International Journal of Information Acquisition

A MULTI-CAMERA VIEW DIRECTION PLANNING STRATEGY FOR MOBILE ROBOTS

Tingting Xu, Kolja Kühnlenz, Martin Buss
Institute of Automatic Control Engineering (LSR)
Technische Universität München
D-80290 Munich, Germany

Received (to be inserted by publisher)

A multi-camera view direction planning strategy for mobile robots is discussed. Two concurrent tasks for efficient and safe locomotion in dynamic environments are considered: self-localization and dynamic object tracking. The approach is to assign different tasks to different cameras, such that for each task an individual optimal view direction is selected based on information gain maximization. Thereby, the individual task performance is improved significantly. The performance of the proposed strategy is evaluated in simulations considering a humanoid robot navigation scenario and compared with two conventional multi-camera view direction planning strategies.

Keywords: view direction planning, multi-camera system, information gain maximization, robotics, computer vision

1. Introduction

For the purpose of locomotion in an unknown real-world environment, a cognitive mobile robot should also have the human-like ability to react to the dynamic world. Humans, like all animals, respond very readily to novel objects and fast changes in their environment. They can be, for example, significantly attracted by motion in real-world environments, which can help to gather task-relevant information and increase their own safety.

To achieve an efficient and safe behavior, particularly useful is an active vision head controlled by a visual attention system that selects task-relevant viewpoints in the environment. Aiming at that, in earlier works we proposed a high-performance multi-focal system, consisting of two independently controllable stereo cameras (see Fig. 1), which is applied in urban environments [Lidoris *et al.*, 2007].

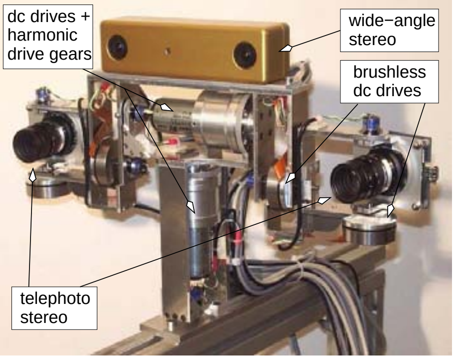

Fig. 1. Multi-focal vision system [Kühnlenz *et al.*, 2006]

In this paper a view direction planning strategy for this camera platform is discussed. Two typical concurrent tasks for effective locomotion in dynamic environments are considered: self-localization and object tracking. An accurate position estimation of mobile robot itself and dynamic objects in the surroundings is the precondition for a successful navigation in real-world environments.

The approach is to assign different tasks to different cameras, such that for each task an individual optimal view direction is selected based on information gain maximization. The performance of the proposed strategy is evaluated in simulations considering a humanoid robot navigation scenario and compared with two conventional multi-camera view direction planning strategies. Using the proposed view direction planning strategy, the accuracies of the robot motion estimation and the object motion estimation are improved significantly in a dynamic environment.

The paper is organized as follows: In Section 2, related work about view direction planning for mobile robot and multi-camera systems is described. The proposed view direction planning technique is presented in Section 3. In Section 4, simulations and results are shown and discussed. Conclusions and future work are given in Section 5.

2. Related Work

Considering improvement of robot navigation provided by active vision, a variety of approaches for view direction planning is proposed. A real-time EKF-based SLAM system using active vision is combined with information from odometry and a roll/pitch accelerometer sensor in [Davison & Kita, 2001] to perform accurate, repeatable localization while traversing an undulating course. In [Pellkofer & Dickmanns, 2000] an approach to an optimal gaze control system for autonomous vehicles is proposed in which the perceptive situation and also subjective situation are predicted. In [Arbel & Ferrie, 1999] a viewpoint selection strategy based on entropy maps is introduced, based on a sequential recognition strategy in which object hypotheses are represented as conditional probability density functions. Selection of viewpoints is accomplished using an active vision approach that selects on the basis of minimizing ambiguity of recognition. For gaze control of humanoid robot the idea of [Seara & Schmidt, 2005] is based on maximization of the predicted visual information content of a view situation. A task decision strategy is applied to view direction selection for multiple tasks.

Due to limited field of view and accuracy of single or stereo camera systems, many works proposed the use of multi-camera vision systems. In [Lippiello *et al.*, 2003] a vision system consisting of two fixed cameras for the position and orientation tracking of a moving object is proposed. In [Ahmedali & Clark, 2006] collaborative multi-camera surveillance with automated person detection is used. Multiple cameras with overlapping fields of view collaborate to confirm results and to implicitly handle non-overlapping regions. In [Keyes *et al.*, 2006] the impact of camera location and multi-camera fusion with real robots in an urban search and rescue task is studied. It is found that having two cameras, one forward-facing and one rear-facing, results in improved situation awareness. In [Krumm *et al.*, 2000] a practical person-tracking system is proposed using two sets of color stereo cameras for locating people with the stereo images and for maintaining their identities with the color images. To achieve both wide field of view and high accuracy, in [Kühnlenz, 2006] multi-focal vision is used for measurement and robotics applications. Advantages are significant improvements of control performance and localization accuracy as well as an extension of the workspace compared to conventional approaches.

