the Jacobian of f w.r.t ω_k

2) Inverse Motion Model: The inverse motion model is given by $X_{k-1} = g(X_k, \mu_k, \omega_k)$. Note that for same time steps, ω_k has the same value as its instance appearing in the motion model. After linearization we get

$$X_{k-1} = g(\hat{X}_k, \mu_k, 0) + G_{X_k}(X_k - \hat{X}_k) + G_{\omega_k}\omega_k,$$

where G_{X_k} is the Jacobian of g w.r.t X_k, and G_{ω_k} is the Jacobian of g w.r.t ω_k.

3) Measurement Model: The measurement model is given by $Y_k = h(X_k, X_{k-1}) + v_k$ where v_k is the sensor noise. Its covariance is given by R_k. After linearization we get

$$\begin{aligned} Y_k =& h(\hat{X}_k, \hat{X}_{k-1}) + H_{X_k}(X_k - \hat{X}_k) \\ &+ H_{X_{k-1}}(X_{k-1} - \hat{X}_{k-1}) + v_k, \end{aligned}$$

where H_{X_k} is the Jacobian of h w.r.t X_k, and $H_{X_{k-1}}$ is the Jacobian of h w.r.t X_{k-1}.

B. Generalized Kalman Filter Equations

Our final equations for the motion and measurement models just presented are given as follows:

$$\hat{X}_{k|k-1} = f(\hat{X}_{k-1|k-1}, \mu_k, 0), \tag{6}$$

$$P_{k|k-1} = F_{X_{k-1}} P_{k-1|k-1} {F_{X_{k-1}}}^T + F_{\omega_k} Q_k {F_{\omega_k}}^T, \tag{7}$$

$$W_k = R_k + H_{X_{k-1}} G_{\omega_k} Q_k {G_{\omega_k}}^T {H_{X_{k-1}}}^T, \tag{8}$$

$$K_k = P_{k|k-1} {H_{X_k}}^T (H_{X_k} P_{k|k-1} {H_{X_k}}^T + W_k)^{-1}, \tag{9}$$

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k(Y_k - h(\hat{X}_{k|k-1}, \hat{X}_{k-1|k-1})), \tag{10}$$

$$P_{k|k} = (I - K_k H_{X_k}) P_{k|k-1}, \tag{11}$$

where $\hat{X}_k$ is the estimate of the state X, P_k is the covariance of $\hat{X}_k$, μ_k is the control vector, $\hat{Y}_k$ is the expected measurement, K_k is the Kalman gain, W_k is the covariance of $\hat{Y}_k$, $f(X_{k-1}, \mu_{k-1})$ is the motion model, $g(X_k, \mu_k)$ is the inverse motion model and $h(X_k, X_{k-1})$ is the measurement model. All quantities are evaluated at time k.

The difference between the conventional EKF and the one given by Equations 6-11 lies in the calculation of Equation 8. For simple measurement models, W_k will be exactly equal to the sensor noise R_k with the rest of the equations remaining unchanged. Note that for linear systems, the inverse motion model will just be equal to F^{-1}. We assume that such an inversion exists.

C. Mathematical Derivation

The equations are derived first by extending the basic Bayes Filter Algorithm [18], to include the previous state in the measurement step. Then the probability distributions are solved for the required models. To include the dependency of Y_k on X_{k-1} we include an intermediate step in the Bayes Filter through the theorem of total probability. The steps are as follows:

Step 1: $\overline{bel(X_k)} = p(X_k | Y_{0:k-1}, \mu_k) =$

$$\int p(X_k | X_{k-1}, Y_{0:k-1}, \mu_k) p(X_{k-1} | Y_{0:k-1}, \mu_k) \mathrm{d}X_{k-1}, \tag{12}$$

Step 2: $p(Y_k|X_k, Y_{0:k-1}, \mu_k) =$

$$\int p(Y_k|X_k, Y_{0:k-1}, \mu_k, X_{k-1}) \; p(X_{k-1}|X_k, Y_{0:k-1}, \mu_k)\mathrm{d}X_{k-1}, \tag{13}$$

Step 3: $bel(X_k) = \eta p(Y_k|X_k, Y_{0:k-1}, \mu_k)p(X_k|Y_{0:k-1}, \mu_k)$, (14)

where η is a normalizing constant. Note that the conventional Bayes filter [18] consists only of steps 1 and 3. We present the derivation, with this extended form of the Bayes filter.

