We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Multi-Focal Visual Servoing Strategies

Kolja Kühnlenz and Martin Buss

Institute of Automatic Control Engineering (LSR), Technische Universität München
Germany

1. Introduction

Multi-focal vision provides two or more vision devices with different fields of view and measurement accuracies. A main advantage of this concept is a flexible allocation of these sensor resources accounting for the current situational and task performance requirements. Particularly, vision devices with large fields of view and low accuracies can be used together. Thereby, a coarse overview of the scene is provided, e.g. in order to be able to perceive activities or structures of potential interest in the local surroundings. Selected smaller regions can be observed with high-accuracy vision devices in order to improve task performance, e.g. localization accuracy, or examine objects of interest. Potential target systems and applications cover the whole range of machine vision from visual perception over active vision and vision-based control to higher-level attention functions.

This chapter is concerned with multi-focal vision on the vision-based feedback control level. Novel vision-based control concepts for multi-focal active vision systems are presented. Of particular interest is the performance of multi-focal approaches in contrast to conventional approaches which is assessed in comparative studies on selected problems.

In vision-based feedback control of the active vision system pose, several options to make use of the individual vision devices of a multi-focal system exist: a) only one of the vision devices is used at a time by switching between the vision devices, b) two or more vision devices are used at the same time, or c) the latter option is combined with individual switching of one or several of the devices. Major benefit of these strategies is an improvement of the control quality, e.g. tracking performance, in contrast to conventional methods. A particular advantage of the switching strategies is the possible avoidance of singular configurations due to field of view limitations and an instantaneous improvement of measurement sensitivity which is beneficial near singular configurations of the visual controller and for increasing distances to observed objects. Another advantage is the possibility to dynamically switch to a different vision device, e.g. in case of sensor breakdown or if the one currently active is to be used otherwise.

The chapter is organized as follows: In Section 2 the general configuration, application areas, data fusion approaches, and measurement performance of multi-focal vision systems are discussed; the focus of Section 3 are vision-based strategies to control the pose of multi-focal active vision systems and comparative evaluation studies assessing their performance in contrast to conventional approaches; conclusions are given in Section 4.
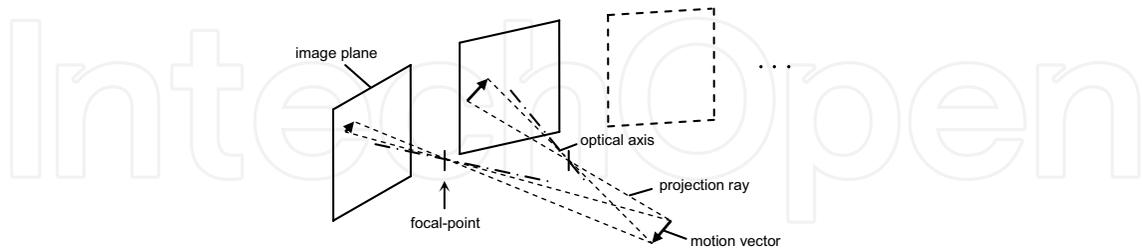
Figure 1. Schematical structure of a general multi-focal vision system consisting of several vision devices with different focal-lengths; projections of a Cartesian motion vector into the image planes of the individual vision devices

2. Multi-Focal Vision

2.1 General Vision System Structure

A multi-focal vision system comprises several vision devices with different fields of view and measurement accuracies. The field of view and accuracy of an individual vision device is mainly determined by the focal-length of the optics in good approximation and by the size and quantization (pixel sizes) of the sensor-chip. Neglecting the gathered quantity of light, choosing a finer quantization has approximately the same effect as choosing a larger focal-length. Therefore, sensor quantization is considered fixed and equal for all vision devices in this chapter. The projections of an environment point or motion vector on the image planes of the individual vision devices are scaled differently depending on the respective focal-lengths. Figure 1 schematically shows a general multi-focal vision system configuration and the projections of a motion vector.

2.2 Systems and Applications

Cameras consisting of a CCD- or CMOS-sensor and lens or mirror optics are the most common vision devices used in multi-focal vision. Typical embodiments of multi-focal vision systems are *foveated* (bi-focal) systems of humanoid robots with two different cameras combined in each eye which are aligned in parallel, e.g. (Brooks et al., 1999; Ude et al., 2006; Vijayakumar et al., 2004). Such systems are the most common types of multi-focal systems. Systems for ground vehicles, e.g. (Apostoloff & Zelinsky, 2002; Maurer et al., 1996) are another prominent class whereas the works of (Pellkofer & Dickmanns, 2000) covering situation-dependent coordination of the individual vision devices are probably the most advanced implementations known. An upcoming area are surveillance systems which strongly benefit from the combination of large scene overview and selective observation with high accuracy, e.g. (Bodor et al., 2004; Davis & Chen, 2003; Elder et al., 2004; Jankovic & Naish, 2005; Horaud et al., 2006).