To accomplish self-localization and tracking of moving objects simultaneously, an active vision system could plan its view direction using the decision strategy mentioned in [Seara & Schmidt, 2005], namely a view direction selected at each time step only for one task. However, due to the simultaneous movements of object and robot, the object position estimation error has an impact on the robot position estimation. Vice versa, the robot position estimation error also affects the object position estimation. To achieve a satisfying accuracy for both tasks, in this paper a view direction planning strategy based on multi-camera vision is proposed. Different stereo cameras are used in order to accomplish different tasks. The view directions of two stereo cameras are controlled independently based on information gain maximization. Thereby, the individual task performance is improved significantly. The performance of the proposed strategy is evaluated in simulations and compared with two conventional strategies using the task decision concept mentioned in [Seara & Schmidt, 2005] considering a humanoid robot navigation scenario.

3. Task Definition

3.1. *Problem description*

Self-localization is one of the most important tasks for a mobile robot. From visual information in a natural scene, the robot can choose appropriate landmarks such as point features or edges for its localization computation. In a real-world environment moving objects also exist, considered as interesting regions such as humans and vehicles. If the robot's attention is directed to them to acquire better accuracy for object position estimation, the robot ego-motion estimation is corrupted due to simultaneous movements of robot and objects. The object position estimation error and the robot position estimation error have negative impacts on each other. To solve this problem, we propose a view direction planning strategy based on multi-camera vision.

For simplicity, in this paper we only refer to moving objects as objects. All other static objects in the environment could be regarded as landmarks and be used for the localization task. We assume that the robot can distinguish between static and dynamic things using general motion detection algorithms. For example, in [Xu *et al.*, 2008] an information-based strategy for detection of static outliers and temporal novelty is presented.

3.2. *View direction planning strategy*

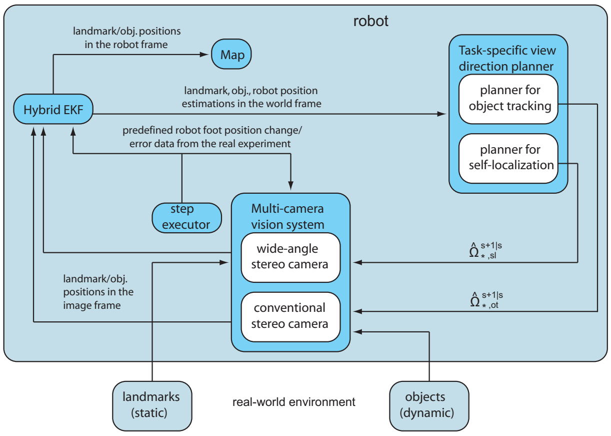

Fig. 2. Multi-focal view direction planning architecture and simulation layout

A view direction planning strategy based on multi-camera vision is proposed. In our approach, different tasks are assigned to different cameras. The view direction planner is divided into several planners, according to the number of tasks. Each planner acquires different information from the environment, processes it separately, utilizing its individual task-relevant criterion, and chooses the most appropriate view direction for its respective camera, such that each task can be accomplished optimally. For purpose of robot self-localization and objects tracking in the environment, two view direction planners are needed (see Fig. 2).

We use an information-based concept for the planners. The idea is to choose the view direction which is predicted to provide maximum information gain.

One planner is applied for the self-localization (sl) task. The information measure v^s_{sl} for this task at step s is defined as follows:

$$v^s_{sl} = \frac{1}{2} \sum_{j=1}^{2} \sqrt{e^s_j}, \tag{1}$$

where e^s_j is the eigenvalue of the robot position estimation covariance matrix and j is the index for x- and y-direction. We use a hybrid Extended Kalman filter [Seara & Schmidt, 2005] to predict the robot position and calculate covariance matrices. The error covariance matrix is a measure of the estimation accuracy. The smaller the eigenvalue of the covariance matrix is, the more accurate is the position estimation. For possible view directions $\hat{\Omega}$ of the respective camera, the respective information measures $\hat{v}^{s+1|s}_{sl}$ will be computed. The view direction $\hat{\Omega}^{s+1|s}_{*,sl}$ with the strongest information content increase, which is defined as the covariance decrease, will be chosen as the optimal view direction by the planner and applied for the next time step $s+1$ to the respective camera.

$$\hat{\Omega}^{s+1|s}_{*,sl} = \arg\max_{\hat{\Omega}} (v^s_{sl} - \hat{v}^{s+1|s}_{sl}). \tag{2}$$

In the view direction planner for the object tracking (ot) task, the distances between the robot and the objects will be checked firstly. The eigenvalues of the position estimation covariance matrix of the object nearest to the robot is considered as information measure v^s_{ot} for this task at step s:

$$v^s_{ot} = \sum_{j=1}^{3} \sqrt{e^s_j}, \tag{3}$$

where e_j^s is the eigenvalue of the object position covariance matrix and j is the index for x-, y- and z-direction. Analogously, a view direction with the strongest information content increase is regarded as the optimal one and should be focused by the other camera.

$$\hat{\Omega}_{*,ot}^{s+1|s} = \arg\max_{\hat{\Omega}}(v_{ot}^{s} - \hat{v}_{ot}^{s+1|s}). \qquad (4)$$

Two view direction planners use two different information measure definitions. There is no overlapping in the computation of the view directions. Applying this strategy, the view direction planner of the self-localization task is independent of the view direction planner of the object tracking task. In other words, the position estimation of the robot ego-motion does not depend on the object movement any more, so that the accuracy of the robot ego-motion estimation will be improved. Consequently, the position estimation of the object will, respectively, also be improved.

