Geometric Map-Assisted Localization for Mobile Robots Based on Uniform-Gaussian Distribution

Article *in* IEEE Robotics and Automation Letters · January 2017

DOI: 10.1109/LRA.2017.2651385

5 authors, including:

Rui Jiang
National University of Singapore
21 PUBLICATIONS **129** CITATIONS
SEE PROFILE

Han Wang
University of Virginia
45 PUBLICATIONS **325** CITATIONS
SEE PROFILE

Geometric Map-Assisted Localization for Mobile Robots Based on Uniform-Gaussian Distribution

Rui Jiang[1], Shuai Yang[2], Shuzhi Sam Ge[1*], Han Wang[2], and Tong Heng Lee[1]

***Abstract*—Drift and scale ambiguity are two main issues which reduce localization accuracy in Monocular Visual Odometry (MVO). It is necessary to propose a unified model to represent these measurement uncertainties. In this paper, we present a geometric map-assisted localization approach for mobile robots equipped with MVO. We model the measurement of MVO as a group of particles, which obey Uniform-Gaussian Distribution and cover measurement uncertainties including scale ambiguity and measurement randomness. The *saliency* of each particle can be obtained from the distribution to indicate raw measurement certainty of MVO. Geometric map-assisted shape matching is implemented as the measurement model to assign *consistency* to the particles generated from the distribution. Both *saliency* and *consistency* are considered in particle weights determination. Furthermore, based on the statistical properties of the probability distribution, a parameter estimation scheme is proposed to narrow down the scale ambiguity of MVO while resampling particles. Experiments with KITTI dataset have demonstrated that the proposed approach greatly enhances positioning accuracy, with average localization error of 6.54 meters in over 15.89 kilometers run.**

***Index Terms*—Autonomous Vehicle Navigation, Localization, Sensor Fusion.**

I. Introduction

THE problem of mobile robot localization is well-studied over the last few years [1]. Although Global Positioning System (GPS)-based approach has become one of the most popular positioning solutions, it is necessary to develop non-GPS based approaches due to inherent drawbacks of GPS, such as the instability in urban canyon environments [2] and the poor accuracy at short distance [3]. The camera as a cost-effective, easy-to-use and lightweight sensor, has been widely used to solve the localization problem, and impressive results have been achieved [4], [5].

Visual Odometry (VO) as the most commonly used method, is a self-contained dead-reckoning localization process and has been successfully implemented in Mars Exploration Rovers [6]. It is known that VO suffers from drift, which is the primary factor limiting its long-range applications. Compared to Stereo Visual Odometry (SVO), Monocular Visual Odometry (MVO) shows advantages in cost and size. However, the scale ambiguity makes MVO even harder for pose recovery without adding constraints [7]. Map-assisted positioning provides a new perspective for robot localization, especially in urban areas. With the assumption that the robot is always on the road, the geometric shape of road networks can be used as constraints to assist with position estimation [8]–[10].

The objective of this work is to globally localize a vehicle equipped with one single camera and a freely available geometric map. To complete this goal, a framework that combines MVO measurement and a geometric map is presented. In this framework, MVO is used to generate particles, i.e. possible trajectories, and shape matching between trajectories and the map is used to weight the particles.

The drift and scale ambiguity of MVO are both considered as measurement uncertainties and are incorporated into a unified probability distribution. Compared to the relevant work, the main contributions of the paper are summarized as follows:

1) Uniform-Gaussian Distribution (UGD) is proposed to describe measurement uncertainties in VO. The model can be used for both SVO and MVO by combining the drift and scale ambiguity in an intuitive way.
2) A parameter estimation scheme is presented to refine the probability distribution according to sample particles. The parameter estimation result is used for generating particles iteratively, similar to Monte Carlo localization framework.
3) Combined with particle *saliency* representing VO measurement certainty degree and particle *consistency* denoting accordance with the geometric map, a geometric map-assisted localization framework is proposed to reduce drift and scale ambiguity of VO.

II. Related Work

Visual odometry methods have been widely used for mobile robot localization thanks to their constantly improving performance. However, several challenges still needs to be resolved. Error accumulation or the so-called drift issue is the first challenge preventing VO from being used in long range navigation. Concretely, the motion estimated from SVO drifts in six degrees of freedom, while the motion estimated from MVO drifts in seven degrees of freedom, with an additional scale drift. To correct the drift, loop closure detection, followed by a Global Bundle Adjustment or graph optimization [11]

Manuscript received: September 5, 2016; Revised November 21, 2016; Accepted December 19, 2016.

This paper was recommended for publication by Editor Cyrill Stachniss upon evaluation of the Associate Editor and Reviewers' comments. This work is supported by Defence Innovative Research Programme (DIRP), the Ministry of Defence, Singapore under grant R-263-000-B08-592.

[1]Rui Jiang, Shuzhi Sam Ge* (corresponding author) and Tong Heng Lee are with the Department of Electrical and Computer Engineering, National University of Singapore, 4 Engineering Drive 3, Singapore 117583 {rui_jiang,samge,eleleeth}@nus.edu.sg

[2]Shuai Yang and Han Wang are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 {sayang,hw}@ntu.edu.sg

The first two authors contributed equally.

Digital Object Identifier (DOI): see top of this page.

step, is widely used, and impressive results have been achieved [12]. Unfortunately, loops do not necessarily exist in practical driving conditions. Even though when loops do exist, the corrected motion is still a delayed result for the route before loop closure. Thus, loop closing method is not appropriate for applications where instantaneous decisions are desired, such as an autonomous vehicle.