An embodiment with independent motion control of three vision devices and a total of 6 degrees-of-freedom (DoF) is the camera head of the humanoid robot *LOLA* developed at our laboratory which is shown in Figure 2, cf. e.g. (Kühnlenz et al., 2006). It provides a flexible allocation of these vision devices and, due to directly driven gimbals, very fast camera saccades outperforming known systems.

Most known methods for active vision control in the field of multi-focal vision are concerned with decision-based mechanisms to coordinate the view direction of a telephoto vision device based on evaluations of visual data of a wide-angle device. For a survey on existing methods cf. (Kühnlenz, 2007).
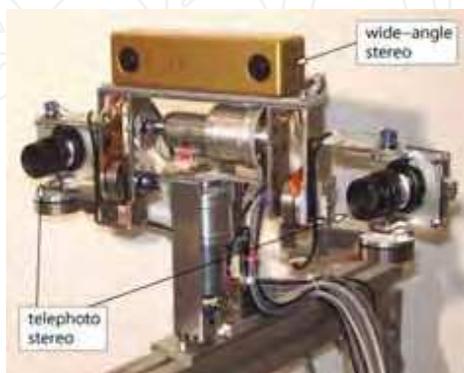

Figure 2. Multi-focal vision system of humanoid *LOLA* (Kühnlenz et al., 2006)

2.3 Fusion of Multi-Focal Visual Data

Several options exist in order to fuse the multi-resolution data of a multi-focal vision system: on pixel level, range-image or 3D representation level, and on higher abstraction levels, e.g. using prototypical environment representations. Each of these is covered by known literature and a variety of methods are known. However, most works do not explicitly account for multi-focal systems. The objective of the first two options is the 3D reconstruction of Cartesian structures whereas the third option may also cover higher-level information, e.g. photometric attributes, symbolic descriptors, etc.

The fusion of the visual data of the individual vision devices on pixel level leads to a common multiple view or multi-sensor data fusion problem for which a large body of literature exists, cf. e.g. (Hartley & Zisserman, 2000; Hall & Llinas, 2001). Common tools in this context are, e.g., projective factorization and bundle adjustment as well as multi-focal tensor methods (Hartley & Zisserman, 2000). Most methods allow for different sensor characteristics to be considered and the contribution of individual sensors can be weighted, e.g. accounting for their accuracy by evaluating measurement covariances (Hall & Llinas, 2001).

In multi-focal vision fusion of range-images requires a representation which covers multiple accuracies. Common methods for fusing range-images are surface models based on triangular meshes and volumetric models based on voxel data, cf. e.g. (Soucy & Laurendeau, 1992; Dorai et al., 1998; Sagawa et al., 2001). Fusion on raw range-point level is also common, however, suffers from several shortcomings which render such methods less suited for multi-focal vision, e.g. not accounting for different measurement accuracies. Several steps have to be accounted for: detection of overlapping regions of the images, establishment of correspondences in these regions between the images, integration of corresponding elements in order to obtain a seamless and nonredundant surface or volumetric model, and reconstruction of new patches in the overlapping areas. In order to optimally integrate corresponding elements, the different accuracies have to be considered (Soucy & Lauredau, 1995), e.g. evaluating measurement covariances (Morooka &

Nagahashi, 2006). The measurement performance of multi-focal vision systems has recently been investigated by (Kühnlenz, 2007).

2.4 Measurement Performance of Multi-Focal Vision Systems

The different focal-lengths of the individual vision devices result in different abilities (sensitivities) to resolve Cartesian information. The combination of several vision devices with different focal-lengths raises the question on the overall measurement performance of the total system. Evaluation studies for single- and multi-camera configurations with equal vision device characteristics have been conducted by (Nelson & Khosla, 1993) assessing the overall sensitivity of the vision system. Generalizing investigations considering multi-focal vision system configurations and first comparative studies have recently been conducted in our laboratory (Kühnlenz, 2007).
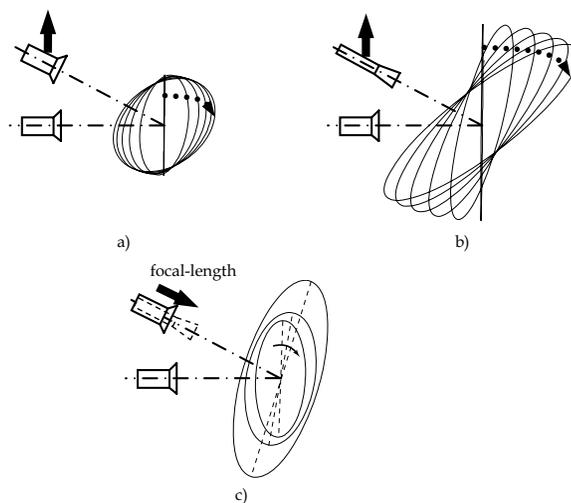

Figure 3. Qualitative change of approximated sensitivity ellipsoids of a two-camera system observing a Cartesian motion vector as measures to resolve Cartesian motion; a) two wide-angle cameras and b) a wide-angle and a telephoto camera with increasing stereo-base, c) two-camera system with fixed stereo-base and increasing focal-length of upper camera

