2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Daejeon Convention Center
October 9-14, 2016, Daejeon, Korea

# Achievement of Localization System for Humanoid Robots with Virtual Horizontal Scan Relative to Improved Odometry Fusing Internal Sensors and Visual Information

Iori Kumagai, Ryohei Ueda, Fumihito Sugai, Shunichi Nozawa,
Yohei Kakiuchi, Kei Okada and Masayuki Inaba[1]

***Abstract*— To achieve tasks in unknown environments with high reliability, highly accurate localization during task execution is necessary for humanoid robots. In this paper, we discuss a localization system which can be applied to a humanoid robot when executing tasks in the real world. During such tasks, humanoid robots typically do not possess a referential to a constant horizontal plane which can in turn be used as part of fast and cost efficient localization methods. We solve this problem by first computing an improved odometry estimate through fusing visual odometry, feedforward commands from gait generator and orientation from inertia sensors. This estimate is used to generate a 3D point cloud from the accumulation of successive laser scans and such point cloud is then properly sliced to create a constant height horizontal virtual scan. Finally, this slice is used as an observation base and fed to a 2D SLAM method. The fusion process uses a velocity error model to achieve greater accuracy, which parameters are measured on the real robot.**

**We evaluate our localization system in a real world task execution experiment using the JAXON robot and show how our system can be used as a practical solution for humanoid robots localization during complex tasks execution processes.**

## I. Introduction

Stimulated by the recent DARPA Robotics Challenge, robots for disaster response scenarios have been greatly attracting attention and efforts. However, construction methods of a reliable robot localization system for disaster like unknown environments are still an active research topic. Especially, humanoid robots have a problem that they do not possess a referential to a constant plane for efficient localization methods during tasks where several disturbances occur such as change of their posture in a manipulation or oscillations in locomotion. It is also not recommended for humanoid robots to suspend its tasks to build an environmental map or run the recognition process to compensate localization errors in consideration for the operation efficiency.

Therefore, in this paper, we discuss a localization system for a humaonid robot which works in unknown environments as shown in Fig. 1. We propose a method which first consists in computing an improved odometry estimate through the fusion between visual odometry, feedforward commands from gait generator and orientation from inertia sensors. We used this estimate to compute a 3D point cloud by accumulating successive 2D laser scans acquired from the rotating laser scanner. The point cloud is successively sliced to create a constant height horizontal virtual scan. This computed slice is used as an observation base and fed to a 2D SLAM method. We argue that this method gives good results even when the process becomes more error prone as the robot carries on various tasks in the 3D world, even if the robot does not posess a referential to a constant plane during the tasks. Although the strict error model of a humanoid robot during task execution is complex to derive due to non-linearities, we propose to improve the accuracy of the odometry using an approximate velocity error model measured on the real robot. Our localization method improves the autonomy of humanoid robots as it is experimentally shown in the case of our humanoid robot, JAXON in Fig. 2.

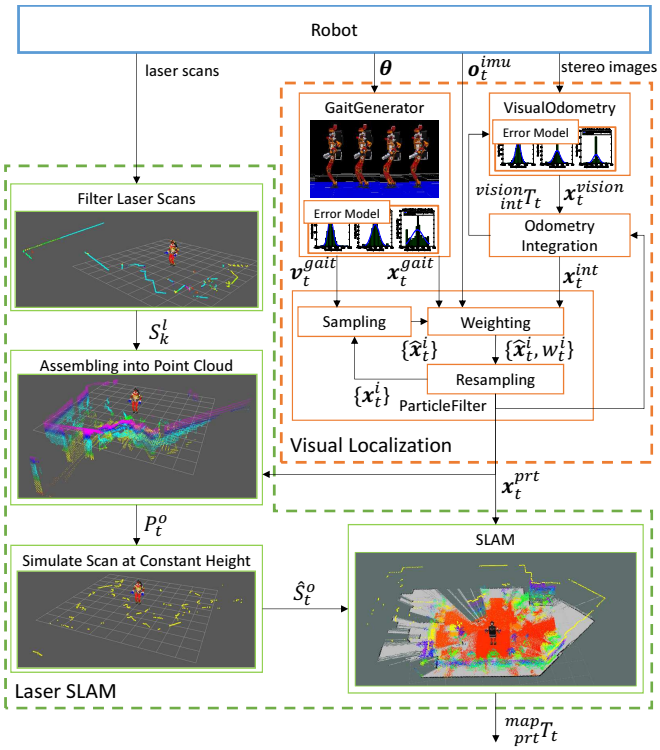


Fig. 1. Overview of the proposed localization system for a humanoid robot

## II. Localization Methods for Hunmanoid Robots

### A. Related works

Many methods using visual odometry with cameras mounted on a humanoid robot have been proposed in the

[1]Graduate School of Information Science and Technology, The University of Tokyo 7-3-1 Hongo, Bunkyo-ku, 113-8656 Tokyo, Japan {iori, ueda, nozawa, youhei, k-okada, inaba}@jsk.t.u-tokyo.ac.jp

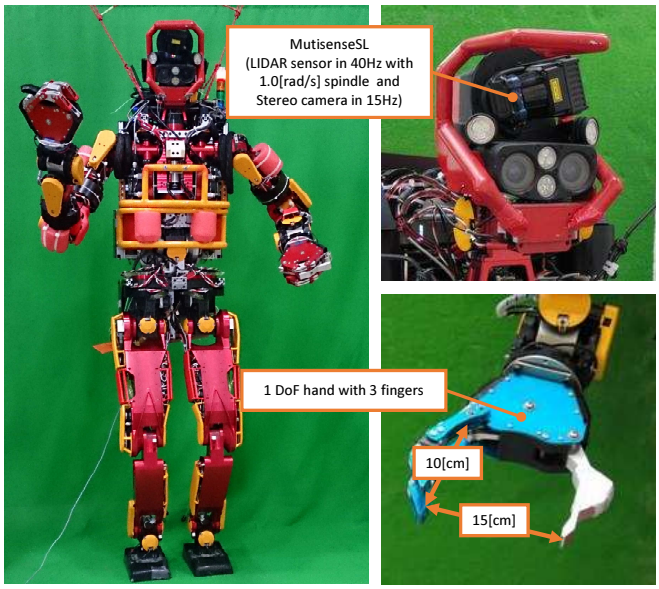


Fig. 2. Overview of the high-power humanoid robot JAXON [1]

past [2], [3]. If visual odometry based results usually provide with reliable solutions, they can suffer from some limitations such as drift during long time experiments or fast motions. In order to alleviate these effects, it is common to fuse visual odometry results with information from internal sensors or target postures from gait generator through filtering methods such as an Extended Kalman Filter [4], [5]. Such association is effective in order to improve the accuracy and robustness of the localization but such improvement still depends on the sensor in use.