4. Performance Evaluation

To evaluate the performance of the view direction planning strategy proposed in Section 3, several simulations are conducted. A humanoid robot navigation scenario is considered.

4.1. *Simulation description*

In our simulation, a walking humanoid robot is equipped with two active stereo cameras, providing the robot with visual information. A wide-angle stereo camera, which has aperture angles of 60° in each direction with focal lengths of 2mm, is responsible for the robot self-localization task, since it can locate more landmarks concurrently in its wide field of view. A conventional stereo camera corresponding to the task of object position estimation has aperture angles of 30° in each direction with focal lengths of 20mm and can provide high accuracy of the position estimation of the nearest moving object. We assume that two stereo cameras are independently controllable like the multi-focal camera system introduced in Section 1, such that we can plan their view directions individually. The multi-camera vision system should be controlled such that the position estimation errors of the robot and the present objects are minimized.

Our view direction planning architecture and simulation layout is shown in Fig. 2.

Based on [Seara & Schmidt, 2005] and [Kühnlenz, 2006], the step executor models the biped walking of a humanoid robot in the world frame. The robot frame is placed in the current supporting foot. A hybrid Extended Kalman Filter (EKF) is the central part of the whole system. An EKF is an efficient recursive filter which gives the state estimation of a dynamic system from a series of noisy measurements. The perception errors are thoroughly studied in [Lorch *et al.*, 2002], in which the hybrid EKF is proposed, accounting for biped locomotion. The system state $\mathbf{x}_{k+1}$ at step $k+1$ is represented as follows:

$$\begin{aligned} \mathbf{x}_{k+1} = & \begin{bmatrix} {}_0\mathbf{x}_{0_{k+1}} \\ {}_F\mathbf{x}_{0_{k+1}} \\ \vdots \end{bmatrix} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_{k+1}, \mathbf{w}_k) \\ & = \mathbf{x}_k(1-\gamma_{k+1}) + \mathbf{f}_s(\mathbf{x}_k, \mathbf{u}_{k+1}, \mathbf{w}_k)\gamma_{k+1} \end{aligned} \qquad (5)$$

where $\gamma_{k+1} \in \{0,1\}$ and $\gamma = 1$ represents the robot foot changes. ${}_0\mathbf{x}_{0_{k+1}}$ and ${}_F\mathbf{x}_{0_{k+1}}$ are the state vectors in the world frame and in the robot frame, respectively. $\mathbf{u}_{k+1}$ is the current control vector, provided by the step executor and $\mathbf{w}$ is the random variable representing the process noise. $\mathbf{f}_s$ is the transformed system state after the foot change.

The approach is to assign different tasks to different cameras. Different environment informations are captured by the multi-camera vision system and provided to the hybrid EKF. The wide-angle stereo camera provides the position estimator, the hybrid EKF, with the landmark positions in the image frame, while the conventional stereo camera only brings the position informations of objects to the position estimator. In this hybrid EKF the landmark positions, the object positions, the robot supporting foot positions and the robot camera head positions in the world frame are predicted. This prediction is evaluated by the task-specific view direction planner.

The task-specific view direction planner evaluates the information content increase for all possible view directions in different planners for different tasks. The task-specific optimal view directions are provided to the wide-angle stereo camera and to the conventional stereo camera, respectively.

In our simulations the robot should walk straight forward parallel to the positive x-axis (see Fig. 3). The real path of the camera head is shown with the blue line. The robot walks 24 steps in total. Because of the walking error data which has been taken from real experiments with a humanoid robot [Seara, 2004], the path is not a straight line. The robot foot position in each step is presented by blue stars. The start point of the right foot is $(0, 0.9)$m. At each step the robot is moving 0.30m in the x-direction. The distance between two robot feet is 0.20m.

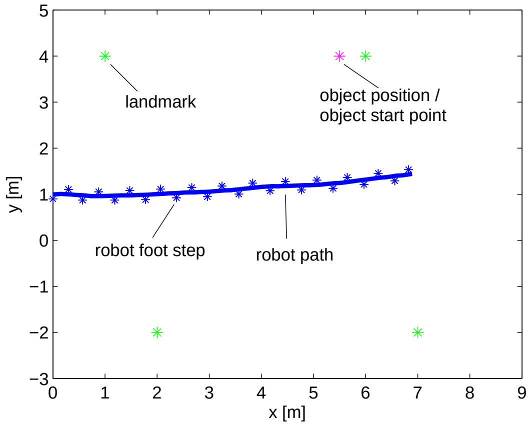

Fig. 3. Simulation scenario: the robot walking path, the landmark positions, the object position.