The multi-focal image space can be considered composed of several subspaces corresponding to the image spaces of the individual vision devices. The sensitivity of the multi-focal mapping of Cartesian to image space coordinates can be approximated by an ellipsoid. Figure 3a and 3b qualitatively show the resulting sensitivity ellipsoids in Cartesian space for a conventional and a multi-focal two-camera system, respectively, with varied distances between the cameras. Two main results are pointed out: Increasing the focal-length of an individual vision device results in larger main axes of the sensitivity ellipsoid and, thus, in improved resolvability in Cartesian space. This improvement, however, is nonuniform in the individual Cartesian directions resulting in a weaker conditioned mapping of the multi-focal system. Another aspect shown in Figure 3c is an additional rotation of the ellipsoid with variation of the focal-length of an individual vision device. This effect can also be exploited in order to achieve a better sensitivity in a particular direction if the camera poses are not variable.

In summary, multi-focal vision provides a better measurement sensitivity and, thus, a higher accuracy, but a weaker condition than conventional vision. These findings are fundamental aspects to be considered in the design and application of multi-focal active vision systems.

3. Multi-Focal Active Vision Control

3.1 Vision-Based Control Strategies

Vision-based feedback control, also called visual servoing, refers to the use of visual data within a feedback loop in order to control a manipulating device. There is a large body of literature which is surveyed in a few comprehensive review articles, e.g. cf. (Chaumette et al., 2004; Corke, 1994; Hutchinson et al., 1996; Kragic & Christensen, 2002). Many applications are known covering, e.g., basic object tracking tasks, control of industrial robots, and guidance of ground and aerial vehicles.

Most approaches are based on geometrical control strategies using inverse kinematics of robot manipulator and vision device. Manipulator dynamics are rarely considered. A commanded torque is computed from the control error in image space projected into Cartesian space by the image Jacobian and a control gain.

Several works on visual servoing with more than one vision device allow for the use of several vision devices differing in measurement accuracy. These works include for instance the consideration of multiple view geometry, e.g. (Hollighurst & Cipolla, 1994; Nelson & Khosla, 1995; Cowan, 2002) and eye-in-hand/eye-to-hand cooperation strategies, e.g. (Flandin et al., 2000; Lipiello et al., 2005). A more general multi-camera approach is (Malis et al., 2000) introducing weighting coefficients of the individual sensors to be tuned according to the multiple sensor accuracies. However, no method to determine the coefficients is given. Control in invariance regions is known resulting in independence of intrinsic camera parameters and allowing for visual servoing over several different vision devices, e.g. (Hager, 1995; Malis, 2001). The use of zooming cameras for control is also known, e.g. (Hayman, 2000; Hosoda et al., 1995), which, however, cannot provide both, large field of view and high measurement accuracy, at the same time.

Multi-focal approaches to visual servoing have recently been proposed by our laboratory in order to overcome common drawbacks of conventional visual servoing (Kühnlenz & Buss, 2005; Kühnlenz & Buss, 2006; Kühnlenz, 2007). Main shortcomings of conventional approaches are dependency of control performance on distance between vision device and observed target and limitations of the field of view. This chapter discusses three control strategies making use of the individual vision devices of a multi-focal vision system in various ways. A switching strategy dynamically selects a particular vision device from a set in order to satisfy conditions on control performance and/or field of view, thereby, assuring a defined performance over the operating distance range. This sensor switching strategy also facilitates visual servoing if a particular vision device has to be used for other tasks or in case of sensor breakdown. A second strategy introduces vision devices with high accuracy observing selected partial target regions in addition to wide-angle devices observing the remaining scene. The advantages of both sensor types are combined: increase of sensitivity resulting in improved control performance and the observation of sufficient features in order to avoid singularities of the visual controller. A third strategy combines both strategies allowing independent switches of individual vision devices simultaneously observing the scene. These strategies are presented in the following sections.

3.2 Sensor Switching Control Strategy

A multi-focal active vision system provides two or more vision devices with different measurement accuracies and fields of view. Each of these vision devices can be used in a feedback control loop in order to control the pose of the active vision system evaluating visual information. A possible strategy is to switch between these vision devices accounting for requirements on control performance and field of view or other situation-dependent conditions. This strategy is discussed in the current section.
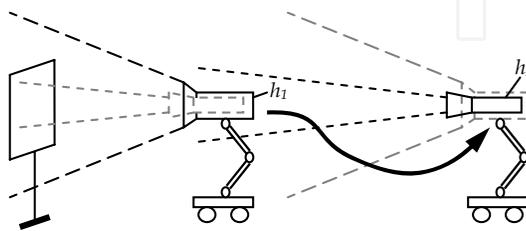
Figure 4. Visual servoing scenario with multi-focal active vision system consisting of a wide-angle camera (h_1) and a telephoto camera (h_2); two vision system poses with switch of active vision device