Therefore, using additional laser based solutions can provide with interesting complementary properties. During localization with laser scaners, it is common to derive a maximum likelihood solution using a reference map [6], [7]. However, more challenges can arise when using 2D laser scans on humanoid robots which don't have constant 2D sensor frame during task execution. Tellez et al [8] achieved simultaneous localization and mapping in the humanoid robot Reem-B using laser sensors mounted on its feet. However, this solution did not consider tasks where height changes like terrain walking and laser scans on the feet are apt to moving objects like humans. Fallon et al [9] achieved drift-free localization by fusing kinematics of the robot, orientations from inertial measurement unit and a 3D environmental map generated from laser scans. This solution is very effective but it requires to wait to build environmental map before executing tasks. They also leave integrating visual informations as a future work.

Localization methods without external sensors like cameras or laser scans was also proposed in the littterature. Masuya et al [10] achieved accurate dead reckoning without external sensors fusing postures from forward kinematics and double integration of acceleration from inertial measurement unit and compensating rotation around the contact point to the ground in the supporting foot. However, the localization method without external sensors requires high accuracy in the model of the robot such as mass property or link length and high rigidity to the robot. We consider that it is essential to compensate localization errors by external sensors since modeling errors and deflection of links account for a large percentage of errors.

### B. Overview of the proposed method

Based on the above discussion, we propose the localization system for a humanoid robot. Our localization system enables a humanoid robot to localize itself with high accuracy and frequency in an unknown environment by fusing joint angles, orientations from inertial measurement unit and the result of visual odometry into 2D laser based SLAM. The proposed localization method assumes a humanoid robot which executes both locomotion and tasks such as a manipulation in an unknown envorinment.

First, we obtain improved odometry fusing the target posture and velocity from the gait generator, the orientation from inertial measurement unit and the estimation result of the visual odometry using a particle filter. The fusion reduces the errors from the difference of the robot model, effects of slippage or oscillations in walking and drift in the estimation of the visual odometry. It is generally difficult to design an exact error model of the robot motion due to the nonlinear error elements such as stabilization control, slippage of the feet and rigidity of the links. In our case, we use an approximate estimate using a velocity error model assuming a normal distribution and measuring its parameters on the real robot.

Then, we adopt a 2D laser SLAM with low computational costs for more accurate localization. It is difficult to apply exisiting 2D SLAM methods directly to a humanoid robot since it does not provide a constant horizontal plane reference during task execution. In order to solve this problem, we transform the laser scans from sensor coordinates to a robot pose independant coordinates estimated by the improved odometry. These laser scans are accumulated into a point-cloud to subsequently generate a virtual 2D scan by slicing at an appropriate horizontal constant plane.

## III. Model Identification with Covariance

### A. Visual odometry

Visual odometry is a typical method for localization in an unknown environment. There are several types of visual odometry such as methods using monocular cameras [11], [12], fusing images and depth information from RGB-D cameras [13] or using sequential stereo images [14]. In this paper, we adopt a visual odometry solution through sequential stereo images [14] since monocular cameras based methods have scale ambiguity and fusing RGB-D information requires increased computational costs.

Visual odometry provides estimates of the camera posture, whereas our objective is to estimate the posture of the base link of the robot. A humanoid robot does not have a constant transformation from a camera to its base link $T_b^c$. Therefore, we calculate a homogeneous transformation matrix for the posture $H_w^b$ using that of the camera from visual odometry

$H_w^c$ and transformation from its base link to the camera by forward kinematics $T_b^c$ as Eq.1.

$$H_w^b = H_w^c (T_b^c)^{-1} \tag{1}$$

### B. Localization error model for a humanoid robot

It remains difficult to design strictly correct error models for humanoid locomotion based on the localization algorithm because there are various error sources such as joint angle measurement errors, posture difference caused by stabilization control and position errors caused by slippage in addition to the errors which can be estimated from the localization algorithm itself. In this paper, we propose a localization error model for humanoid robots based on the measured velocities.

*1) Derivation of an analytical covariance:* We assume that state $\boldsymbol{x}(t)$ is updated according to Eq.2.

$$\boldsymbol{x}(t + \Delta t) = f(\boldsymbol{x}(t), \dot{\boldsymbol{x}}(t)) \tag{2}$$

The state $\boldsymbol{x}(t)$ contains the error $\Delta\boldsymbol{x}(t)$ with respect to the true state $\hat{\boldsymbol{x}}(t)$. We derive the update formula for the error $\Delta\boldsymbol{x}(t + \Delta t)$ at the time $t + \Delta t$ by applying primary Taylor expansion to Eq.2 around true states $\hat{\boldsymbol{x}}(t)$, $\hat{\dot{\boldsymbol{x}}}(t)$. We omit the assignment notations for the partial differentiations.

$$\boldsymbol{x}(t + \Delta t) = f(\hat{\boldsymbol{x}}(t), \hat{\dot{\boldsymbol{x}}}(t)) + \frac{\partial f}{\partial \boldsymbol{x}} \Delta\boldsymbol{x}(t) + \frac{\partial f}{\partial \dot{\boldsymbol{x}}} \Delta\dot{\boldsymbol{x}}(t) \tag{3}$$

The covariance of the state $\Sigma_x(t)$ can be expressed by Eq.4 with $\Delta\boldsymbol{x}(t)$ according to the definition of a covariance.

$$\Sigma_x(t) = E[\Delta\boldsymbol{x}(t) \Delta\boldsymbol{x}^T(t)] \tag{4}$$

Assuming that $\Delta\boldsymbol{x}(t)$ and $\Delta\dot{\boldsymbol{x}}(t)$ are independent, we derive Eq.5 from Eq.4 based on $E[\Delta\boldsymbol{x}(t)\Delta\dot{\boldsymbol{x}}^T(t)] = E[\Delta\dot{\boldsymbol{x}}(t)\Delta\boldsymbol{x}^T(t)] = 0$.

$$\Sigma_x(t + \Delta t) = \frac{\partial f}{\partial \boldsymbol{x}} \Sigma_x(t) \frac{\partial f}{\partial \boldsymbol{x}}^T + \frac{\partial f}{\partial \dot{\boldsymbol{x}}} \Sigma_{\dot{x}}(t) \frac{\partial f}{\partial \dot{\boldsymbol{x}}}^T \tag{5}$$

We use Eq.5 as an update formula for analytical covariance from this [15].