The second challenge is how to obtain the metric scale of the motion estimated from a monocular system. The most straight forward approach is fuse information from IMU, GPS or other sensors [13], which of course will increase the cost and complexity of the system. Another popular method to estimate the metric scale is assuming that the camera is moving at a known and fixed height over ground [14]. However, the result of this kind of method relies heavily on the accuracy of ground plane detection. Some other researchers proposed to use objects with known size to give the absolute scale of monocular results [15]. Nevertheless, it is quite tough to ensure the object appear and be detected in all frames.

To solve the drift and scale ambiguity challenge, we turn to freely available geometric maps. Geometric maps, such as OpenStreeMaps (OSM) and Google Maps have plenty of information which can be used for localization. On the one hand, since street segments and the connectivity of roads can be intuitively expressed as nodes and edges of a graph model, then a geometric map can be represented by a directed graph, which is much simpler to work with in localization problem. In [8], a graph-based representation of the map was defined and a probabilistic map localization approach was proposed. They achieved an accuracy of 3 meters within a city-level road map by using VO measurements and OSM data as the only inputs. To reduce the high computational cost, a simplified approach which used wheel speed odometry instead of VO was proposed in [10] and real-time performance were achieved. On the other hand, the geometric shape of road networks can be considered as constraints to assist with position estimation. Authors of [16] succeeded in localizing a mobile robot on sidewalks by using a combination of SVO and satellite map matching. Top view images generated from stereo frames were matched with satellite map to correct incremental drifts. Shape matching method was utilized to evaluate the alignment of different trajectories to the map in [9] and a shape matching process was considered as the measurement model of Monte Carlo localization framework. Similarly, shape matching is also used in our work. However, there is a difference that we focus on monocular camera localization while they worked on the stereo case. Thus, not only motion drift but also scale ambiguity are considered in this paper.

III. Geometric Map-Assisted Localization

The framework of the proposed approach is shown in Fig. 1. The map preparation process requires a rough positioning result such that the geometric map can be prepared. If the rough position is not available, any other position indication such as a block name will do as long as the map of the particular area can be prepared. The duration of the initialization process depends on practical road conditions. Since only sufficiently complex trajectory can be used for shape matching, the main loop will not be executed unless the trajectory satisfies certain geometric conditions.

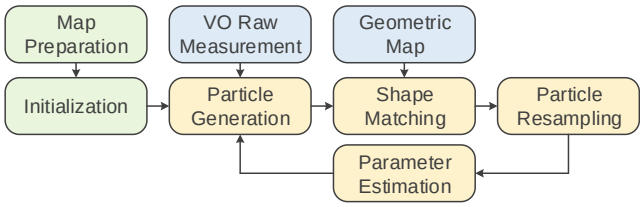

Fig. 1. Framework of the geometric map-assisted localization approach.

After each time particles are generated based on UGD, shape matching is implemented to evaluate the similarity between geometric map and all probable trajectories by examining each possible trajectory consists of historical and current estimations. Particle *saliency* and *consistency* are obtained to indicate the representativeness of a particle for VO measurement and the degree of similarity between estimated trajectory and geometric map, respectively. Then particle weights are determined by considering both *saliency* and *consistency*. Next, a resampling process is carried out to select particles with higher weights and ignore particles with lower weights. The remaining particles are used to update parameters of UGD. By the framework mentioned above, the robot is localized while VO measurement uncertainties is reduced simultaneously.

A. Measurement Equations of Visual Odometry

It is impossible to obtain the measurement without uncertainty. Conventional Gaussian probability assumption makes sense in dealing with MVO measurement randomness. However, MVO is not able to recover trajectory's absolute scale and thus scale ambiguity exists. This section demonstrates measurement equations of MVO and makes an assumption in preparations for modeling measurement uncertainties.

Visual odometry outputs raw transformation matrix by measuring camera pose difference between consecutive image frames [17]:

$$\mathbf{T}^{\text{raw}}_{k,k-1} = \begin{bmatrix} \mathbf{R}^{\text{raw}}_{k,k-1} & \mathbf{t}^{\text{raw}}_{k,k-1} \\ \mathbf{0} & 1 \end{bmatrix} \tag{1}$$

where $\mathbf{R}^{\text{raw}}_{k,k-1} \in SO(3)$, $\mathbf{t}^{\text{raw}}_{k,k-1} \in \mathbb{R}^3$ denote the rotation and translation parts of $\mathbf{T}^{\text{raw}}_{k,k-1}$. The raw camera pose $\mathbf{C}^{\text{raw}}_k = \left[\mathbf{R}^{\text{raw}}_k | \mathbf{t}^{\text{raw}}_k\right]$ can be computed by the raw measurement equation $\mathbf{C}^{\text{raw}}_k = \mathbf{C}^{\text{raw}}_{k-1}\mathbf{T}^{\text{raw}}_{k,k-1}$ where $\mathbf{R}^{\text{raw}}_k$ and $\mathbf{t}^{\text{raw}}_k$ are raw rotation matrix and raw translation vector at time instant k, respectively.

Since maps are two-dimensional projections from real geographic information, to incorporate road constraints into VO measurement, two-dimensional simplification is necessary. Moreover, as the focus of the paper is translation rather than rotation in the localization problem, relation between raw measurement and scaled measurement (VO aligned with actual scale) can be expressed as

$$\mathbf{t}_{k,k-1} = s_k \mathbf{t}^{\text{raw}}_{k,k-1} \tag{2}$$

where $s_k \in \mathbb{R}^+$ is a positive scaling factor at time instant k. Eqn. (2) is more generalized and can be applied to SVO

by letting $s_k = 1$. To push forward, we make the following assumption:

Assumption 1: In MVO measurement model (2), for every time instant k, s_k is independent with each element in $\mathbf{t}_{k,k-1}$, and the elements in $\mathbf{t}_{k,k-1}$ are also mutually independent.