The proposed sensor switching control strategy is visualized in Figure 5. Assumed is a physical vision device mapping observed feature points concatenated in vector r to an image space vector ξ

$$\xi = h(r, x(q)),\tag{1}$$

at some Cartesian sensor pose x relative to the observed feature points which is dependent on the joint angle configuration q of the active vision device. Consider further a velocity relationship between image space coordinates ξ and joint space coordinates q

$$\dot{\xi}(q) = J(\xi(q), \dot{q})\dot{q},\tag{2}$$

with differential kinematics $J=J_vRJ_g$ corresponding to a particular combination of vision device and manipulator, visual Jacobian J_v, matrix $R=diag(R_c,...,R_c)$ with rotation matrix R_c of camera frame with respect to robot frame, and the geometric Jacobian of the manipulator J_g, cf. (Kelly et al., 2000). A common approach to control the pose of an active vision system evaluating visual information is a basic resolved rate controller computing joint torques from a control error ξ^d-$\xi(t)$ in image space in combination with a joint-level controller

$$\tau = J^+ K_p (\xi^d - \xi) - K_v \dot{q} + g,\tag{3}$$

with positive semi-definite control gain matrices K_p and K_v, a desired feature point configuration ξ^d, joint angles q, gravitational torques g, and joint torques τ. The computed torques are fed into the dynamics of the active vision system which can be written

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + g(q) = \tau,\tag{4}$$

with the inertia matrix M and C summarizing Coriolis and friction forces, gravitational torques g, joint angles q, and joint torques τ.

Now consider a set of n vision devices $\mathcal{H}=\{h_1,h_2,\ldots,h_n\}$ mounted on the same manipulator and the corresponding set of differential kinematics $J=\{J_1,J_2,\ldots,J_n\}$. An active vision controller is proposed which substitutes the conventional visual controller by a switching controller

$$\tau = J^{\eta^+} K_p (\xi^d - \xi) - K_v \dot{q} + g \,, \tag{5}$$

with a switched tuple of vision device h^η and corresponding differential kinematics J^η

$$< J^\eta \in J, h^\eta \in \mathcal{H} >, \ \eta \in \{1,2,\ldots n\} \,, \tag{6}$$

selected from the sets J and $\mathcal{H}$.
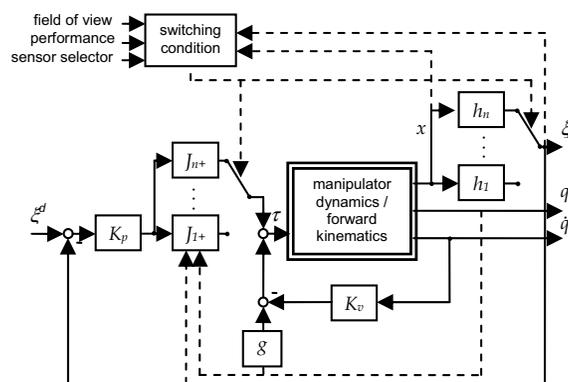
Figure 5. Block diagram of multi-focal switching visual servoing strategy; vision devices are switched directly or by conditions on field of view and/or control performance

This switching control strategy has been shown locally asymptotically stable by proving the existence of a common Lyapunov function under the assumption that no parameter perturbations exist (Kühnlenz, 2007). In case of parameter perturbations, e.g. focal-lengths or control gains are not known exactly, stability can be assured by, e.g., invoking multiple Lyapunov functions and the dwell-time approach (Kühnlenz, 2007).

A major benefit of the proposed control strategy is the possibility to dynamically switch between several vision devices if the control performance decreases. This is, e.g., the case at or near singular configurations of the visual controller. Most important cases are the exceedance of the image plane limits by observed feature points and large distances between vision device and observed environmental structure. In these cases a vision device with a larger field of view or a larger focal-length, respectively, can be selected.

Main conditions for switching of vision devices and visual controller may consider requirements on control performance and field of view. A straight forward formulation dynamically selects the vision device with the highest necessary sensitivity in order to provide a sufficient control performance, e.g. evaluating the pose error variance, in the current situation. As a side-condition field of view requirements can be considered, e.g. always selecting the vision device providing sufficient control performance with maximum field of view. Alternatively, if no measurements of the vision device pose are available the sensitivity or condition of the visual controller can be evaluated. A discussion of selected switching conditions is given in (Kühnlenz, 2007).