There are four landmarks around the robot path. Their positions are (1, 4)m, (2, -2)m, (6, 4)m and (7, -2)m. They are marked with green stars. An object is located in the environment at the position (5.5, 4)m, marked with a pink cross.

We compare three multi-camera planning strategies evaluating the absolute position estimation errors and the estimation covariance of the robot ego-motion as well as the object motion based on visual information:

- Strategy 1 (S1): Two stereo cameras are rigidly connected and not independently controllable. The task decision strategy (see [Seara & Schmidt, 2005]) is used to choose the task for both cameras at each step. The view direction planner computes either the optimal view direction for self-localization or the optimal view direction for object position estimation at each step.
- Strategy 2 (S2): Two stereo cameras are independently controllable. The task decision strategy is also applied to the view direction planners. Both stereo cameras execute the same task at each step, but their view directions may be different.
- Strategy 3 (S3): Two stereo cameras are independently controllable. The view direction strategy proposed in Section 3 is applied. Different tasks are assigned to different stereo cameras. The wide-angle stereo camera is responsible for the self-localization task, while the conventional stereo camera is accountable for the object position estimation task.

Pinhole camera model is used. All the cameras are assumed to be ideal. Lens distortion and quantization effects are not considered.

The other system parameters are set as follows: At each step the positions are measured and estimated three times. Step parameters covariance equals $0.005^2 m^2$ in each direction, x and y. The square of the variance of the robot orientation in each step has the value of $\sqrt{3}\ rad^2$. The initial position variance of the object equals $0.02^2 m^2$, while the initial position variance of the robot also equals $0.02^2 m^2$.

4.2. *Simulation results*

4.2.1. *In a static environment*

Firstly, we consider the situation that the object at $(5.5, 4)m$ is not moving, such that this object is also treated as a landmark.

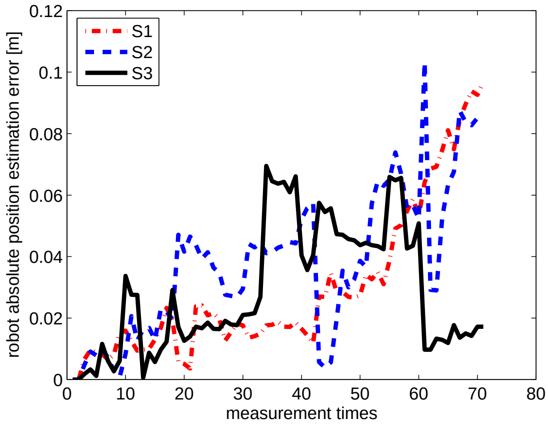

Fig. 4. Absolute position estimation errors of the robot ego-motion in a static environment. Two stereo cameras are: rigidly connected (S1), independently controllable with the same task (S2), or independently controllable with individual task (S3).

In Fig. 4 the absolute position estimation errors of the robot ego-motion using view direction planning strategy S1, S2 and S3 are compared. The estimation using S3 has no apparent improvement in comparison to the estimations using S1 and S2. The mean value of the robot position estimation error using S3 is 0.0285m, while the mean values using S1 and S2 are 0.0384m and 0.030m, respectively.

Fig. 5 shows a comparison of the object absolute position estimation errors in this static environment using three strategies. The mean values of the object absolute position estimation errors using S1, S2 and S3 are 0.0229m, 0.0250m and 0.0214m, respectively. The object estimation applying S3 has no apparently better performance than those using S1 and S2, either.

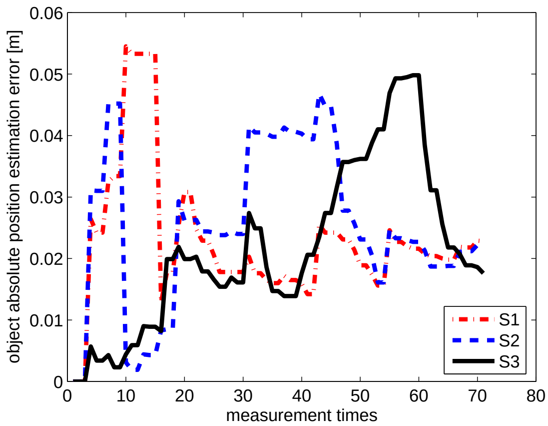

Fig. 5. Absolute position estimation errors of the object in a static environment. Two stereo cameras are: rigidly connected (S1), independently controllable with the same task (S2), or independently controllable with individual task (S3).

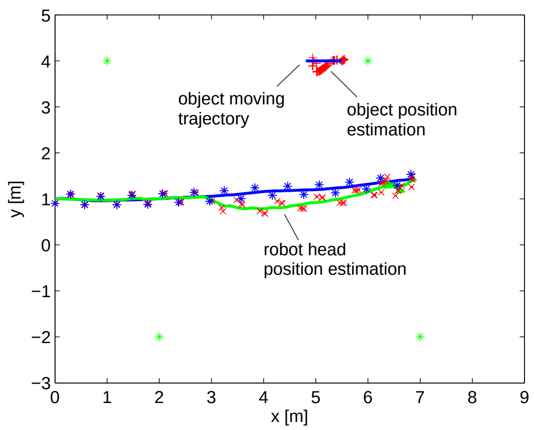

Fig. 6. Simulation result displayed in an overlooking view using S1 in a dynamic environment.