*2) Formulation of a localization error model for a humanoid robot:* The update formula of states $\boldsymbol{x}, \dot{\boldsymbol{x}} \in \mathbb{R}^6$ for a humanoid robot can be expressed as Eq.6 assuming that holonomic movement model can be applied in a infinitesimal time $\Delta t$ and all elements are independent.

$$\boldsymbol{x}(t + \Delta t) = \boldsymbol{x}(t) + \dot{\boldsymbol{x}}(t) \Delta t \tag{6}$$

Therefore, we can derive the covariance of pose $\Sigma_x(t)$ for a biped robot analytically using Eq.7 according to the general update formula in Eq.5.

$$\Sigma_x(t + \Delta t) = \Sigma_x(t) + \Sigma_{\dot{x}}(t) \Delta t^2 \tag{7}$$

We adopt the error model for $\Sigma_{\dot{x}}(t)$ in which all elements of the linear and angular velocities follow normal distributions independently. We define the means $\boldsymbol{\mu} = (\mu_x \cdots \mu_\gamma)^T$ and variances $\sigma_x^2 \cdots \sigma_\gamma^2$ as constant values for each element based on the measured velocity errors in a real robot and use a normal distribution $\mathcal{N}(\dot{\boldsymbol{x}}; \tilde{\dot{\boldsymbol{x}}} + \boldsymbol{\mu}, \Sigma_{\dot{x}}(t))$ as the velocity error model of $\tilde{\dot{\boldsymbol{x}}}$. $\Sigma_{\dot{x}}(t)$ can be expressed as $\Sigma_{\dot{x}}(t) = \mathrm{diag}\left(\sigma_x^2, \cdots, \sigma_\gamma^2\right)$ because we assume all elements are independent.

*3) Identification of velocity error models in a real robot:* We identify the velocity error models based on measured velocity errors on a real robot. We sampled errors of linear velocity of x and y axis direction and angular velocity around z axis comparing ground truth velocities based on checkerboard recognition and velocities from gait generator and visual odometry. Our humanoid robot JAXON [1] was walking in accordance with the manually given trajectory by feedforward control as shown in the left of Fig. 3.

We assume a normal distribution for the measured errors and estimate means and standard deviations using maximum likelihood method. The results are shown on the right of Fig. 3. Although usually high errors can be observed around zero since the stopping state of the robot is more stable than the moving state, it can be seen that the rest of the measured errors generally follow normal distributions. We selected the parameters for the velocity error models of JAXON as shown in TABLE. I. These parameters have larger standard deviations than measured ones for practical purpose.

TABLE I
VELOCITY ERROR MODEL PARAMETERS FOR JAXON

| Direction | Variable | Gait Generator | Visual Odometry |
|---|---|---|---|
| x translation | $\mu$ | 0.006[m] | 0.0[m] |
| | $\sigma$ | 0.075[m] | 0.05[m] |
| y translation | $\mu$ | 0.001[m] | -0.002[m] |
| | $\sigma$ | 0.15[m] | 0.1[m] |
| z translation | $\mu$ | 0.0[m] | 0.0[m] |
| | $\sigma$ | 0.015[m] | 0.030[m] |
| roll rotation | $\mu$ | 0.0[rad] | 0.0[rad] |
| | $\sigma$ | 0.001[rad] | 0.3[rad] |
| pitch rotation | $\mu$ | 0.001[rad] | 0.3[rad] |
| | $\sigma$ | 0.15[rad] | 0.06[rad] |
| yaw rotation | $\mu$ | 0.0[rad] | -0.002[rad] |
| | $\sigma$ | 0.15[rad] | 0.06[rad] |

### C. Estimated posture feedback with integration of normal distributions

We use integration of normal distributions to compensate for a drift in the visual odometry. In this section, we discuss how to estimate the maximum likelihood distribution $\boldsymbol{x}_t^{int}$ using the posture distributions from visual odometry $\boldsymbol{x}_t^{vision}$ and posture and velocity distributions from feedback odometry $\boldsymbol{x}_{t-\Delta t}^{fb}$, $\boldsymbol{v}_{t-\Delta t}^{fb}$.

*1) Maximum likelihood estimation with independent measurement:* When two normal distributions $\boldsymbol{z}_1(t_1) \sim \mathcal{N}_1(\boldsymbol{x}; \boldsymbol{\mu}_1, \Sigma_1)$ and $\boldsymbol{z}_2(t_2) \sim \mathcal{N}_2(\boldsymbol{x}; \boldsymbol{\mu}_2, \Sigma_2)$ are obtained independently for the state $\boldsymbol{x}(t)$, we can determine the maximum likelihood probability distribution $\mathcal{N}_{opt}(\boldsymbol{x}; \boldsymbol{\mu}_{opt}, \Sigma_{opt})$ for the current state $\boldsymbol{x}(t)$. Generally, the conditional probability $p(x|z_1, z_2)$ of the state $x$ can be calculated as Eq.8 based on Bayes rule.

$$p(x|z_1, z_2) = \frac{p(x, z_1, z_2)}{p(z_1, z_2)} = p(x|z_1) p(x|z_2) \tag{8}$$

We can determine the mean and standard deviation of the maximum likelihood probability distribution as Eq.9 and

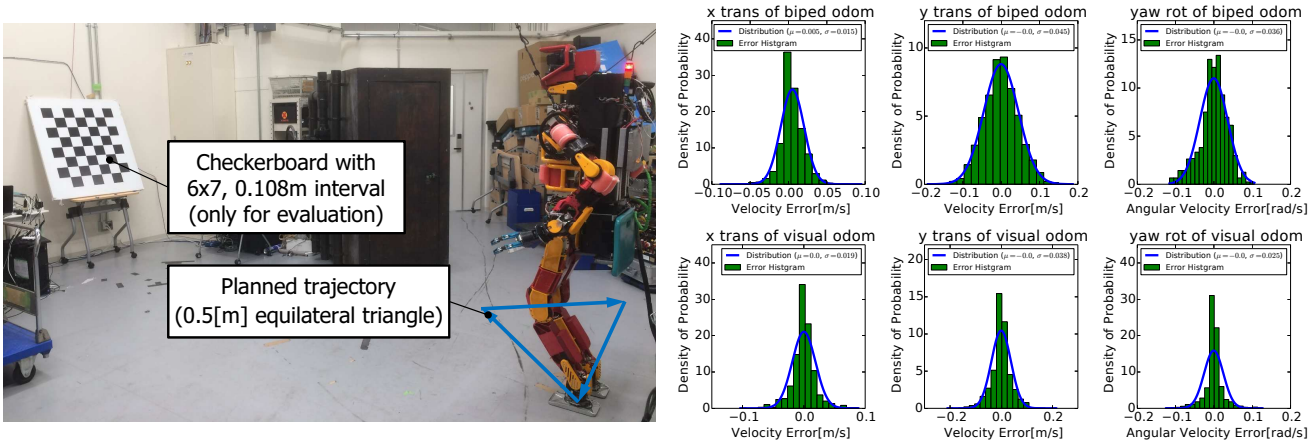


Fig. 3. Left: Experimental environment for velocity model sampling, Right: Measured error histograms and estimated Gaussian probability distributions of JAXON [1]. Green histograms show measured velocity errors and blue lines show fitted Gaussian distributions. Upper graphs are result of velocity errors from walking pattern generator and lower ones are from visual odometry.