3.3 Comparative Evaluation Study of Sensor Switching Control Strategy

The impact of the proposed switching visual servoing strategy on control performance is evaluated in simulations using a standard trajectory following task along the optical axis. The manipulator dynamics are modeled as a simple decoupled mass-damper-system. Manipulator geometry is neglected. Joint and Cartesian spaces are, thus, equivalent. The manipulator inertia matrix is M=0.05diag(1kg, 1kg, 1kg, 1kgm², 1kgm², 1kgm²) and matrices K_v+C=0.2diag(1kgs⁻¹, 1kgs⁻¹, 1kgs⁻¹, 1kgms⁻¹, 1kgms⁻¹, 1kgms⁻¹). The control gain K_p is set such that the system settles in 2s for a static ξ^d. A set of three sensors with different focal-lengths of $\mathcal{H}$={10mm, 20mm, 40mm} and a set of corresponding differential kinematics J={J_1, J_2, J_3} based on the visual Jacobian are defined. The vision devices are assumed coincident. A feedback quantization of 0.00001m and a sensor noise power of 0.00001²m² are assumed. A square object is observed with edge lengths of 0.5m at an initial distance of 1m from the vision system. The desired trajectory is

$$x^d(t) = \begin{bmatrix} 0 & 0 & \dfrac{7}{2}\sin\left(\dfrac{1}{5}t - \dfrac{\pi}{2}\right) - \dfrac{7}{2} & 0 & 0 & \dfrac{1}{5}t \end{bmatrix}^T, \tag{7}$$

with a sinusoidal translation along the optical axes and a uniform rotation around the optical axes. The corresponding desired feature point vector ξ^d is computed using a pinhole camera model.
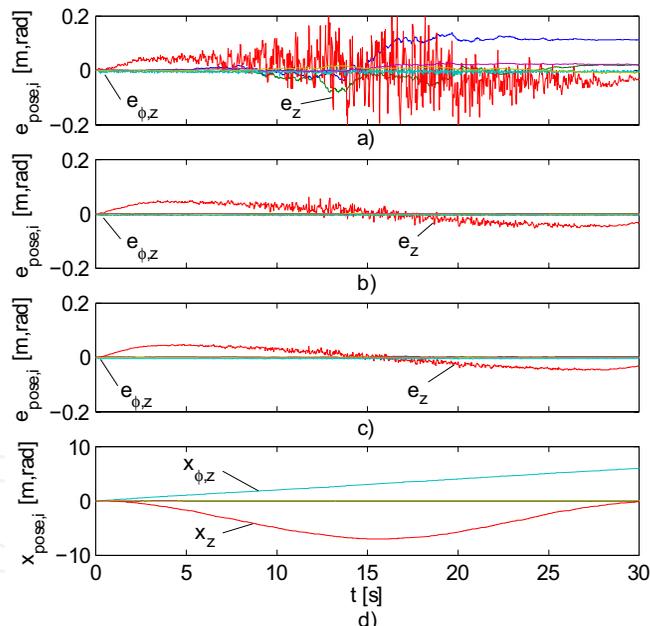

Figure 6. Tracking errors $e_{pose,i}$ and trajectory $x_{pose,i}$ of visual servoing trajectory following task; sinusoidal translation along optical (x_z-)axis with uniform rotation ($x_{\phi,z}$) ; focal-lengths a) 10mm, b) 20mm, c) 40mm

For comparison the task is performed with each of the vision devices independently and afterwards utilizing the proposed switching strategy. A switching condition is defined with

a pose error variance band of $\sigma^2 = 6.25 \cdot 10^{-6} \text{m}^2$ and a side-condition to provide a maximum field of view. Thus, whenever this variance band is exceeded the next vision device providing the maximum possible field of view is selected.
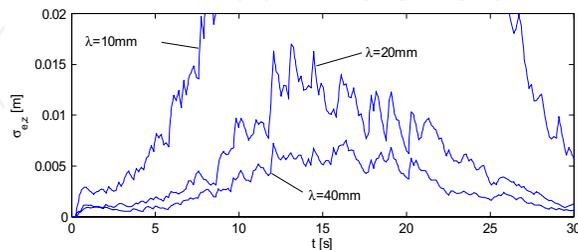

Figure 7. Corresponding tracking error standard deviation estimates for trajectory following tasks (Figure 6) with different cameras; three samples estimation window
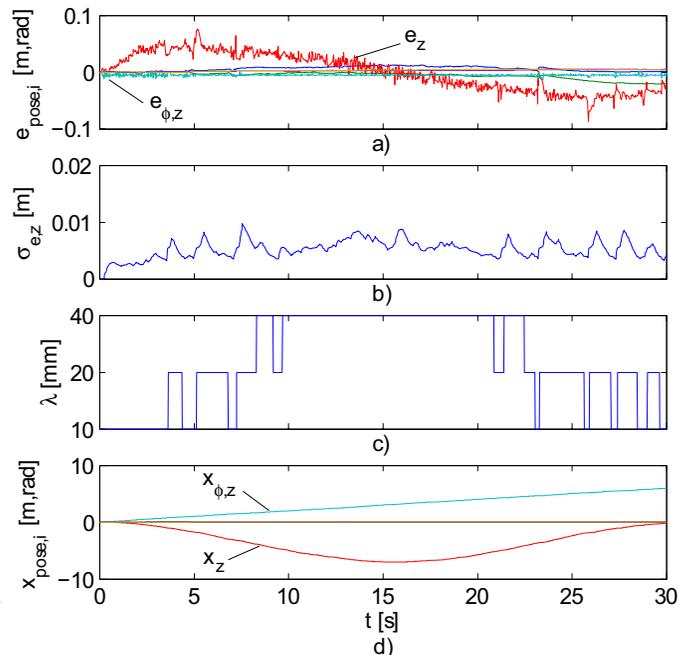