B. Uniform-Gaussian Distribution

The UGD is proposed to represent MVO measurement with respect to drift and scale ambiguity. With (2), we use a Gaussian-distributed random vector $\mathbf{X}$ to represent raw translation measurement $\mathbf{t}^{\text{raw}}_{k,k-1}$. As we have no a priori knowledge about the scaling factor s_k, it is reasonable to assume that s_k can be modeled using a random variable S that obeys uniform distribution in the interval $[a, b]$, where $[a, b]$ denotes current probable range of s_k. Thus, the scaled translation measurement $\mathbf{t}_{k,k-1}$ is expressed as a random vector $\mathbf{Y} = S\mathbf{X}$, which leads to the definition of UGD as follows:

Definition 1 (Uniform-Gaussian Distribution, UGD): Given random variable S and random vector $\mathbf{X}$, the variate $\mathbf{Y} = S\mathbf{X}$ obeys UGD $UG(a, b, E\mathbf{X}, \Sigma)$ if $S \sim U(a, b)$ and $\mathbf{X} \sim N(E\mathbf{X}, \Sigma)$, where $U(\cdot)$ and $N(\cdot)$ are abbreviations for uniform and Gaussian distribution, respectively.

The above definition generalizes Gaussian distribution: If $a = b$, then S will be a constant, which results $\mathbf{Y}$ in a multivariate bell-shaped distribution. Based on Assumption 1, S is independent with each variable in $\mathbf{X}$, and variables in $\mathbf{X}$ are mutually independent. Several important properties, which will be used for parameter estimation, are discussed here.

Property 1 (PDF of UGD): The Probability Density Function (PDF) of j-th random variable in $\mathbf{Y}$ can be represented as [18]

$$f_{Y_j}(y) = \int_{-\infty}^{+\infty} f_S(s) f_{X_j}(\frac{y}{s}) \frac{1}{|s|} \mathrm{d}s \tag{3}$$

If $0 < a \leq b$, above PDF can be simplified as

$$f_{Y_j}(y) = \frac{1}{(b-a)\sigma_j\sqrt{2\pi}} \int_a^b \frac{1}{s} \exp\left\{ -\frac{[y/s - E(X_j)]^2}{2\sigma_j^2} \right\} \mathrm{d}s \tag{4}$$

where σ_j^2 denotes the j-th diagonal element in covariance matrix Σ.

Property 2 (Expectation of UGD): The expectation of $\mathbf{Y}$ is the multiplication of expectations of S and $\mathbf{X}$:

$$E\mathbf{Y} = \frac{a+b}{2} E\mathbf{X} \tag{5}$$

Property 3 (Variance of UGD): The variance of j-th random variable in $\mathbf{Y}$ can be represented as

$$D(Y_j) = \frac{a^2 + b^2 - 2ab}{12} E(X_j)^2 + \frac{a^2 + b^2 + ab}{3} \sigma_j^2 \tag{6}$$

Given a UGD with certain parameters, samples can be generated according to the Algorithm 1.

Algorithm 1 UGD Sampling

Input: UGD parameters $(a, b, E\mathbf{X}, \Sigma)$, number of samples N.
Output: Samples $\mathbf{y}_i, i \in \{1, ..., N\}$
1: **for** $i \leftarrow 1$ to N **do**
2: $s_i \leftarrow Uniform(a, b)$ {generating uniform-distributed s_i}
3: $\mathbf{x}_i \leftarrow Norm(E\mathbf{X}, \Sigma)$ {generating Gaussian-distributed $\mathbf{x}_i$}
4: $\mathbf{y}_i \leftarrow s_i \mathbf{x}_i$
5: **end for**

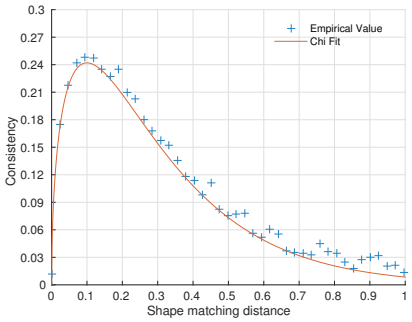

Fig. 2. Particle *consistency* distribution with regard to shape matching distance.

C. Measurement Representation with UGD

Scaled measurement of MVO can be regarded as Uniform-Gaussian distributed samples $\mathbf{y}_i$ that are the multiplications of uniformly distributed samples s_i and Gaussian distributed samples $\mathbf{x}_i$. Since only single measurement can be obtained at a certain time, in the meaning of Maximum Likelihood Estimation, the raw measurement is estimated as the expectation of Gaussian distribution $E\mathbf{X}$. The covariance matrix Σ, which represents the uncertainty of VO result, can be obtained from the exponential drift model [19] as $\Sigma = \gamma_0^{d_{\text{VO}}} \Sigma_0$, where γ_0 and Σ_0 are platform-dependent constants, which denote the base growth rate and the initial covariance matrix, respectively; d_{VO} is the distance measured by VO. Parameters a and b can be evaluated iteratively based on the estimation algorithm, which will be detailed in Section III-E.