Eq.10 by solving Eq.8 for the exponent part [16].

$$\Sigma_{opt} = \Sigma_1 - \Sigma_1 \left(\Sigma_1 + \Sigma_2\right)^{-1} \Sigma_1 \quad (9)$$

$$\boldsymbol{\mu}_{opt} = \mu_1 + \Sigma_1 \left(\Sigma_1 + \Sigma_2\right)^{-1} \left(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1\right) \quad (10)$$

*2) Feedback for odometry distribution:* The integration method for maximum likelihood probability distribution discussed in Subsection III-C.1 assumes that two normal distributions are obtained at the same time. Therefore, we update the feedback state using the velocity integration as Eq.11.

$$\bar{\boldsymbol{x}}_t^{fb} = \boldsymbol{x}_{t-\Delta t}^{fb} + \boldsymbol{v}_{t-\Delta t}^{fb} \Delta t \quad (11)$$

We derive the maximum likelihood odometry distribution $\boldsymbol{x}_t^{int}$ by applying Eq.10 and Eq.9 to $\boldsymbol{x}_t^{vision}$ and $\bar{\boldsymbol{x}}_t^{fb}$ We also use the transformation ${}^{vision}_{int}T_t$ from $\boldsymbol{x}_t^{vision}$ to $\boldsymbol{x}_t^{int}$ as the offset for visual odometry.

*3) Feedback interval:* Odometry feedback can compensate the drift in visual odometry estimation but a time delay can occur in the integrated odometry. Moreover, excessive feedback can cause degeneration of the odometry distribution when estimation frequency of the visual odmetry is slow compared to the feedback odometry. Therefore, we enable odometry feedback when the standard deviation of the visual odometry exceeds a threshold $\sigma_{th}$ or the result of feedback odometry is not in the range of $3\sigma$ of the visual odometry. In this research, we set $\sigma_{th}$ to be 1.0[m] for the robot's position and 1.0[rad] for the robot's orientation.

## IV. Improved Odometry by Particle Filter

Extended Kalman Filter has been widely adopted as a sensor fusion method in the literature [5]. Extended Kalman Filter assumes that the state transition and observation models of the system are linear and tends to have large errors in systems that have probability density functions with large covariance or local non-linearity [17]. While the state transition model given in Eq.6 is indeed linear, the observation model which uses visual odometry has large covariance and a non-linear relation to the state as discussed in Section III. Therefore, we adopt a particle filter solution counterpart which can bear a non-linear system.

### A. Sampling

We take $N$ samples of velocities $\hat{\boldsymbol{v}}_t^i$ from target velocities from gait generator according to Eq.12.

$$\hat{\boldsymbol{v}}_t^i \leftarrow \mathcal{N}_t^{\boldsymbol{v}^{gait}}(\boldsymbol{v}; \boldsymbol{\mu}_t^{\boldsymbol{v}^{gait}}, \Sigma_t^{\boldsymbol{v}^{gait}}) \quad (12)$$

We update the state of a particle $\hat{\boldsymbol{x}}_t^i$ with Eq.13 using the sampled velocity $\hat{\boldsymbol{v}}_t^i$.

$$\hat{\boldsymbol{x}}_t^i = \boldsymbol{x}_{t-1}^i + \hat{\boldsymbol{v}}_t^i \Delta t \quad (13)$$

### B. Weighting

We determine the weight $w_t^i$ for the particle $\hat{\boldsymbol{x}}_t^i$. Generally, the weight of a particle is calculated as Eq.14 based on the posterior probability of the observation $\boldsymbol{z}_t^i$ and normalized.

$$w_t^i \sim p(\boldsymbol{z}_t | \hat{\boldsymbol{x}}_t^i) \quad (14)$$

In out method, we use the resulting pose distribution of the visual odometry $\boldsymbol{x}_t^{int}$ as the observation. However, given the nature of our motions, the resulting visual odometry has large drift in vertical position and orientation around roll and pitch axis. Therefore, we obtain stable odometry estimation by fusing the target position at z-axis from the gait generator $z_t^{gait}$ and roll and pitch orientations from inertial measurement unit $\boldsymbol{o}_t^{imu}$ into the resulting distribution of visual odometry. Assuming that $\boldsymbol{x}_t^{int}$, $z_t^{gait}$, and $\boldsymbol{o}_t^{imu}$ are independent, we can derive the posterior probability $p(\boldsymbol{z}_t | \hat{\boldsymbol{x}}_t^i)$ as Eq.15 [7].

$$p(\boldsymbol{z}_t | \hat{\boldsymbol{x}}_t^i) = p(\boldsymbol{x}_t^{int} | \hat{\boldsymbol{x}}_t^i) p(z_t^{gait} | \hat{\boldsymbol{x}}_t^i) p(\boldsymbol{o}_t^{imu} | \hat{\boldsymbol{x}}_t^i) \quad (15)$$

*1) Observation model based on visual odometry:* When the pose distribution from the visual odometry $\boldsymbol{x}_t^{int} \sim \mathcal{N}_t^{\boldsymbol{x}^{int}}(\boldsymbol{x}; \boldsymbol{\mu}_t^{\boldsymbol{x}^{int}}, \Sigma_t^{\boldsymbol{x}^{int}})$ is observed, we can derive its posterior probability density function as Eq.16 with the normalization constant $\eta$.

$$p(\boldsymbol{x} | \boldsymbol{x}_t^{int}) = \eta e^{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_t^{\boldsymbol{x}^{int}})^T (\Sigma_t^{\boldsymbol{x}^{int}})^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_t^{\boldsymbol{x}^{int}})} \quad (16)$$

Assuming that the particle $\hat{\boldsymbol{x}}_t^i$ can uniformly span the space, the observation model can be derived as Eq.17 from Bayes rule.

$$p(\boldsymbol{x}_t^{int}|\hat{\boldsymbol{x}}_t^i) \sim e^{-\frac{1}{2}(\hat{\boldsymbol{x}}_t^i-\boldsymbol{\mu}_t^{\boldsymbol{x}^{int}})^T(\Sigma_t^{\boldsymbol{x}^{int}})^{-1}(\hat{\boldsymbol{x}}_t^i-\boldsymbol{\mu}_t^{\boldsymbol{x}^{int}})} \quad (17)$$

*2) Observation model for virtical direction:* We use the z-axis element of the target position from the gait generator $z_t^{gait}$ for the vertical direction. We can derive observation model as Eq.18 for the z-axis element of the particle $\hat{z}_t^i$ using its standard deviation $\sigma_t^z$ [7].