It is established that S3 has no obviously better performance in position estimation in a static environment than S1 and S2.

4.2.2. *In a dynamic environment*

Now we investigate the case that an object is simultaneously moving along a straight line toward the negative x-direction (see Fig. 6, 7 and 8). The start point is (5.5, 4)m. The trajectory of the object is displayed by a blue line. During the time interval of one step of the robot, the object moves 0.03m in the negative x-direction. The object position estimation is shown in red crosses. The green line is the estimated trajectory of the robot head. The red stars near the green line are the position estimations of the robot foot currently supporting at each step.

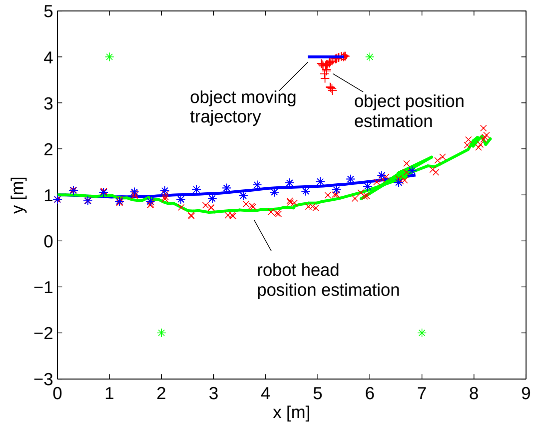

Fig. 7. Simulation result displayed in an overlooking view using S2 in a dynamic environment.

Fig. 6 and Fig. 7 are the overlooking views of the robot motion and the estimation results using S1 and S2. The position estimations of the robot head using S1 and S2 are far away from the real path for the reason that the object seen by the robot is also moving. Using S3 we have a significant improvement in the robot position estimation and the object position estimation (see Fig. 8).

Fig. 9 shows the quantitative comparison of the robot position estimation errors over the simulation time. The mean value of the robot absolute position estimation errors using S1 equals 0.3119m, while the absolute robot position estimation errors applying S2 have the mean value of 0.6841m. If S3 is used, the mean value of the estimation errors is

0.0623m which is obviously smaller than those using S1 and S2. The accuracy is much higher using S3 than those using S1 and S2.

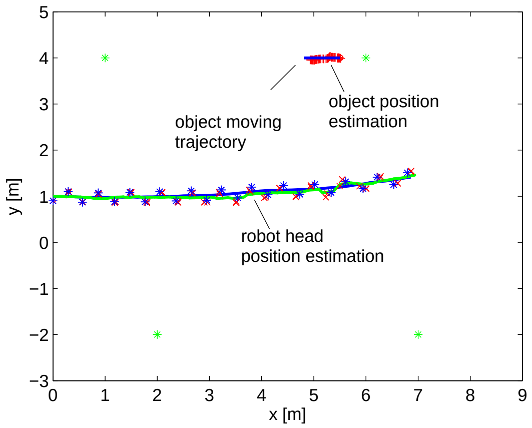

Fig. 8. Simulation result displayed in a overlooking view using S3 in a dynamic environment.

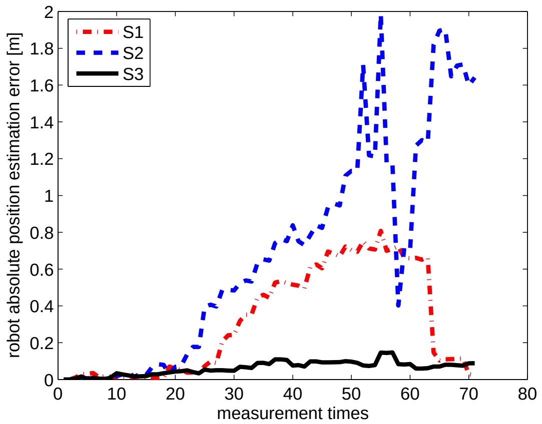

Fig. 9. Absolute position estimation errors of the robot ego-motion in a dynamic environment. Two stereo cameras are: rigidly connected (S1), independently controllable with the same task (S2), or independently controllable with individual task (S3).

Fig. 10 illustrates a comparison of robot position estimation covariances using S1, S2 and S3 over 71 measure times (three times per step). The covariance of S3 is much smaller and a little smaller than those of S1 and S2. The oscillation is caused by three measurements and estimation using the hybrid EKF at each step. By the second measurement the covariance of the ego motion is obviously reduced.

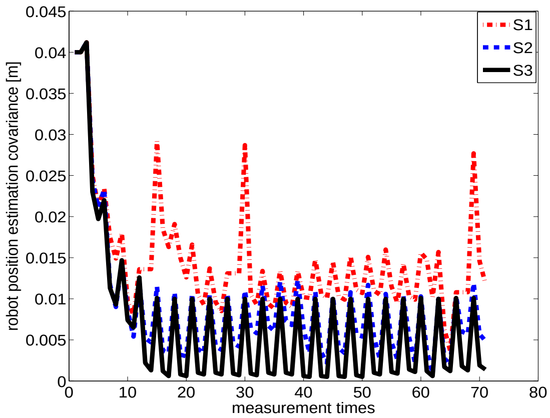

Fig. 10. Robot ego-motion estimation covariance. Two stereo cameras are: rigidly connected (S1), independently controllable with the same task (S2), or independently controllable with individual task (S3).