Figure 8. Results of sensor switching visual servoing strategy with multi-focal vision; sinusoidal translation along optical (x_z-)axis with uniform rotation ($x_{\phi,z}$); a) tracking errors, b) tracking error standard deviation estimates, c) current focal-length, d) pose trajectory

Figure 6 shows the resulting tracking errors for the trajectory following task for each of the individual vision devices. In spite of very low control error variances in image space of about 0.01 pixels² large pose error variances in Cartesian space can be noted which vary over the whole operating distance as shown in Figure 7. The distance dependent sensitivity of the visual controller and quantization effects result in varying pose error variances over the operating range caused by sensor noise. These effects remain a particular problem for wide range visual servoing rendering conventional visual servoing strategies unsuitable.

Figure 8 shows the results of the switching control strategy. The standard deviation (Figure 8b) is kept within a small band reaching from about 0.004m to 0.008m. The overall variability is significantly lower compared to the single-camera tasks (Figure 7). The spikes, which can be noted in the standard deviation diagram, are caused by the switches due to the delay of the feedback signal. After a switch the desired feature value changes with the sensor, but the current value is still taken from the previous sensor. Thus, the control error at this time instance jumps. This effect can be reduced by mapping the previous value of the feature vector to the image space of the new sensor or by definition of a narrower variance band as switching condition.

Figure 9 exemplarily illustrates the progression of the fields of view over time for a uniform single-camera translation task and the corresponding camera switching task. The field of view is defined by the visible part of the plane extending the surface of the observed object in *x*-direction. The variability achieved with the switching strategy is significantly lower.

The effectiveness of the proposed multi-focal switching strategy has been shown successfully. The contributions of this novel approach are a guaranteed control performance by means of a bounded pose error variance, a low variability of the performance over the whole operating range, and the consideration of situational side-conditions as, e.g., a maximum field of view.

3.4 Multi-Camera Control Strategy

If two or more vision devices of a multi-focal system are available simultaneously these devices can be used together in order to control the pose of the vision system. In this section a multi-focal multi-camera strategy is proposed in order to make use of several available vision devices with different fields of view and measurement accuracies. Major benefit is an improved control performance compared to single-camera strategies whereas only a partial observation of the reference object with high accuracy is necessary.
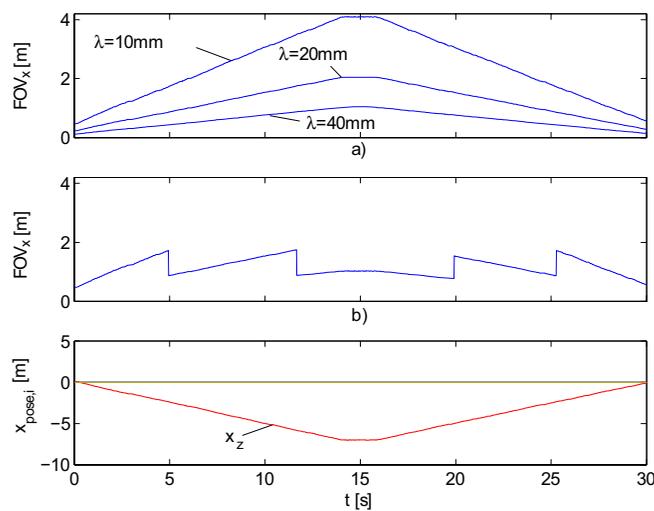

Figure 9. Progression of the extension of the field of view orthogonal to the optical axis of the observing vision device; uniform translation along optical (x_z-)axis; a) single-camera tasks, b) sensor switching strategy with multi-focal vision, c) pose trajectory

A vision-based controller computing joint torques from a control error in image space requires sufficient observed feature points to be mapped to the six Cartesian degrees of freedom. A minimum of three feature points composed of two elements in image space is needed in order to render the controller full rank. If the field of view of the observing vision device is too small to cover all feature points the controller becomes singular. However, high-sensitivity sensors needed in order to achieve high control performance only provide small fields of view.

A multi-camera strategy is proposed combining the advantages of vision devices with different characteristics. High-sensitivity devices are used for improving control performance and wide-field-of-view devices in order to observe the required number of remaining feature points to render the controller full rank.
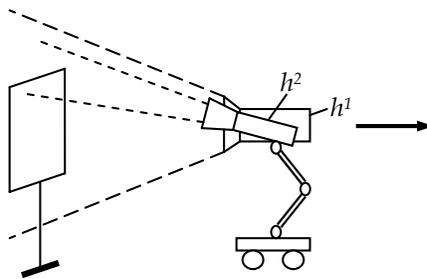

Figure 10. Visual servoing scenario with multi-focal active vision system consisting of a wide-angle camera (h^1) and a telephoto camera (h^2); both vision devices are observing different feature points of a reference object accounting for field of view constraints