$$p(z_t^{gait}|\hat{\boldsymbol{x}}_t^i) \sim e^{-\frac{(z_t^{gait}-\hat{z}_t^i)^2}{2(\sigma_t^z)^2}} \quad (18)$$

*3) Observation model for roll and pitch orientation:* We use the orientation from the inertial measurement unit $\boldsymbol{o}_t^{imu} = (\phi_t^{imu}, \psi_t^{imu}, \theta_t^{imu})$ for the orientations around roll and pitch axis. Assuming that the probability distributions are independent, we can derive the observation model as Eq.19 using their standard deviation $\sigma_t^{\phi}$, $\sigma_t^{\psi}$ [7].

$$p(\boldsymbol{o}_t^{imu}|\hat{\boldsymbol{x}}_t^i) \sim e^{-\frac{(\phi_t^{imu}-\hat{\phi}_t^i)^2}{2(\sigma_t^{\phi})^2}-\frac{(\psi_t^{imu}-\hat{\psi}_t^i)^2}{2(\sigma_t^{\psi})^2}} \quad (19)$$

### C. Resampling

We adopt low variance resampling in consideration for computational cost [17]. A new particle $\boldsymbol{x}_t^i$ is selected from $\{\hat{\boldsymbol{x}}_t^i\}$ using the total number of particles $N$ and a random number $r$ selected from $[0; N^{-1}]$.

$$i = \underset{j}{\operatorname{argmin}} \sum_{n=1}^{j} w_t^n \geq r + (k-1)N^{-1} \quad (20)$$

We calculate weighted mean $\boldsymbol{\mu}_t^{\boldsymbol{x}^{prt}}$ and covariance $\Sigma_t^{\boldsymbol{x}^{prt}}$ from obtained particles and weights $\{\boldsymbol{x}_t^i, w_t^i\}$ assuming that the result distribution follows a normal distribution and estimate current pose distribution $\mathcal{N}_t^{\boldsymbol{x}^{prt}}(\boldsymbol{x}; \boldsymbol{\mu}_t^{\boldsymbol{x}^{prt}}, \Sigma_t^{\boldsymbol{x}^{prt}})$.

### D. Evaluation of localization accuracy in simulation

We evaluated the proposed improved odometry with JAXON walking according to the given feedforward trajectory in simulation environment and comparing the results with the ground truth, odometry from the gait generator, raw visual odometry, visual odometry with proposed feedback in Section III and result of Extended Kalman Filter with odometry from gait generator and raw visual odometry. We added the Gaussian noise following the normal distributions shown in the right of Fig. 3 to x and y positions and yaw orientation. We set the number of the particle to be $N$ =100 and execution rate to be 40[Hz] based on the frequency of the laser scanner.

The result is shown in Fig. 4. The proposed odometry can estimate the position in x and y direction with the accuracy within approximately 0.1[m] and orientation around yaw axis with the accuracy within approximately 2[deg] over the task of walking 40[m]. The raw visual odometry suffers drift in the z position and around roll and pitch orientations. However, it was reduced by odometry feedback and particle filters to approximately 1[cm] for position and 1[deg] for orientation.

## V. Pointcloud Accumulation and 2D SLAM

We have discussed the method for improving odometry by fusing internal sensors and stereo camera so far, but there are limitations to the accuracy improvement in it because of its vulnerability to drifts and oscillations in walking or task execution. Therefore, we adopt a SLAM method with laser scans to more accurately localize the robot.

### A. 2D SLAM in a humanoid robot

Many SLAM methods have been studied so far in the field of computer vision and robotics. Grisetti et al [18] proposed the efficient 2D SLAM method called Gmapping, which uses a Rao-Blackwellized Particle Filter preventing particle depletion by adoptive resampling and computational cost by considering recent observation. Recently, SLAM methods for three dimentional space are proposed, such as stereo vision based rtabmap [19] or lidar based V-LOAM [20].

In a real world, a humanoid robot has six DoF for its posture but we can obtain accurate enough orientation around roll and pitch axis from an inertial measurement unit, and calculate the base link's height from forward kinematics using joint angles when the robot's feet are in contact with the ground [7]. Moreover, 3D SLAM methods require high computational costs, which can be seen from the fact that average CPU usage of Gmapping is 4.57[%] while that of rtabmap is 22.03[%] on the on-board computer of JAXON (Intel(R) Core(TM) i7-4770R CPU @ 3.20GHz). Therefore, we consider 2D SLAM to be sufficient for localization of a humanoid robot and adopt Gmapping to correct the position in x and y directions and the orientation around yaw axis.

### B. Generation of a virtual scan from the accumulated point cloud

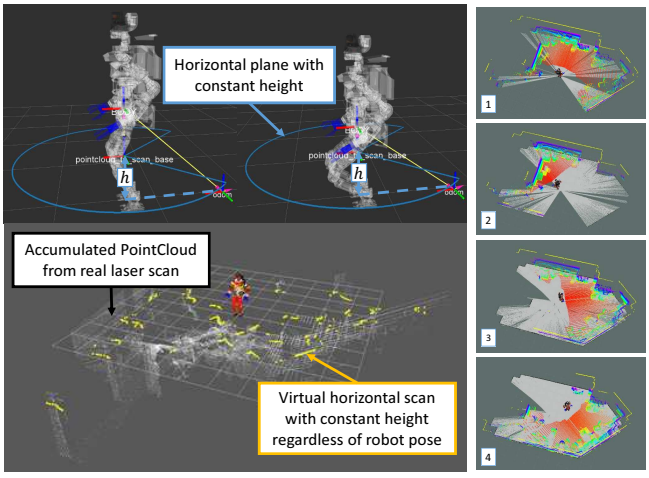


Fig. 5. Left: A viatual constant height scan frame and simulated horizontal scan from accumulated pointcloud, Right: Simultaneous localization and mapping by Gmapping with proposed simulated scan in unknown environment

Generally, 2D SLAM methods assume the constant height laser scanner from the ground. However, humanoid robots which execute various tasks do not have constant sensor frame with respect to the ground. Therefore, we generate a