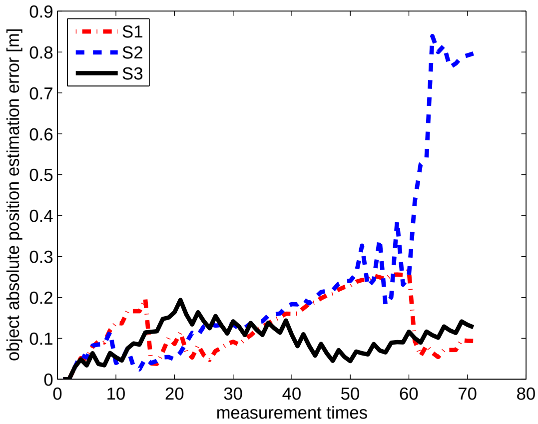

Fig. 11. Absolute position estimation errors of the object motion in dynamic environment. Two stereo cameras are: rigidly connected (S1), independently controllable with the same task (S2), or independently controllable with individual task (S3).

Considering the comparison of the absolute errors and covariances, S3 performs better than S1 and S2 for robot ego-motion estimation.

Respectively, S3 also has a better performance of the object position estimation. A comparison among S1, S2 and S3 is presented in Figure 11. The mean values of the absolute object position estimation errors using S1, S2 and S3 are 0.1295m, 0.2300m and 0.0976m, respectively. The average position estimation error using S3 is about 0.03m and 0.13m smaller than those using S2 and S1.

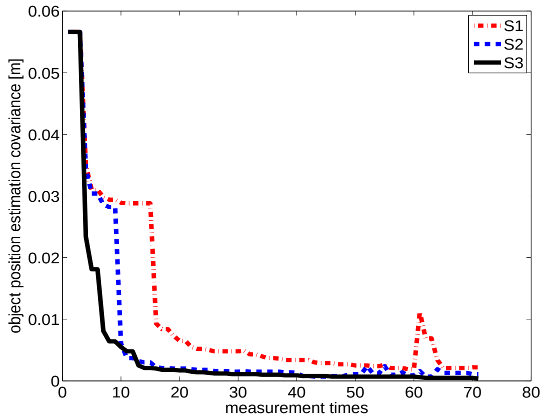

Fig. 12. Position estimation covariance of the object motion in dynamic environment. Two stereo cameras are: rigidly connected (S1), independently controllable with the same task (S2), or independently controllable with individual task (S3).

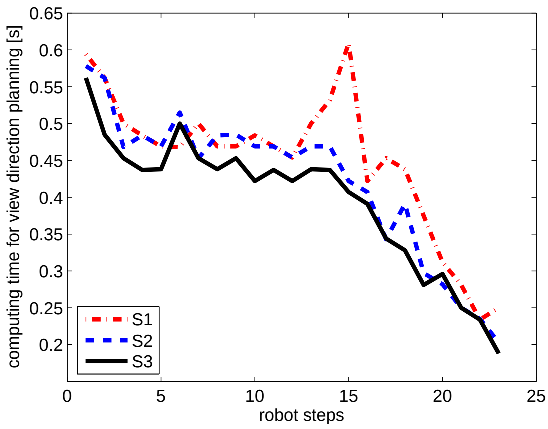

Fig. 13. Computational cost for view direction planning. Two stereo cameras are: rigidly connected (S1), independently controllable with the same task (S2), or independently controllable with individual task (S3).

Furthermore, the object position estimation covariances using S1, S2 and S3 are compared in Fig. 12. A same result is achieved as the comparison of robot position estimation. S3 performs much better than S1 and a little better than S2.

From the simulation results above, we can draw a conclusion as follows:

- $S3 > S1 > S2$ for absolute position estimation error,
- $S3 > S2 > S1$ for estimation covariance.

The accuracies of the position estimation of the robot motion and the object motion in a dynamic environment are improved apparently using S3.

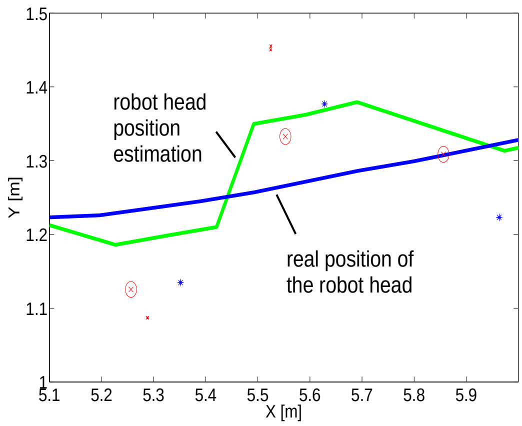

Fig. 14. Robot position estimation: three measurements per step.