A particle with position coordinates is an intuitive representation of a sample. Thus, given samples from UGD, a group of particles is created to represent scaled MVO measurement. In order to measure the representativeness of a particular particle, *saliency* $\mu_{\text{VO}}(\mathbf{y}_i) \in (0, 1]$ can be defined uniquely for each particle as

$$\mu_{\text{VO}}(\mathbf{y}_i) = \exp\left\{ -\frac{1}{2} \left[\frac{\mathbf{y}_i}{s_i} - E\left(\frac{\mathbf{y}_i}{s_i}\right) \right]^T \Sigma^{-1} \left[\frac{\mathbf{y}_i}{s_i} - E\left(\frac{\mathbf{y}_i}{s_i}\right) \right] \right\} \tag{7}$$

$$= \exp\left\{ -\frac{1}{2} \left[\mathbf{x}_i - E\mathbf{X} \right]^T \Sigma^{-1} \left[\mathbf{x}_i - E\mathbf{X} \right] \right\} \tag{8}$$

Above definition is rational since the particle which satisfies $\mathbf{x}_i = E\mathbf{X}$ is regarded as the most representative particle for VO measurement. In practice, the scaled measurement from VO can be represented as a set, which contains a group of particles and their corresponding *saliency*.

D. Geometric Map-assisted Shape Matching

The geometric map that contains road shape information is used to correct VO trajectory. This idea is quite similar to shape matching, whose goal is to find the best alignment between two edge maps. To use the geometric map effectively, we first convert it to an edge map $\mathcal{U} = \{\mathbf{u}_p\}$, where all the street segments in the geometric map are represented by edge points. Then each particle $\mathbf{y}_i$ representing possible trajectory is plotted and also converted to an edge map $\mathcal{V}_i = \{\mathbf{v}_q\}$. The shape matching distance between $\mathcal{U}$ and $\mathcal{V}_i$ is given by the

average distance between each point $\mathbf{v}_q \in \mathcal{V}_i$ and its nearest edge in $\mathcal{U}$ [20]:

$$d_{\text{SM}}(\mathcal{V}_i,\mathcal{U}) = \frac{1}{n_{\text{SM}}} \sum_{\mathbf{v}_q \in \mathcal{V}_i} \min_{\mathbf{u}_p \in \mathcal{U}} |\mathbf{v}_q - \mathbf{u}_p| \quad (9)$$

where $n_{\text{SM}} = |\mathcal{V}_i|$ denotes the number of points in $\mathcal{V}_i$.

To measure in what degree the particle $\mathbf{y}_i$ accord with vehicle's trajectory, the *consistency* $\mu_{\text{SM}}(\mathbf{y}_i)$ is proposed. Since it is obvious that the *consistency* is negatively associated with the shape matching distance, one would think that the ideal distribution of $d_{\text{SM}}(\mathcal{V}_i,\mathcal{U})$ follows an exponential model [9]. However, as all geometric maps are modeled with the road center line, there is always an offset between vehicle's actual trajectory and map-represented roads. As a consequence, the maximum value of μ_{SM} appears when $d_{\text{SM}}(\mathcal{V}_i,\mathcal{U}) > 0$.

To find the fittest function to model the relation between shape matching distance and *consistency*, for a real road on the map, a set of pseudo trajectories is generated to simulate vehicle's possible trajectories. The shape matching distance can be calculated from (9) while the similarity of pseudo trajectories and the real trajectory is represented as *consistency*. The fitting result is demonstrated in Fig. 2. It is noted that the empirical distribution closely follows a Chi-squared distribution, and $\mu_{\text{SM}}(\mathbf{y}_i)$ can be expressed as $\mu_{\text{SM}}(\mathbf{y}_i) = \chi(d_{\text{SM}}(\mathcal{V}_i,\mathcal{U}),k)$, where $\chi(d_{\text{SM}}(\mathcal{V}_i,\mathcal{U}),k)$ denotes the estimated Chi-squared distribution with $k = 3$ degrees of freedom; $d_{\text{SM}}(\mathcal{V}_i,\mathcal{U})$ is the shape matching distance for particle $\mathbf{y}_i$.

E. Particle Resampling and Parameter Estimation

The weight of each particle before normalization can be obtained from

$$\tilde{w}(\mathbf{y}_i) = \mu_{\text{VO}}(\mathbf{y}_i)\mu_{\text{SM}}(\mathbf{y}_i) \quad (10)$$

Herein, N particles with top weights will be selected for location and parameter estimation. After resampling, weights need to be normalized for selected particles, and the normalization rule can be written as $w(\mathbf{y}_i) = \frac{\tilde{w}(\mathbf{y}_i)}{\sum_{i=1}^{N} \tilde{w}(\mathbf{y}_i)}$.

Estimating all parameters in UGD leads to higher degree polynomial equations with no algebraic solution. In this particular resampling step, as the raw VO result is quite accurate over a short distance, scaling factor plays a major role in selecting particles. To simplify the problem, we assume that after resampling, the expectation and variance of samples remain the same. In other words, shape matching filters out particles with improper scaling factors, and the distribution of raw MVO measurement does not change much after resampling process. The simplified parameter estimation problem can be formulated as: Given N samples $\mathbf{y}_i (i = 1, \cdots, N)$, estimate UGD parameters a, b with known parameters $E\mathbf{X}, \Sigma$.