1) Part 1: Prediction: We begin with step 1 of the Bayes filter. The terms of Equation (12) are normally distributed with mean and covariance as

$$\underbrace{p(X_k|X_{k-1}, Y_{0:k-1}, \mu_k)}_{N\sim(X_k; f(\hat{X}_{k-1},\mu_k,0)+F_{X_{k-1}}(X_{k-1}-\hat{X}_{k-1}); F_{\omega_k}Q_k{F_{\omega_k}}^T)} \quad \text{and,}$$

$$\underbrace{p(X_{k-1}|Y_{0:k-1}, \mu_k)}_{N\sim(X_{k-1};\hat{X}_{k-1};P_{k-1})}.$$

After evaluation of this equation, the outcome is also a Gaussian with mean $\hat{X}_{k|k-1}$ and covariance $P_{k|k-1}$. The interested reader is directed to [18] for a full proof. The mean and variance of the resulting distribution are given as $\hat{X}_{k|k-1} = f(X_{k-1}, \mu_k, 0)$ and $P_{k|k-1} = (F_{\omega_k}Q_k{F_{\omega_k}}^T) + F_{X_{k-1}}P_{k-1}{F_{X_{k-1}}}^T$ respectively.

2) Part 2: Measurement Covariance: Here we derive step 2 of the extended Bayes filter given by (13). The distributions of the terms involved are given as follows

$$\underbrace{p(Y_k|X_k, Y_{0:k-1}, \mu_k, X_{k-1})}_{(\sim N(Y_K; h(\hat{X}_k,\hat{X}_{k-1})+H_{x_k}(X_k-\hat{X}_k)+H_{X_{k-1}}(X_{k-1}-\hat{X}_{k-1}); R_k))} \quad \text{and,}$$

$$\underbrace{p(X_{k-1}|X_k, Y_{0:k-1}, \mu_k)}_{(\sim N(X_{k-1}; g(\hat{x}_k,\mu_k,0)+G_{X_k}(X_k-\hat{X}_k); G_{\omega_k}Q_k{G_{\omega_k}}^T))}.$$

The derivation is essentially the same as that of step 1 given in [18]. In this case the integral in Equation 13 can be expressed as $\int \exp\{-M_k\}\mathrm{d}X_{k-1}$, where

$$\begin{aligned} M_k = & \\ & \frac{1}{2}\Big(Y_k - h(\hat{X}_k,\hat{X}_{k-1}) - H_{x_k}(X_k-\hat{X}_k) - H_{X_{k-1}}(X_{k-1}-\hat{X}_{k-1})\Big)^T \\ & {R_k}^{-1}\Big(Y_k - h(\hat{X}_k,\hat{X}_{k-1}) - H_{x_k}(X_k-\hat{X}_k) - H_{X_{k-1}}(X_{k-1}-\hat{X}_{k-1})\Big) \\ & + \frac{1}{2}\Big(X_{k-1} - g(\hat{x}_k,\mu_k,0) - G_{X_k}(X_k-\hat{X}_k)\Big)^T \Big(G_{\omega_k}Q_k{G_{\omega_k}}^T\Big)^{-1} \\ & \Big(X_{k-1} - g(\hat{x}_k,\mu_k,0) - G_{X_k}(X_k-\hat{X}_k)\Big). \end{aligned} \tag{15}$$

We can decompose M_k into two functions. These functions are given by

$$M_k = M(Y_k, X_k, X_{k-1}) + M(Y_k, X_k). \tag{16}$$

By differenciating twice with respect to X_{k-1}, we determine the minimum and covariance of M_k so as to form a quadratic function similar to the exponential of the normal distribution. Thus we get

$$\begin{aligned} & M(Y_k, X_k, X_{k-1}) = \\ & \frac{1}{2}\Big(X_{k-1} - \Phi_k\Big({H_{X_{k-1}}}^T {R_k}^{-1}(Y_k - h(\hat{X}_k,\hat{X}_{k-1}) - H_{x_k}(X_k-\hat{X}_k) \\ & - H_{X_{k-1}}(X_{k-1}-\hat{X}_{k-1})) \\ & + (G_{\omega_k}Q_k{G_{\omega_k}}^T)^{-1}(g(\hat{x}_k,\mu_k,0) - G_{X_k}(X_k-\hat{X}_k))\Big)\Big)^T \Phi_k^{-1} \\ & \Big(X_{k-1} - \Phi_k\Big({H_{X_{k-1}}}^T {R_k}^{-1}(Y_k - h(\hat{X}_k,\hat{X}_{k-1}) - H_{x_k}(X_k-\hat{X}_k) \\ & - H_{X_{k-1}}(X_{k-1}-\hat{X}_{k-1})) \\ & + (G_{\omega_k}Q_k{G_{\omega_k}}^T)^{-1}(g(\hat{x}_k,\mu_k,0) - G_{X_k}(X_k-\hat{X}_k))\Big)\Big), \end{aligned} \tag{17}$$