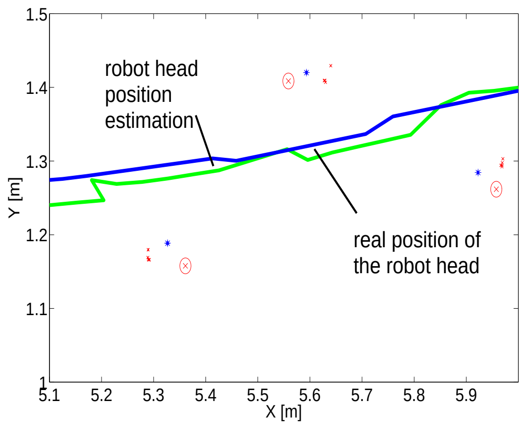

Fig. 15. Robot position estimation: six measurements per step.

In Fig. 13 a comparison of the computational cost using Matlab R2008a and AMD Dual Core 3800+ is illustrated. It takes, respectively, 0.449s, 0.419s and 0.395s using S1, S2 and S3 for one robot step, namely three measurements.

If we increase the measurement frequency per step, more accuracy is achieved. Fig. 14 and Fig. 15 show the enlarged estimation results for three and six measurements in one robot step. The blue lines indicate the real position of the robot head, while the green lines represent the position estimation of the robot head. Since an Gaussian distributed error is generated and added to the walking error, the real position of the robot head, the blue lines, in two simulations are not the same. Nevertheless, the position estimation error with six measurements per step is much smaller than the one with three measurements per step. Obviously, a high-speed image processing is the precondition of applications in dynamic environments.

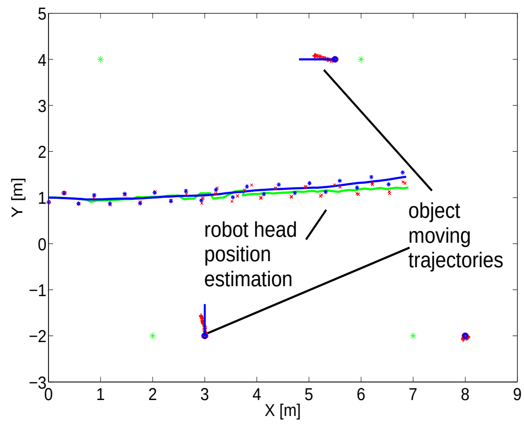

Fig. 16. Simulation result displayed in an overlooking view using S3 in a dynamic environment (more than one object).

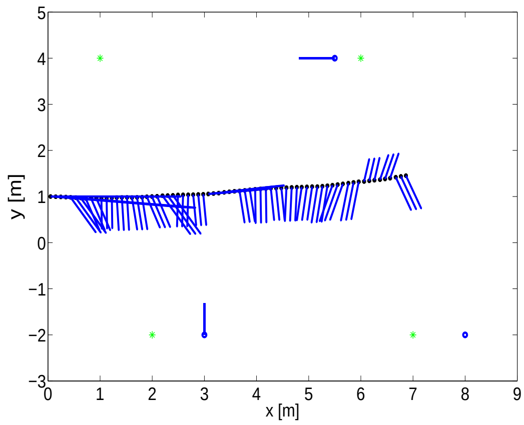

Fig. 17. View directions planned for the self-localization.

To test our strategy in environments with more than one object, in Fig. 16 another object, starting from the position (3, -2)m with vertical velocity towards the robot, is added, while at the position (8, -2)m a static object which can be seen as landmark is also included. Fig. 17 and Fig. 18 shows the planned view directions for the self-localization task, in which the attention of the wide-angle stereo camera is attracted by landmarks, and for the object position estimation, in which the conventional stereo camera directs its attention towards the nearest moving objects.

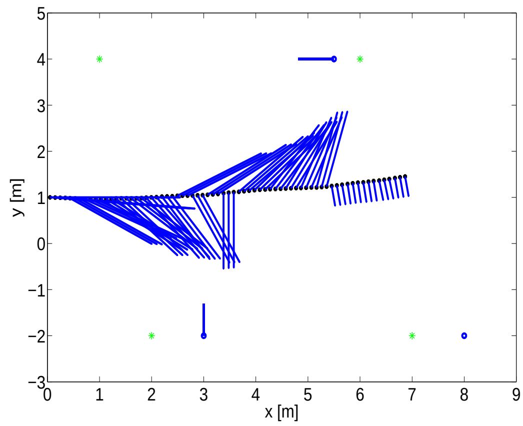

Fig. 18. View directions planned for the object position estimation.

4.3. *Discussion*

The simulation results above show a better performance in measuring the position estimation applying the view direction planning strategy S3 than those applying S1 and S2 in a dynamic environment. Without moving objects the accuracy of the position estimation using S3 is not improved significantly. But in a dynamic environment, the robot absolute position estimation error using S3 is even five times smaller than that using S1 and ten times smaller than that using S2. The small estimation covariance of S3 also provides a better accuracy of the estimation. The reason is the decoupled processing of different visual informations for different tasks. Without the influence of the object motion, the wide-angle stereo camera estimates the robot motion more accurately.