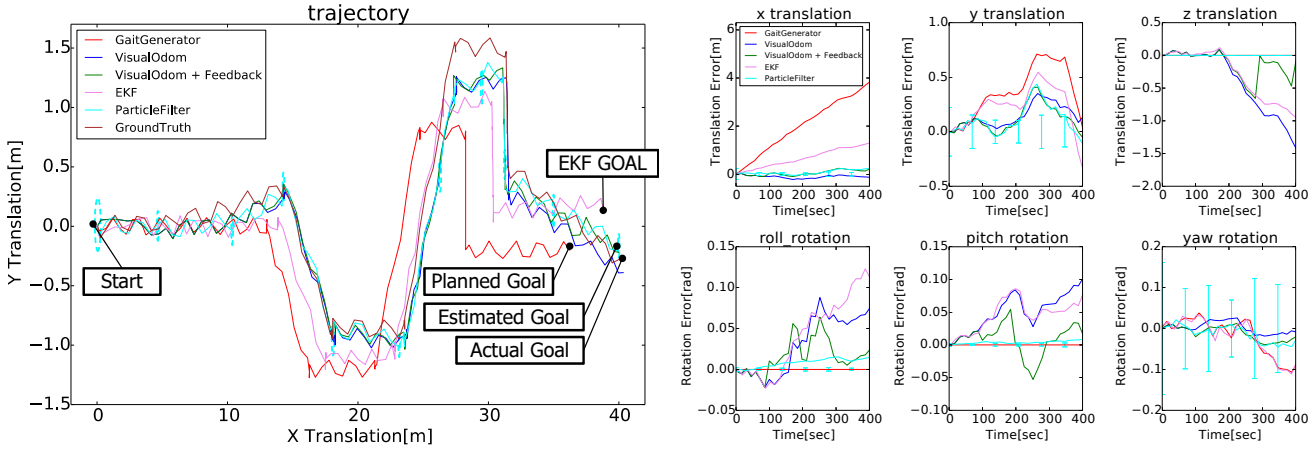


Fig. 4. Odometry Evaluation by JAXON in simulation. Left: Walking trajectory of JAXON in the evaluation experiment. Ellipses show $3\sigma$ errors of the Particle Filter estimation. Right: Errors from the ground truth in the evalluation experiment. Errorbars show $3\sigma$ errors of ParticleFilter estimation.

virtual horizontal constant height scan from an accumulated point cloud constructed from the raw laser scans.

We convert a raw laser scan at time $k$ into a point cloud and obtain the filtered point cloud $S_k^l$ in which the noise and the part of robot itself and ground are removed. Then, we transform $S_k^l$ into $S_k^o$ which is relative to the origin of the improved odometry obtained in Section IV to prevent dynamic motion of the robot from affecting accumulation process and accumulate them into the full environment point cloud $P_t^o = \sum_k^t S_k^o$ until the time $t$ in a quasistatic manner.

The virtual scan plane, which is horizontal and has constant height $h$ from the ground and fixed on the base link of the robot as shown in the upper left of Fig. 5, can be obtained using the improved odometry and the coordinate of the base link calculated using forward kinematics according to joint angles. We slice the accumulated environment point cloud $P_t^o$ by the virtual scan plane and generate a virtual horizontal constant height scan $\hat{S}_t^o$ as shown in the lower left of Fig. 5. We use this virtual scan $\hat{S}_t^o$ as the input for Gmapping and obtain the transformation from improved odometry to the world origin ${}^{map}_{prt}T_t$. In this paper, we set the height $h$ as 2.5[m] to detect only static walls because we aim at improving localization accuracy. The experimental result of SLAM in real JAXON with the proposed virtual scan is shown in the right of the Fig. 5.

### C. Accuracy evaluation for 2D SLAM on real robot

We evaluated the effectiveness of the proposed method with real JAXON in the experimental environment shown in the left of Fig. 3. We set the number of the particles in improved odometry as $N$ =60 in consideration for computational cost. We apply feedforward walking trajectory three times to JAXON and compared the result of proposed localization with ground truth calculated using the checkerboard. It should be noted that we used the checkerboard only for evaluation and the calculated ground truth is indepenent of the proposed localization system of the JAXON.

The result is shown in Fig. 6. As with the case of the simulation experiment shown in Section IV, proposed localization succeeded in improving the localization accuracy reducing the drift of visual odometry in the real JAXON. Especially, the errors of the position in x and y direction and the orientation around yaw axis falls within the range of the errorbars and this proves that the error models are successfully estimated.

The errors are shown in TABLE. II. It can be seen that the combination of improved odometry and 2D SLAM to virtual scan reduces the average errors compared with the result of using only improved odometry. Especially, 2D SLAM compensated the error in y direction which is strongly affected by the error of visual odometry.

TABLE II
EVALUATION OF LOCALIZATION ERRORS

| Direction | Method | Avg Abs Errors | Final Errors |
|---|---|---|---|
| x translation | Gait Generator | 1.03[m] | 2.31[m] |
| | Particle Filter | 0.054[m] | 0.025[m] |
| | PF + SLAM | 0.038[m] | -0.023[m] |
| y translation | Gait Generator | 0.070[m] | 0.043[m] |
| | Particle Filter | 0.064[m] | 0.15[m] |
| | PF + SLAM | 0.055[m] | 0.058[m] |
| yaw rotation | Gait Generator | 0.14[rad] | -0.29[rad] |
| | Particle Filter | 0.019[rad] | 0.020[rad] |
| | PF + SLAM | 0.025[rad] | 0.037[rad] |

## VI. EXPERIMENTS

### A. Evaluation of localization in terrain walking

We conducted an experiment where JAXON goes up and down a terrain with planned trajectory to evaluate the effectiveness of the proposed method in the situation where the height of the robot changes. Results are shown in Fig. 7. Even in terrain walking, where the height of the base link changes, the proposed localization method can estimate its position within approximately 3[cm] and orientation within approximately 1[deg] with the constant map generated from the virtual constant height horizontal scans.

### B. Evaluation of localization during valve approach

We integrated the proposed localization method and our teleoperation system [21] in order to evaluate its effectiveness