where $\Phi_k = \Big({H_{X_{k-1}}}^T {R_k}^{-1} H_{X_{k-1}} + (G_{\omega_k}Q_k{G_{\omega_k}}^T)^{-1}\Big)^{-1}$. Hence from Equations (15), (16) and (17), we get $M(Y_k, X_k)$. As this does not have any dependence on X_{k-1}, we can move it out of the integral. What remains inside the integral is $\exp\{-M(Y_k, X_k, X_{k-1}\}$ which is a valid probability density function hence integrates to a constant. Thus $\exp\{-M(Y_k, X_k\}$ gives the required distribution. The mean and covariance of this distribution is given by the minimumm and inverse of the curvature of $M(Y_k, X_k)$ respectively. Differentiating with respect to Y_k, we get the mean as $h(\hat{X}_k, \hat{X}_{k-1}) + H_{X_k}(X_k - \hat{X}_k)$, and the covariance, the inverse of the curvature as $W_k = R_k + H_{X_{k-1}}G_{\omega_k}Q_k{G_{\omega_k}}^T{H_{X_{k-1}}}^T$. This proves the correctness of (8).

3) Part3: Correction / Update: The distributions of the terms involved in the 3rd step of the Bayes Filter are

$$\underbrace{p(Y_k|X_k, Y_{0:k-1}, \mu_k)}_{\Big(\sim N(Y_k; h(\hat{X}_k,\hat{X}_{k-1})+H_{X_{k-1}}(X_k-\hat{X}_k))\Big)} \quad \text{and,}$$

$$\underbrace{p(X_k|Y_{0:k-1}, \mu_k)}_{\Big(\sim N(X_k; f(\hat{X}_{k-1},\mu_k,0); F_{X_{k-1}}P_{k-1}{F_{X_{k-1}}}^T + F_{\omega_k}Q_k{F_{\omega_k}}^T)\Big)}.$$

The resulting distribution follows from [18] by replacing R_k with W_k so that we get the remaining equations.

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k\Big(Y_k - h(\hat{X}_k, \hat{X}_{k-1})\Big),$$

$$K_k = P_{k|k-1}{H_{X_k}}^T(H_{X_k}P_{k|k-1}{H_{X_k}}^T + W_k)^{-1},$$

$$P_{k|k} = (I - K_kH_{X_k})P_{k|k-1}.$$

IV. Planar Monocular SLAM

In this section, we setup the EKF presented in Section III for planar monocular SLAM and show through simulations, the effect of disparity measurements.

A. System description

Consider Figure 2, in which a camera performs SLAM in a 2-D world. The state vector is given by

$$X_k = \begin{bmatrix} {}^cx_k & {}^cx'_k & {}^cy_k & {}^cy'_k & {}^c\theta_k & {}^c\theta'_k & {}^{L2}x_k & {}^{L2}y_k \end{bmatrix}^T.$$

Here we assume a constant velocity motion model for state and covariance propagation. The measurement equation

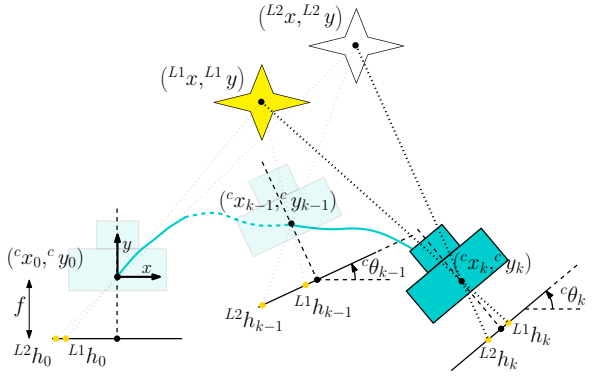

Fig. 2. Monocular state sequence in 2-D

requires ${}^{L1}x$, ${}^{L1}y$, ${}^{L2}x$ and ${}^{L2}y$ to be described in the camera coordinate frame. Transforming the world coordinates gives us ${}^{C}L1 = \begin{bmatrix} {}^{L1}\overline{x} & {}^{L1}\overline{y} \end{bmatrix}^T$ as