According to Property 2-3, with the method of moments applied, the parameters can be estimated based on the following equations:

$$M_{1j} = \frac{\hat{a}+\hat{b}}{2} EX_j \quad (11)$$

$$M_{2j} = \frac{\hat{a}^2+\hat{b}^2-2\hat{a}\hat{b}}{12} E(X_j)^2 + \frac{\hat{a}^2+\hat{b}^2+\hat{a}\hat{b}}{3}\sigma_j^2 \quad (12)$$

where $M_{1j} = \frac{1}{N}\sum_{i=1}^{N} y_{ij}$ and $M_{2j} = \frac{1}{N-1}\sum_{i=1}^{N}(y_{ij} - M_{1j})^2$ denote the expectation and variance of sample particles respectively. The estimation results can be expressed as

$$\hat{a}_j = \frac{M_{1j}}{EX_j} - \sqrt{3\frac{M_{2j} - \frac{\sigma_j^2}{EX_j^2}M_{1j}^2}{{EX_j}^2 + \sigma_j^2}} \quad (13)$$

$$\hat{b}_j = \frac{M_{1j}}{EX_j} + \sqrt{3\frac{M_{2j} - \frac{\sigma_j^2}{EX_j^2}M_{1j}^2}{{EX_j}^2 + \sigma_j^2}} \quad (14)$$

where estimated scale range $[\hat{a}, \hat{b}]$ are used to generate particles obeying $UG(\hat{a}, \hat{b}, E\mathbf{X}, \Sigma)$ at the next iteration.

IV. Experimental Validation

Two sets of experiments are conducted and the results are shown to demonstrate the performance of the proposed approach. Although the original intention of this work is to reduce MVO measurement uncertainties which include drift and scale ambiguity, evaluation with SVO data is conducted first for the following reasons:

1) Localization with SVO can be regarded as a particular case with MVO when $a = b = 1$ is known. Evaluation with a known scale validates the effectiveness of parameter estimation scheme and is the prerequisite of the experiment with MVO.
2) Although no scale ambiguity exists in the stereo case, drift caused by calibration error, feature selection error, matching error may make the estimated scale fluctuate around the true value. The proposed approach provides a universally applicable framework to deal with these measurement uncertainties.

In the second experiment, we implement the proposed approach on MVO raw measurement. All the experiments are conducted on the challenging KITTI dataset. Some sequences of KITTI, either have a very short or straight driving trajectory, which makes the vehicle difficult to be localized. The so-called "fundamental ambiguities" problem is also discussed in [8].

A. Configurations

In all experiments, we use OSM as geometric maps. At the beginning, a rough positioning result from GPS is needed for map downloading. Then the map is revised by preserving drivable roads only. Before VO raw measurement is used for particle generation, an initialization step is required. As the shape matching needs a trajectory with enough geometric characteristics, the first batch of particles will not be generated until a minimum distance $d_{\text{min}} = 300$ meters and a turning angle $\alpha_{\text{min}} = \frac{\pi}{4}$ is reached. As for particle generation, 2000 particles are generated. The map sizes for Sequence 00, 02, 05, 08 and 09 are 725×836, 1159×1165, 730×583, 794×915, and 854×820 meters respectively. A low variance sampler is implemented in particle resampling. Parameters γ_0 and Σ_0 are set to $\gamma_0 = 1.02$, $\Sigma_0 = \text{diag}(2, 2)$ for SVO and $\gamma_0 = 1.005$, $\Sigma_0 = \text{diag}(0.1, 0.1)$ for MVO. During the experiments, most of the computation time is reserved for shape matching and

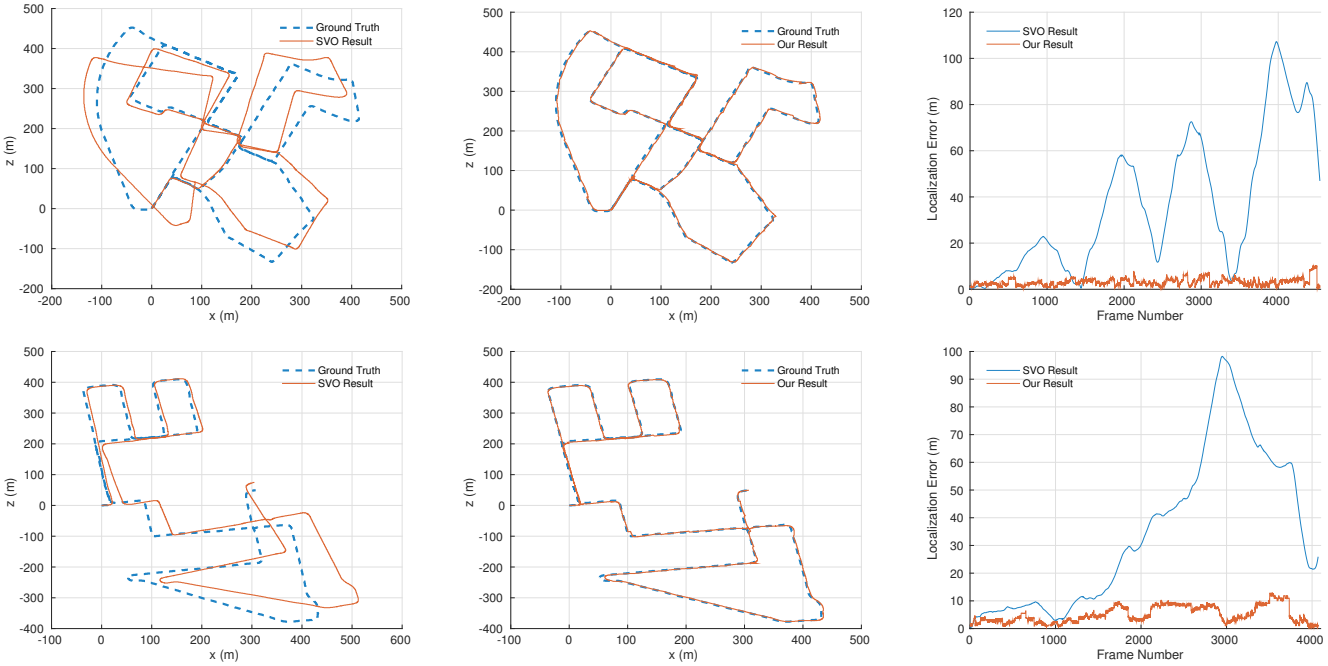

Fig. 3. Sequence 00 and 08 from VO benchmark of the KITTI dataset. Left: trajectory estimated from SVO and ground truth. Center: trajectory estimated from our map-assisted approach and ground truth. Right: localization error comparison between SVO and our method.