The sensor equation (1) extends such that individual feature points are observed with different vision sensors

$$\begin{bmatrix} \xi_1^{\;T} & \cdots & \xi_i^{\;T} & \xi_j^{\;T} & \cdots \end{bmatrix}^T = \begin{bmatrix} h_1\left(\begin{bmatrix} r_1^{\;T} & \cdots & r_i^{\;T} \end{bmatrix}^T, x_1(q) \right) & h_2\left(\begin{bmatrix} r_j^{\;T} & \cdots \end{bmatrix}^T, x_2(q) \right) & \cdots \end{bmatrix}^T, \quad (8)$$

where a Cartesian point r_k is mapped to an image point ξ_l by vision device h_m. The proposed visual controller is given by

$$\tau = \begin{bmatrix} J_1^{\;T} & \cdots & J_1^{\;T} & J_2^{\;T} & \cdots \end{bmatrix}^{T^+} K_p(\xi^d - \xi) - K_v \dot{q} + g , \quad (9)$$

with image feature vector $\xi = [\; \xi_1 \; ... \; \xi_i \; \xi_j \; ...]^T$ and differential kinematics J_m corresponding to vision device h_m.

Substituting the composition of individual differential kinematics J_m by a generalized differential kinematics J^* the proposed control strategy can be expressed by

$$\tau = J^{*+} K_p(\xi^d - \xi) - K_v \dot{q} + g , \quad (10)$$

which has been proven locally asymptotically stable (Kelly et al., 2000).

Utilizing the proposed multi-camera strategy an improved control performance is achieved even though only parts of the observed reference structure are visible for the high-sensitivity vision devices. This multi-camera strategy can be combined with the switching

strategy discussed in Section 3.2 allowing switches of the individual vision devices of a multi-focal vision system. Such a multi-camera switching strategy is discussed in the following section.

3.5 Multi-Camera Switching Control Strategy

In the previous sections two concepts to make use of the individual vision devices of a multi-focal vision system have been presented: a sensor switching and a multi-camera vision-based control strategy. This section proposes the integration of both strategies, thus, allowing switches of one or more vision devices observing parts of a reference structure simultaneously. Thereby, the benefits of both strategies are combined.

The sensor equation (8) is extended writing

$$\begin{bmatrix} \xi_1^T & \cdots & \xi_i^T & \xi_j^T & \cdots \end{bmatrix}^T = \begin{bmatrix} h_1^{\eta}\left(\begin{bmatrix} r_1^T & \cdots & r_i^T \end{bmatrix}^T, x_1(q)\right) & h_2^{\eta}\left(\begin{bmatrix} r_j^T & \cdots \end{bmatrix}^T, x_2(q)\right) & \cdots \end{bmatrix}^T, \quad (11)$$

allowing the h_m^{η} of (8) to be selected dynamically from a set $\mathcal{H}=\{h_1,h_2,\ldots,h_n\}$. The visual controllers (5) and (10) are integrated writing

$$\tau = J^{\eta*+} K_p(\xi^d - \xi) - K_v \dot{q} + g, \quad (12)$$

where $J^{\eta*}$ is composed of individual differential kinematics J^m

$$J^{\eta*+} = \begin{bmatrix} J_1^{\eta T} & \cdots & J_1^{\eta T} & J_2^{\eta T} & \cdots \end{bmatrix}^T, \quad (13)$$

which are selected dynamically from a set $J=\{J_1,J_2,\ldots,J_n\}$ of differential kinematics corresponding to the set $\mathcal{H}$ of available vision devices.

In the following section the proposed multi-camera strategies are exemplarily evaluated in a standard visual servoing scenario.

3.6 Comparative Evaluation Study of Multi-Camera Control Strategies

In this section a comparative evaluation study is conducted in order to demonstrate the benefits of the proposed multi-camera and multi-camera switching strategies. Considered is again a trajectory following task with a uniform translation along the optical axis of a main camera with a wide field of view (focal-length 5mm) as shown in Figure 10. A square reference object is observed initially located at a distance of 1m to the camera. A second camera observes only one feature point of the object. The characteristics of this camera are switchable. Either the same characteristics as of the wide-angle camera or telephoto characteristics (focal-length 40mm) are selectable. The inertia matrix is set to M=0.5diag(1kg, 1kg, 1kg, 1kgm², 1kgm², 1kgm²) and matrices K_v+C=200diag(1kgs⁻¹, 1kgs⁻¹, 1kgs⁻¹, 1kgms⁻¹, 1kgms⁻¹, 1kgms⁻¹). The other simulation parameters are set equal to section 3.3.

Three simulation scenarios are compared: second camera with wide-angle characteristics, with telephoto characteristics, and switchable. Switches of the second camera are allowed after a time of 2s when a constant tracking error is achieved. A switch is performed when the tracking error standard deviation exceeds a threshold of 0.00004m.