$${}^{C}L1 = \begin{bmatrix} ({}^{L1}x - {}^{c}x)\cos({}^{c}\theta) + ({}^{L1}y - {}^{c}y)\sin({}^{c}\theta) \\ -({}^{L1}x - {}^{c}x)\sin({}^{c}\theta) + ({}^{L1}y - {}^{c}y)\cos({}^{c}\theta) \end{bmatrix},$$

and so, the measurement model for the camera is given by

$${}^{L1}h = f\frac{({}^{L1}x - {}^{c}x)\cos({}^{c}\theta) + ({}^{L1}y - {}^{c}y)\sin({}^{c}\theta)}{-({}^{L1}x - {}^{c}x)\sin({}^{c}\theta) + ({}^{L1}y - {}^{c}y)\cos({}^{c}\theta)},$$

and disparity is given by ${}^{L2}\text{disp}_k = f\left({}^{L2}B_k / {}^{L2}d_k\right)$, where ${}^{L2}B_k$ is the magnitude of the baseline and ${}^{L2}d_k$ is the perpendicular distance of the landmark from the baseline. They are given as

$${}^{L2}d_k = \frac{({}^{c}x_k - {}^{c}x_{k-1})({}^{c}y_{k-1} - {}^{L2}y)}{\sqrt{({}^{c}x_k - {}^{c}x_{k-1})^2 + ({}^{c}y_k - {}^{c}y_{k-1})^2}} - \frac{({}^{c}x_{k-1} - {}^{L2}x)({}^{c}y_k - {}^{c}y_{k-1})}{\sqrt{({}^{c}x_k - {}^{c}x_{k-1})^2 + ({}^{c}y_k - {}^{c}y_{k-1})^2}}, \text{and}$$

$${}^{L2}B_k = \sqrt{({}^{c}x_k - {}^{c}x_{k-1})^2 + ({}^{c}y_k - {}^{c}y_{k-1})^2}.$$

B. Colored Measurement Noise

The disparity measurement is acheived through the following steps: 1) Obtain the measurements ${}^{Li}h_{k-1}, {}^{Li}h_k$ for consecutive time steps. 2) Rectify the measurements to get ${}^{Li}h'_{k-1}, {}^{Li}h'_k$. 3) Get ${}^{L2}\text{disp}_k$ as the difference between ${}^{Li}h'_{k-1}$ and ${}^{Li}h'_k$. These steps result in the following equation for disparity

$${}^{Li}\text{disp}_k = \frac{f\left({}^{Li}h_k\cos(\alpha_k - {}^{c}\theta_k) - f^2\sin(\alpha_k - {}^{c}\theta_k)\right)}{{}^{Li}h_k\sin(\alpha_k - {}^{c}\theta_k) + f\cos(\alpha_k - {}^{c}\theta_k)} - \frac{f\left({}^{Li}h_{k-1}\cos(\alpha_k - {}^{c}\theta_{k-1}) - f^2\sin(\alpha_k - {}^{c}\theta_{k-1})\right)}{{}^{Li}h_{k-1}\sin(\alpha_k - {}^{c}\theta_{k-1}) + f\cos(\alpha_k - {}^{c}\theta_{k-1})}, \quad (18)$$

where $\alpha_k = \arctan\left(({}^{c}y_k - {}^{c}y_{k-1}) / ({}^{c}x_k - {}^{c}x_{k-1})\right)$. Equation 18 shows that the noise in the disparity measurement at time step k is correlated with the projection measurement at time step $k-1$. We can get the noise equation from Equation 18 and approximate it linearly to get, ζ_{disp_k} equal to $H_{\zeta_{h_k}}\zeta_{h_k} + H_{\zeta_{h_{k-1}}}\zeta_{h_{k-1}}$, where ζ is the variable for noise and H is the respective derivative. Once we have these matrices, we use the state augmentation technique presented in [23] to deal with colored measurement noise.

V. Simulation

A. Setup

The simulations for validation of our work has been implemented in MATLAB. In the scenario presented here, the robot follows a rhombus shaped trajectory with smoothed corners as shown in Figure 3. Velocity has been kept higher at the sides and lower at the corners. The focal length of the camera has been set at 20 pixel units. The process and measurement models are as given in Section IV. The process model is a constant velocity one with process noise modeled as zero-mean Gaussian with variances of $0.5\,(\text{m/s}^2)^2$, $0.1\ \text{deg}^2$ and $0.001\ \text{m}^2$ for the linear velocities, angular velocities and landmark positions respectively. Measurement noise variables have also been drawn from zero-mean Gaussian distributions with variances of $0.074\ \text{pixels}^2$ for projective and $74\ \text{pixels}^2$ for disparity measurements. Length of one time step has been kept at 0.1 seconds after which the robot receives projective and disparity measurements of the landmarks and beacons. As seen in Figure 3 an incorrect initial estimate is provided with large uncertainty.