Table I. Quantitative comparison between the proposed approach and SVO localization.

Sequence	Travelling Distance (km)	Our Proposed			SVO			[8]
		Avg Error (m)	Std Dev (m)	Max Error (m)	Avg Error (m)	Std Dev (m)	Max Error (m)	Avg Error (m)
KITTI 00	3.72	3.14	1.16	10.28	37.11	29.63	107.23	2.1
KITTI 02	5.06	11.46	7.47	25.42	66.08	47.37	172.71	4.1
KITTI 05	2.20	4.14	2.12	8.89	14.40	13.30	45.60	2.6
KITTI 08	3.21	5.01	3.01	12.65	34.70	28.12	98.20	2.4
KITTI 09	1.70	5.75	2.29	11.84	16.63	34.91	9.72	4.2
Total	15.89	6.58	-	-	40.51	-	-	3.19

Table II. Quantitative comparison between the proposed approach and MVO localization.

Sequence	Travelling Distance (km)	Our Proposed			MVO			[8]
		Avg Error (m)	Std Dev (m)	Max Error (m)	Avg Error (m)	Std Dev (m)	Max Error (m)	Avg Error (m)
KITTI 00	3.72	4.11	2.40	13.60	10.00	6.62	26.07	15.6
KITTI 02	5.06	9.53	4.36	21.90	151.25	72.64	256.72	8.1
KITTI 05	2.20	6.28	5.34	22.90	7.21	7.20	23.31	5.6
KITTI 08	3.21	4.43	3.38	18.29	475.43	408.21	1300.20	45.2
KITTI 09	1.70	7.28	3.73	19.65	90.81	100.84	347.64	5.4
Total	15.89	6.54	-	-	160.07	-	-	16.72

particle resampling. With C++ implementation, it takes 1.5 ms per particle for shape matching and particle resampling under a mobile workstation with an i7-4710MQ processor. Since particles resampling at every time instant is unnecessary, it is promising to implement the proposed approach in real-time.

B. Localization with Stereo Visual Odometry

For SVO data, *libviso2* package is used for raw image processing [21]. It is assumed that the scale of SVO is unknown, and the initial range of scaling factor is set as $[0.5, 1.5]$. The estimated trajectory and localization error of our method and SVO are shown in Fig. 3 and Table I.

Qualitatively, the SVO drift increases every time the robot turns sharply as depicted in the left column of Fig. 3. The trajectories generated from the map-assisted approach can be found in the middle column of Fig. 3. In contrast, for the proposed approach, the more complicated trajectory is, the better localization results will be. It is not difficult to understand this phenomenon, as every time the robot changes its motion direction, more information is added to robot's trajectory such that the shape matching may further narrow down possible position of the robot. According to the quantitative result, pure SVO localization becomes increasingly unreliable with the growth of travel distance, mainly due to the accumulative drift. The proposed approach suppressed average localization error drastically from 40.51m to 6.58m. Besides, the proposed approach provides more robust positioning compared to pure SVO, as the standard deviations for all sequences are evidently lower.

C. Localization with Monocular Visual Odometry

Next, we evaluate the proposed approach using MVO raw translation vectors without scale information. State-of-the-art