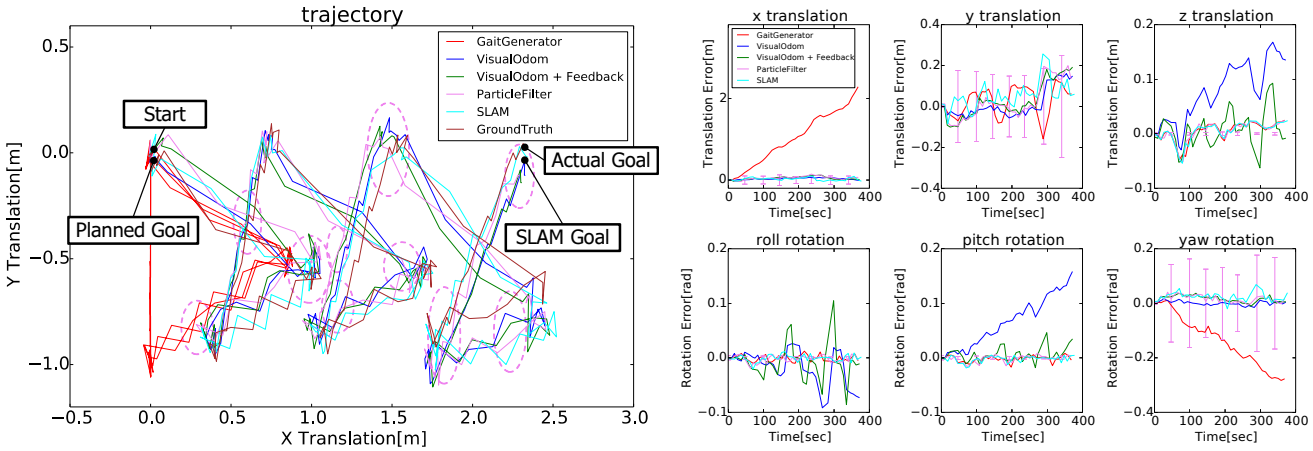


Fig. 6. Evaluation of self localization from 2D SLAM by real JAXON. Left: Walking trajectory of JAXON in the evaluation experiment. Ellipses show $3\sigma$ errors of ParticleFilter estimation. Right: Errors from the ground truth in the evalluation experiment. Errorbars show $3\sigma$ errors of ParticleFilter estimation.

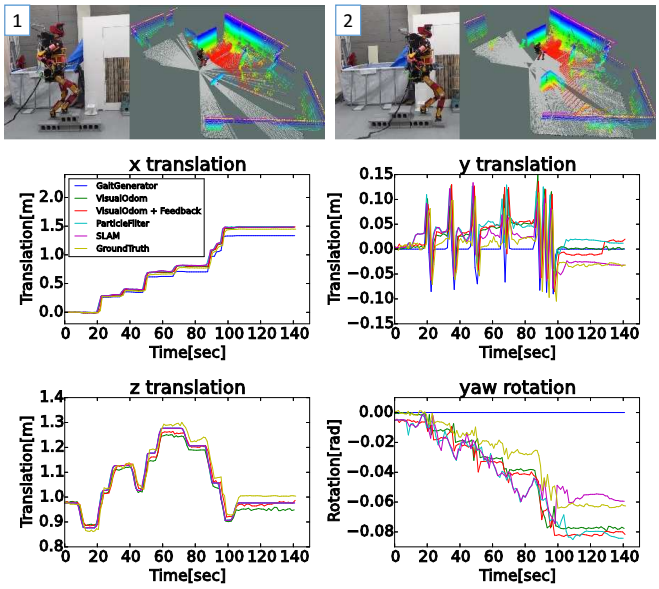


Fig. 7. Evaluation experiment of proposed localization with JAXON going up and down a terrain. Upper: Captures of real JAXON and generated maps. Lower: Translation of x, y and z axis and rotation around z in the experiment.

by comparing its localization result with the ground truth calculated from the checkerboard in a valve turning task. In the valve turning task, we fixed the target posture of the gait generator according to the estimated posture from the proposed localization method. The result is shown in Fig. 8. Using the proposed method and target posture modification, JAXON reached the planned posture within 3.0[cm] and 1.0[deg]. Considering hand measurements of JAXON, these errors are within the required accuracy for turning the valve. In contrast, feedforward odometry from the gait generator has -25[cm] error in x direction, -4.0[cm] error in y direction and 5.0[deg] error in yaw orientation. These errors severely impact the success of the operation and resulted in failure of the task as shown in the lower right of Fig. 8. We further performed 10 additional trials each with and without the proposed method. JAXON succeeded in turning the valve 9 times out of the 10 trials (with one failure due to inaccurate

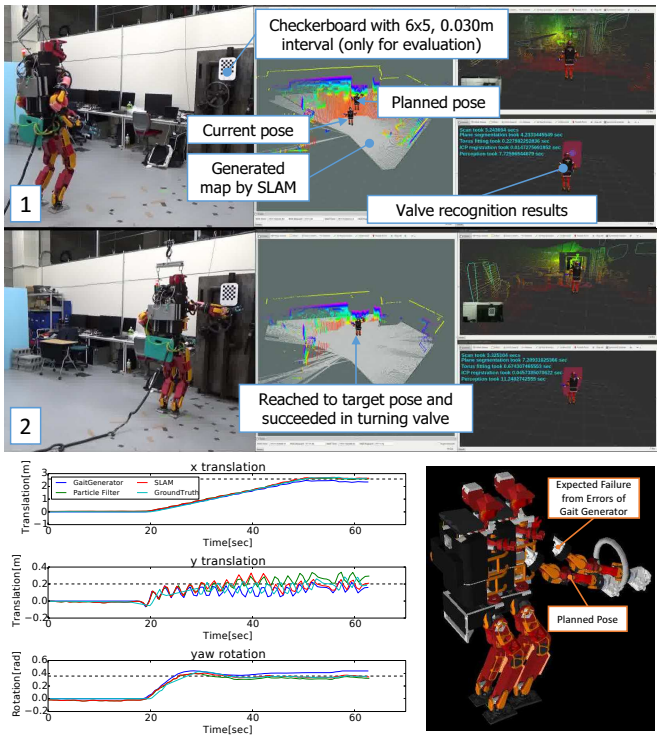


Fig. 8. JAXON turns valve with proposed localization and feedback velocity control. Upper: Captures of real JAXON and its user interface in turning valve. Lower Left: Translation of x and y axis and rotation around z in the experiment. Black dotted lines show the target pose. Lower Right: Estimated pose resulting from the gait generator odometry is used.

estimation of visual odometry) with the proposed method while it succeeded 0 times with only feedforward odometry.

### C. Continuous door opening and valve turning task

We evaluated the feasibility of the proposed localization through continuous door opening and valve turning tasks.The results are shown in Fig. 9. Proposed method succeeded in generating accurate maps in the tasks where several disturbances occur such as change of base link height in the door opening task, oscillation in walking or upper body twist in valve turning. Moreover, due to accurate localization and by fixing target posture of the gait generator, JAXON

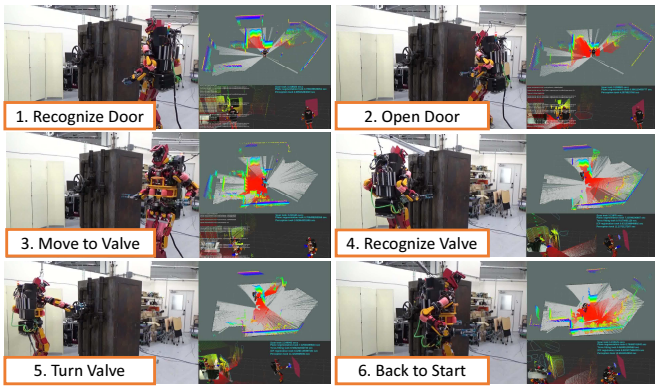


Fig. 9. JAXON opens a door and turns a valve continuously using the proposed localization. In each image, the capture of the experiment is shown in left, the generated map is shown in upper right and recognition results of the teleoperation system [21] are shown in lower right.

could achieve these tasks without running recognition of the target again, manual modification of posture and replanning of motion after reaching its target position, which were all needed before the introduction of the proposed localization. From this result, it can be concluded that the autonomy of the robot was improved by the proposed method.