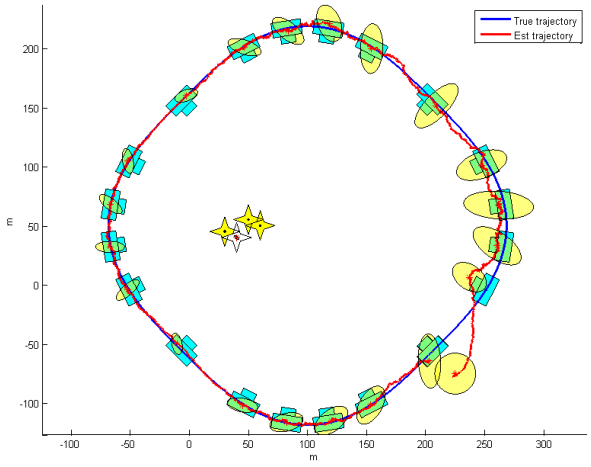

Fig. 3. A robot performing SLAM in 2-D. The cameras depict the true pose of the robot at selected time steps, equally distributed in time. The beacons are drawn in yellow whereas the estimated landmark is in white. The ellipses show the 95% confidence regions. An incorrect estimate with large uncertainty is provided.

B. Results

The results of the proposed framework can be seen in Figures 3 & 4. In Figure 3, the true robot location is almost always captured by the respective uncertainty ellipse. Due to the incorrect initialization, there is a little inconsistency for some time after the filter is initialized. Overall, the filter shows consistent behavior. This can also be seen for the uncertainty ellipse of the estimated landmark.

Figure 4 compares the performance between the conventional EKF and the modified EKF of Section III. The first two graphs show the trace of the position and map covariance

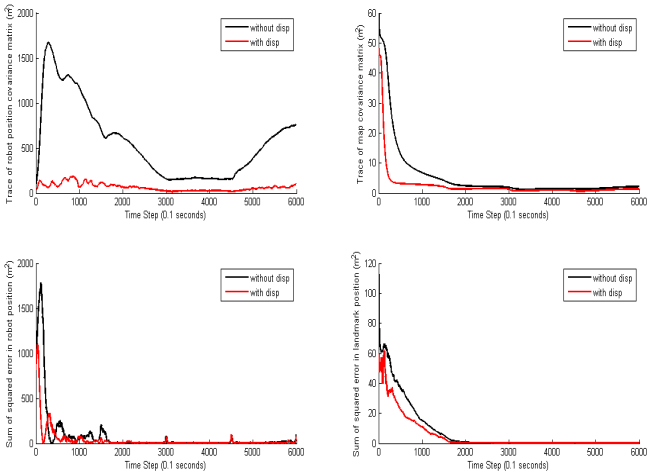

Fig. 4. Performance comparison between the conventional EKF and the generalized EKF presented in Section III. From top-left in counter clockwise order: 1) Trace of covariances for robot location, 2) Trace of covariances for landmark location, 3) SSD for robot location, 4) SSD for landmark location.

matrices. They represent the sum of the eigenvalues and thus the degree of observability of the system [1]. We see that there is considerable improvement with respect to the robot position variables. Although the steady state behavior for the landmark remains the same, the filter run with inclusion of disparity measurements, has a smaller transient. Thus we observe an improvement both for the robot and map variables. Also, we see from Figure 4 that the steady state accuracy with respect to the sum of squared errors is almost the same for both cases, with the filter run with inclusion of disparity measurements having smaller transients.

VI. Conclusion

The idea of including disparity measurements in monocular SLAM presented in this paper has been tested first theoretically through the observability rank condition and then through simulations. During this process, we have also generalized the EKF for measurement models dependent on the previous state which may have applications other than that of disparity measurements in monocular SLAM. We have seen through simulations that the degree of observability is indeed improved both for robot pose and landmark variables.