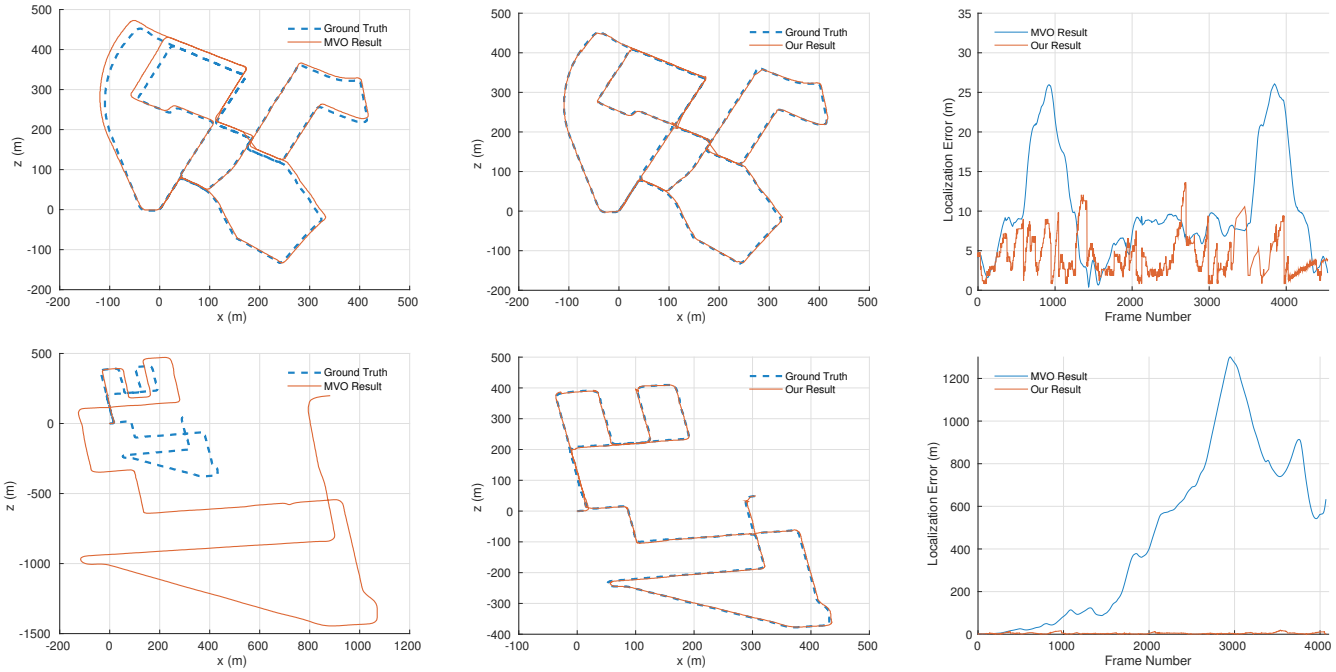

Fig. 4. Sequence 00 and 08 from VO benchmark of the KITTI dataset. Left: trajectory estimated from MVO and ground truth. Center: trajectory estimated from our map-assisted approach and ground truth. Right: localization error comparison between MVO and our method.

ORB-SLAM is used to provide us benchmark in monocular case [12]. Relying on loop closing and pose graph optimization, this system can achieve quite accurate motion estimation and map generation up to an absolute scale. When no loop closure occurs, standard local tracking and mapping are performed such that ORB-SLAM acts like a VO system. Without distinction, all the motion estimations from ORB-SLAM are called MVO in the experiment.

During the evaluations, we compare the performance of our method with MVO method in both conditions with and without loop closure. The trajectories estimated from the two methods are shown in Fig. 4. Since the absolute scale is unknown, to compare monocular motion estimation with the ground truth, we re-scale the MVO raw measurement by aligning the VO trajectories during initialization with the ground truth using a similarity transformation, as shown in the left column of Fig. 4. As can be seen, MVO performs much better on Sequence 00 than on Sequence 08 because loop closures exist only in Sequence 00, while the scale drift destroys motion estimation of Sequence 08. Despite the loop closure optimization, drift issue still exists in Sequence 00. By contrast, our method always performs well regardless of the existence of the loop closure. Most of the results are perfectly constrained near a road thanks to shape matching and parameter estimation scheme. The right column of Fig. 4 shows error curves of the two methods. It can be concluded that the proposed framework is valid for global monocular localization.

The quantitative results are listed in Table II. As KITTI 00 and KITTI 05 have loop closures, MVO results for these sequences are still acceptable. However, MVO results from KITTI 08 and KITTI 09 are not as optimistic as the first two sequences. Our results are much better, as the average localization error is 6.54m over the 15.89km run. Average errors in [8] are also listed for comparison. In summary, compared to [9], the proposed approach is effective for both SVO and MVO; Compared to [8], the proposed approach performs worse in SVO sequences with large drifts. It is noted that the exponential drift model may not represent measurement randomness properly when drifting error is large. It is promising to obtain less error based on a more precise drift model. Moreover, building an unified representative model with angular error considered may also help.

Besides error of VO and particle number, practical road conditions and resolution of the geometric map may influence positioning performance. Error from road conditions cannot be decreased because roads are modeled with center lines while real roads have different widths. Theoretically, error from geometric map precision can be eliminated if the resolution of the map is high enough. However, this is in contradictory with execution time and memory usage, as shape matching is needed for each possible trajectory to obtain particle weights.

D. Scale Estimation Results

The scale estimation results for Sequence 00 and 02 are shown in Fig. 5. Although it has not been proved that the proposed approach ensures the convergence of the estimated scale, for most cases, true scales can be recovered with road constraints. The estimated scale range is narrower when the trajectory contains more geometric information, i.e. turns. To avoid unnecessary estimation error from redundant constraints such as straight roads, scale estimation results based on straight roads are not used.

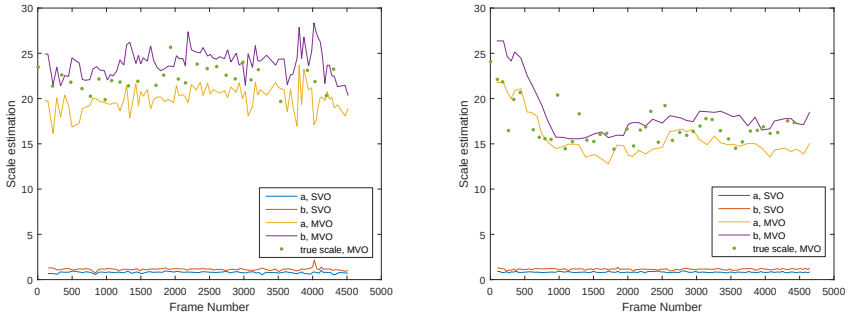

Fig. 5. Scale estimation results for Sequence 00 (left) and 02 (right).

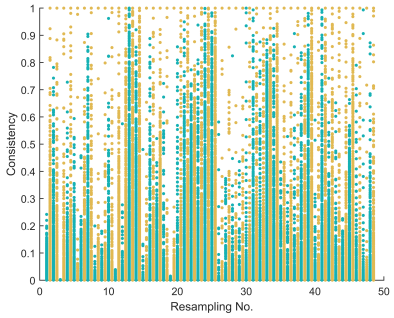

Fig. 6. *Consistency* range variation with (yellow) and without (green) weighting function balancing.