## VII. Conclusion

In this paper, we proposed a localization pipeline for humanoid robots in order to achieve tasks reliably in unknown environments. Our contribution is enabling a humanoid robot to localize itself within approximately 5.0[cm] and 1.0[deg] accuracy in unknown environments through a fusion beetween an imporved odometry consisting in visual odometry, inertial sensors, gait controller odometry and an efficient laser based 2D SLAM. The 2D SLAM world observation builds on the improved odometry with 40[Hz] frequency to extract a virtual horizontal scan. Such scan results from the proper slicing of a point cloud issued from 3D space accumulated successive scans. The sensor fusion process uses a velocity based error model in order to derive more accurate estimates assuming a normal distribution. We experimentally showed that the accuracy of our localization system was enough to increase the autonomy and the success rate of humanoid robots performing various tasks.

Strictly speaking, the velocity errors are dependent on the robot's posture, floor characteristics and the task at hand. We dealt with this problem by using larger standard deviations but we are now trying to estimate velocity error parameters while performing tasks for more accurate localization. As a future work, we will integrate the proposed localization system with our work on trajectory planning [22] in order to further improve the autonomy of humanoid robots.

## References

[1] Kunio Kojima, Tatsuhi Karasawa, Toyotaka Kozuki, Eisoku Kuroiwa, Sou Yukizaki, Satoshi Iwaishi, Tatsuya Ishikawa, Ryo Koyama, Shintaro Noda, Fumihito Sugai, Shunichi Nozawa, Yohei Kakiuchi, Kei Okada, and Masayuki Inaba. Development of life-sized high-power humanoid robot jaxon for real-world use. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots (Humanoids 2015)*, pages 838–843, November 2015.

[2] S. Thompson and S. Kagami. Humanoid robot localisation using stereo vision. *IEEE-RAS International Conference on Humanoid Robots*, pages 19–25, 2005.

[3] Nosan Kwak, Olivier Stasse, Torea Foissotte, and Kazuhito Yokoi. 3D grid and particle based SLAM for a humanoid robot. *IEEE-RAS International Conference on Humanoid Robots*, pages 62–67, 2009.

[4] Olivier Stasse, Andrew J Davison, Ramzi Sellaouti, Kazuhito Yokoi, and Cnrs Aist. Real-time 3D SLAM for Humanoid Robot considering Pattern Generator Information. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 348–355, 2006.

[5] Giuseppe Oriolo, Antonio Paolillo, Lorenzo Rosa, and Marilena Vendittelli. Humanoid odometric localization integrating kinematic, inertial and visual information. *Autonomous Robots*, pages 1–13, 2015.

[6] Simon Thompson, Satoshi Kagami, and Koichi Nishiwaki. Localisation for autonomous humanoid navigation. *Proceedings of IEEE/RAS International Conference on Humanoid Robots*, pages 13–19, 2006.

[7] a. Hornung, K.M. Wurm, and M. Bennewitz. Humanoid robot localization in complex indoor environments. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1690–1695, 2010.

[8] Ricardo Tellez, Francesco Ferro, Dario Mora, Daniel Pinyol, and Davide Faconti. Autonomous humanoid navigation using laser and odometry data. *IEEE/RAS International Conference on Humanoid Robots*, pages 500–506, 2008.

[9] Maurice F Fallon, Matthew Antone, Nicholas Roy, and Seth Teller. Drift-free humanoid state estimation fusing kinematic, inertial and LIDAR sensing. *IEEE/RAS International Conference on Humanoid Robots*, pages 112–119, 2014.

[10] Ken Masuya and Tomomichi Sugihara. Dead reckoning for biped robots that suffers less from foot contact condition based on anchoring pivot estimation. *Advanced Robotics*, (August):1–15, 2015.

[11] G. Klein and D.W. Murray. Parallel Tracking and Mapping for Small AR Workspaces. *IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–10, 2007.

[12] Christian Forster, Matia Pizzoli, and Davide Scaramuzza. SVO : Fast Semi-Direct Monocular Visual Odometry. *IEEE International Conference on Robotics and Automation*, 2014.

[13] A. Bachrach, S. Prentice, R. He, P. Henry, a. S. Huang, M. Krainin, D. Maturana, D. Fox, and N. Roy. Visual Odometry and Mapping for Autonomous Flight Using an RGB-D Camera. *The International Journal of Robotics Research*, 31(11):1320–1343, 2012.

[14] Bernd Kitt, Andreas Geiger, and Henning Lategahn. Visual odometry based on stereo image sequences with RANSAC-based outlier rejection scheme. *Proceedings of IEEE Intelligent Vehicles Symposium*, pages 486–492, 2010.

[15] K. Nagatani, S. Tachibana, M. Sofne, and Y. Tanaka. Improvement of odometry for omnidirectional vehicle using optical flow information. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 468–473, 2000.

[16] a.W. Stroupe, M.C. Martin, and T. Balch. Distributed sensor fusion for object position estimation by multi-robot systems. *Proceedings of IEEE International Conference on Robotics and Automation*, 2, 2001.

[17] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. Intelligent robotics and autonomous agents. MIT Press, 2005.

[18] Giorgio Grisetti, C Stachniss, and W Burgard. Improved Techniques for Grid Mapping with Rao-Blackwellized Particle Filters. *IEEE Transactions on Robotics*, 23:34–46, 2007.

[19] Mathieu Labb and Francois Michaud. Online Global Loop Closure Detection for Large-Scale Multi-Session Graph-Based SLAM. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2661–2666, 2014.

[20] Ji Zhang and Sanjiv Singh. Visual-lidar Odometry and Mapping: Low-drift, Robust, and Fast. *IEEE International Conference on Robotics and Automation*, pages 2174–2181, 2015.

[21] Yu Ohara, Masaki Murooka, Ryohei Ueda, Shunichi Nozawa, Yohei Kakiuchi, Kei Okada, and Masayuki Inaba. Configurable autonomy applicable to humanoid manipulation in unstructured and communication-limited environment. pages 373–380, November 2015.

[22] Ryohei Ueda, Shunichi Nozawa, Kei Okada, and Masayuki Inaba. Biped Humanoid Navigation System Supervised through Interruptible User-Interface with Asynchronous Vision and Foot Sensor Monitoring. In *IEEE-RAS International Conference on Humanoid Robots*, pages 273—-278, 2014.