It must be noted here that disparity is already included implicitly in Monocular SLAM, and that the process is inherently unobservable especially in the scale direction. Thus the inclusion of an explicit disparity measurement cannot convert this unobservable process into an observable one. However as described in this paper, it does improve the degree of observability for all states and hence provides more confident estimates for the EKF. Although we have not conducted a mathematically rigorous proof for our proposition, we hope that this work will move theoreticians to look in this direction.

References

[1] F. M. Ham and R. G. Brown, "Observability, eigenvalues, and kalman filtering," *Aerospace and Electronic Systems, IEEE Transactions on*, no. 2, pp. 269–273, 1983.

[2] A. J. Davison, "Real-time simultaneous localisation and mapping with a single camera," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 1403–1410.

[3] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "Monoslam: Real-time single camera slam," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 6, pp. 1052–1067, 2007.

[4] L. A. Clemente, A. J. Davison, I. Reid, J. Neira, and J. D. Tardós, "Mapping large loops with a single hand-held camera," in *Robotics: Science and Systems*, 2007.

[5] H. Strasdat, J. Montiel, and A. J. Davison, "Scale drift-aware large scale monocular slam," in *Proceedings of Robotics: Science and Systems (RSS)*, vol. 2, no. 3, 2010, p. 5.

[6] L. M. Paz, P. Piniés, J. D. Tardós, and J. Neira, "Large-scale 6-dof slam with stereo-in-hand," *Robotics, IEEE Transactions on*, vol. 24, no. 5, pp. 946–957, 2008.

[7] J. Huber and V. Graefe, "Motion stereo for mobile robots," in *Industrial Electronics, 1993. Conference Proceedings, ISIE'93-Budapest., IEEE International Symposium on*. IEEE, 1993, pp. 263–270.

[8] P. Pritchett and A. Zisserman, "Wide baseline stereo matching," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 754–760.

[9] P. Bonnifait and G. Garcia, "Design and experimental validation of an odometric and goniometric localization system for outdoor robot vehicles," *Robotics and Automation, IEEE Transactions on*, vol. 14, no. 4, pp. 541–548, 1998.

[10] T. Vidal-Calleja, M. Bryson, S. Sukkarieh, A. Sanfeliu, and J. Andrade-Cetto, "On the observability of bearing-only slam," in *Robotics and Automation, 2007 IEEE International Conference on*. IEEE, 2007, pp. 4114–4119.

[11] J. Kim and S. Sukkarieh, "Improving the real-time efficiency of inertial slam and understanding its observability," in *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 1. IEEE, 2004, pp. 21–26.

[12] D. Goshen-Meskin and I. Bar-Itzhack, "Observability analysis of piece-wise constant systems. i. theory," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 28, no. 4, pp. 1056–1067, 1992.

[13] M. J. Tribou, D. W. Wang, and W. J. Wilson, "Observability of planar combined relative pose and target model estimation using monocular vision," in *American Control Conference (ACC), 2010*. IEEE, 2010, pp. 4546–4551.

[14] F. A. Belo, P. Salaris, and A. Bicchi, "3 known landmarks are enough for solving planar bearing slam and fully reconstruct unknown inputs," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 2539–2545.

[15] R. Hermann and A. Krener, "Nonlinear controllability and observability," *Automatic Control, IEEE Transactions on*, vol. 22, no. 5, pp. 728–740, 1977.

[16] L. Matthies, T. Kanade, and R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequences," *International Journal of Computer Vision*, vol. 3, no. 3, pp. 209–238, 1989.

[17] C. F. Olson and H. Abi-Rached, "Wide-baseline stereo vision for terrain mapping," *Machine Vision and Applications*, vol. 21, no. 5, pp. 713–725, 2010.

[18] S. Thrun, W. Burgard, D. Fox, *et al.*, *Probabilistic robotics*. MIT press Cambridge, MA, 2005, vol. 1.

[19] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete-time extended kalman filter," *Automatic Control, IEEE Transactions on*, vol. 44, no. 4, pp. 714–728, 1999.

[20] M. Anguelova, *Observability and identifiability of nonlinear systems with applications in biology*. Chalmers University of Technology, 2007.

[21] J.-J. E. Slotine, W. Li, *et al.*, *Applied nonlinear control*. Prentice-Hall Englewood Cliffs, NJ, 1991, vol. 199, no. 1.

[22] T. Manzoor, "Observability and disparity in monocular slam: A study," Master's thesis, Lahore University of Management Sciences, 2013.

[23] R. S. Bucy and P. D. Joseph, *Filtering for stochastic processes with applications to guidance*. American Mathematical Society, 1987, vol. 326.