E. Weighting Function Balancing

As the weighting function (10) considers both *saliency* and *consistency*, it is necessary to examine the influence on algorithm performance from these two terms. For *saliency* set $\{\mu_{\text{VO}}(\mathbf{y}_i)\}$, it is obvious that $\min\lim_{N\to\infty}\{\mu_{\text{VO}}(\mathbf{y}_i)\} = 0$ and $\max\lim_{N\to\infty}\{\mu_{\text{VO}}(\mathbf{y}_i)\} = 1$ where N denotes particle number. However, this is not the case for *consistency* even after normalization, as the shape matching distance of all particles may not cover a specific interval, especially when the estimated trajectory is off the road. This unbalanced weighting strategy may cause *consistency* less significant and decrease the algorithm robustness. In practice, the following affine transform is used for *consistency* such that distribution shape remains unchanged:

$$\mu'_{\text{SM}}(\mathbf{y}_i) = \frac{\mu_{\text{SM}}(\mathbf{y}_i) - \min\{\mu_{\text{SM}}(\mathbf{y}_i)\}}{\max\{\mu_{\text{SM}}(\mathbf{y}_i)\} - \min\{\mu_{\text{SM}}(\mathbf{y}_i)\}} \tag{15}$$

Fig. 6 shows *consistency* range with and without (15) on MVO Sequence 02. The affine transform ensures $\min\lim_{N\to\infty}\{\mu_{\text{SM}}(\mathbf{y}_i)\} = 0$ and $\max\lim_{N\to\infty}\{\mu_{\text{SM}}(\mathbf{y}_i)\} = 1$. It is noted that the weighting function after balancing avoids the possible significance degradation problem of *consistency*.

V. Conclusions

We have presented a geometric map-assisted localization approach, which utilizes the dead-reckoning trajectory from VO and a geometric map to obtain position estimation. We have proposed a new probability distribution to describe the measurement uncertainties in VO, and a framework to incorporate the motion information from VO and the measurement information from shape matching together. Experiments have shown that with road constraints, the localization error has been significantly reduced compared to an advanced MVO method. Currently, we are conducting more experiments with our dataset, and future work involves the real-time implementation of the positioning approach.

References

[1] J. A. Castellanos and J. D. Tardos, *Mobile robot localization and map building: A multisensor fusion approach*. Springer Science & Business Media, 2012.

[2] Y. Cui and S. S. Ge, "Autonomous vehicle positioning with GPS in urban canyon environments," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 1, pp. 15–25, 2003.

[3] E. Kaplan and C. Hegarty, *Understanding GPS: principles and applications*. Artech house, 2005.

[4] M. Milford, E. Vig, W. Scheirer, and D. Cox, "Vision-based simultaneous localization and mapping in changing outdoor environments," *Journal of Field Robotics*, vol. 31, no. 5, pp. 780–802, 2014.

[5] H. Lim, S. N. Sinha, M. F. Cohen, M. Uyttendaele, and H. J. Kim, "Real-time monocular image-based 6-DoF localization," *The International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 476–492, 2015.

[6] M. Maimone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the mars exploration rovers," *Journal of Field Robotics*, vol. 24, no. 3, pp. 169–186, 2007.

[7] B. M. Kitt, J. Rehder, A. D. Chambers, M. Schonbein, H. Lategahn, and S. Singh, "Monocular visual odometry using a planar road model to solve scale ambiguity," 2011.

[8] M. A. Brubaker, A. Geiger, and R. Urtasun, "Lost! leveraging the crowd for probabilistic visual self-localization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 3057–3064.

[9] G. Floros, B. van der Zander, and B. Leibe, "Openstreetslam: Global vehicle localization using openstreetmaps," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1054–1059.

[10] P. Merriaux, Y. Dupuis, P. Vasseur, and X. Savatier, "Fast and robust vehicle positioning on graph-based representation of drivable maps," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 2787–2793.

[11] H. Strasdat, J. Montiel, and A. J. Davison, "Scale drift-aware large scale monocular SLAM," *Robotics: Science and Systems VI*, 2010.

[12] R. Mur-Artal, J. Montiel, and J. D. Tardós, "ORB-SLAM: a versatile and accurate monocular SLAM system," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[13] G. Nützi, S. Weiss, D. Scaramuzza, and R. Siegwart, "Fusion of imu and vision for absolute scale estimation in monocular slam," *Journal of Intelligent & Robotic Systems*, vol. 61, no. 1-4, pp. 287–299, 2011.

[14] S. Song and M. Chandraker, "Robust scale estimation in real-time monocular SFM for autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1566–1573.

[15] 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.

[16] T. Senlet and A. Elgammal, "Satellite image based precise robot localization on sidewalks," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 2647–2653.

[17] D. Scaramuzza and F. Fraundorfer, "Visual odometry [tutorial]," *IEEE Robotics & Automation Magazine*, vol. 18, no. 4, pp. 80–92, 2011.

[18] V. K. Rohatgi and A. M. E. Saleh, *An introduction to probability and statistics*. John Wiley & Sons, 2015.

[19] R. Jiang, R. Klette, and S. Wang, "Modeling of unbounded long-range drift in visual odometry," in *Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on*. IEEE, 2010, pp. 121–126.

[20] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching," DTIC Document, Tech. Rep., 1977.

[21] A. Geiger, J. Ziegler, and C. Stiller, "Stereoscan: Dense 3d reconstruction in real-time," in *Intelligent Vehicles Symposium (IV)*, 2011.