However, the view direction planning strategy S3 also has its limitations. Firstly, if we have a system in which the robot should accomplish more than two tasks, it is not always comfortable to use more than two cameras. To keep the improvement of accuracy provided by this view direction planning strategy, we can divide the system into several task groups. In each group there are only the tasks which do not have interactive impacts on the computation results of each other.

Secondly, the increase of the absolute position estimation error of the object motion shown in Figure 11 using S1 and S2 is principally due to the increase of the absolute position estimation error of robot motion. It is obvious that the absolute position estimation error of the object applying S3 increases at first. Then it decreases between the 20. measurement and the 50. measurement despite the fact that the absolute position estimation error of the robot motion grows in this time interval. This phenomenon is due to the object lying out of the field of view of the conventional stereo camera in the intervals from 1. to the 20. measurement and from the 50. measurement to the 71. measurement, which causes the increasing estimation error of the object position.

5. Conclusions and Future Work

An information-based multi-camera view direction planning strategy for mobile robots is discussed. Two concurrent tasks are considered: self-localization and object tracking. To avoid the interactive estimation error of moving robot and objects, different tasks are assigned to different cameras. Different cameras acquire different information from the environment, process them separately and plan their view directions individually, using their task-specific criteria. Information measures are defined for different tasks. The view direction with the maximum information gain is chosen as the optimal view direction which ensures an optimal view behavior with respect to task performance.

The performance of our strategy is evaluated in simulations considering a humanoid robot navigation scenario, compared with two conventional view direction planning strategies. Using the proposed view direction planning strategy, the accuracies of the robot motion estimation and the object motion estimation are improved significantly in a dynamic environment. The impact of the object motion on the robot motion estimation is eliminated. The performance of accomplishing concurrent tasks is improved.

It is envisioned to implement this strategy on the multi-focal multi-camera platform mentioned in Section 1. An evaluation in real world environment should be accomplished.

ACKNOWLEDGMENTS

This work is supported in part within the DFG excellence initiative research cluster *Cognition for Technical Systems – CoTeSys*, see also `www.cotesys.org`.

References

A. J. Davison and N. Kita. [2001] *3d simultaneous localisation and map building using active vision for a robot moving on undulating terrain*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

M. Pellkofer and E. Dickmanns. [2000] *EMS-vision: Gaze control in autonomous vehicles*, in Proceedings of the IEEE Intelligent Vehicles Symposium 2000.

T. Arbel and F. P. Ferrie. [1999] *Viewpoint Selection by Navigation through Entropy Maps*, in ICCV, pp. 248-254.

V. Lippiello, B. Siciliano and L. Villani. [2003]. *Robust Visual Tracking Using a Fixed Multi-camera System*, in Proceedings of the 2003 IEEE International Conference on Robotics & Automation, Taipei, Taiwan.

J. Seara and G. Schmidt. [2005] *Gaze control strategy for vision-guided humanoid walking*, at Automatisierungstechnik, vol. 2, pp. 49-58.

T. Ahmedali and J. Clark. [2006] *Collaborative multi-camera surveillance with automated person detection*, in Proceedings of the 3rd Canadian Conference on Computer and Robot Vision (CRV'06).

B. Keyes, R. Casey, H. A. Yanco, B. A. Maxwell and Y. Georgiev. [2006] *Camera Placement and Multi-Camera Fusion for Remote Robot Operation*, in Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics, Gaithersburg, MD, 2006.

J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale and S. Shafer. [2000] *Multi-camera multi-person tracking for EasyLiving*, in Proceedings of the 3rd IEEE International Workshop on Visual Surveillance, 2000.

K. Kühnlenz. [2006] *Aspects of Multi-Focal Vision*, PhD thesis, Technische Universität München.

O. Lorch, J. F. Seara, K. H. Strobl, U. D. Hanebeck and G. Schmidt. [2002] *Perception Errors in Vision Guided Walking: Analysis, Modeling, and Filtering*, Proceedings of the 2002 IEEE International Conference on Robotics and Automation, Washington, DC, May, 2002.

T. Xu, K. Kühnlenz and M. Buss. [2008] *Look at the Surprise: Bottom-Up Attentional Control of An Active Camera System*, Proceedings of the IEEE 10th International Conference on Control, Automation, Robotics and Vision (ICARCV) 2008, to appear.

J. Seara. [2004] *Intelligent Gaze Control for Vision-Guided Humanoid Walking*, PhD thesis, Technische Universität München.

T. Xu, K. Kühnlenz and M. Buss. [2008] *A View Direction Planning Strategy for a Multi-Camera Vision System*, in Proceedings of the IEEE International Conference of Information and Automation, ZhangJiaJie, 2008.

K. Kühnlenz, M. Bachmayer and M. Buss, [2006] *A Multi-Focal High-Performance Vision System*. In Proceedings of the International Conference of Robotics and Automation (ICRA), pp. 150-155, Orlando, USA, May 2006.

G. Lidoris, K. Klasing, A. Bauer, T. Xu, K. Kühnlenz, D. Wollherr and M. Buss. *The Autonomous City Explorer Project: Aims and System Overview*. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2007.