Figure 11 shows the tracking error of the uniform trajectory following task with switched second camera which can be considered constant after about 2s. Figure 12 shows the resulting standard deviations of the tracking error for all three tasks. It can be noted that a

lower standard deviation is achieved by the multi-camera task (second camera with telephoto characteristics) compared to the wide-angle task. The multi-camera switching task additionally achieves a lower variability of the standard deviation of the tracking error.
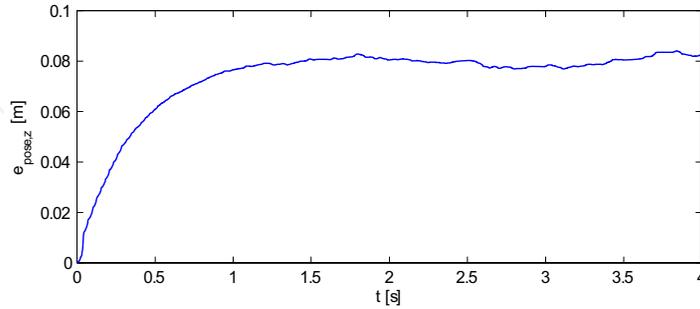

Figure 11. Tracking error of multi-focal two-camera visual servoing task with wide-angle and switchable wide-angle/telephoto camera; desired trajectory x_z^d(t)=-0.2ms^{-1}t-1m
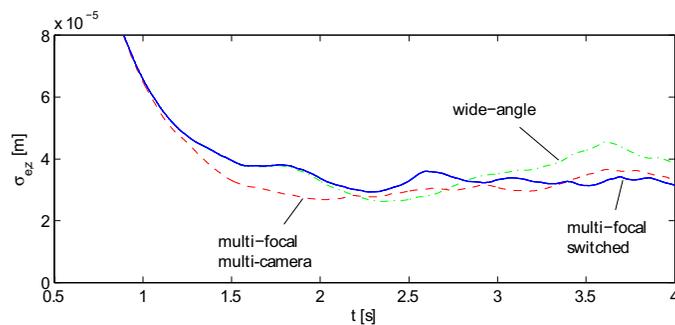

Figure 12. Standard deviation estimates of tracking error of unswitched single-camera task (wide-angle), of unswitched multi-focal multi-camera task with one feature point observed by additional telephoto camera, and of switched multi-focal multi-camera task with additional camera switching from wide-angle to telephoto characteristics at t=2.6s
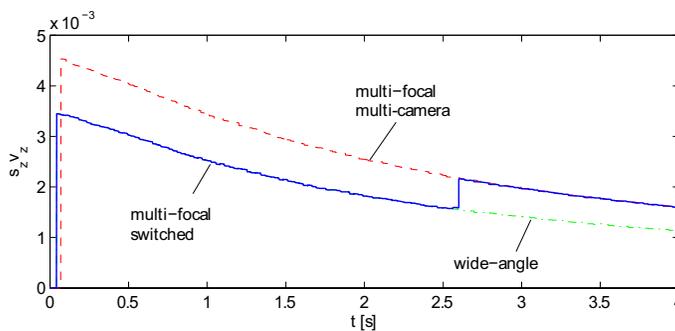
Figure 13. Sensitivities of the visual servoing controller along the optical axis of the central wide-angle camera corresponding to the tasks in Figure 12

Figure 13 shows the sensitivity ($s_z v_z$) of the visual controller for all three tasks along the optical axis of the wide-angle camera. It can be noted that the multi-camera strategies result in a better sensitivity of the controller compared to the wide-angle task.

Summarized, the simulations clearly show the benefits of the proposed multi-camera control strategies for multi-focal vision systems: an exploitation of the field of view and sensitivity characteristics in order to achieve improved control performance and a lower variability of the performance by switching of individual vision devices.

4. Conclusion

In this chapter novel visual servoing strategies have been proposed based on multi-focal active vision systems able to overcome common drawbacks of conventional approaches: a tradeoff between field of view and sensitivity of vision devices and a large variability of the control performance due to distance dependency and singular configurations of the visual controller. Several control approaches to exploit the benefits of multi-focal vision have been proposed and evaluated in simulations: Serial switching between vision devices with different characteristics based on performance- and field-of-view-dependent switching conditions, usage of several of these vision devices at the same time observing different parts of a reference structure, and individual switching of one or more of these simultaneously used sensors. Stability has been discussed utilizing common and multiple Lyapunov functions.

It has been shown that each of the proposed strategies significantly improves the visual servoing performance by reduction of the pose error variance. Depending on the application scenario several guidelines for using multi-focal vision can be given. If only one vision sensor at a time is selectable then a dynamical sensor selection satisfying desired performance constraints and side-conditions is proposed. If several vision sensors can be used simultaneously selected features of a reference object can be observed with high-sensitivity sensors while a large field of view sensor ensures observation of a sufficient number of features in order to render the visual controller full rank. The high-sensitivity sensors should preferably be focused on those feature points resulting in the highest sensitivity of the controller.

5. Acknowledgments

6. References

Apostoloff, N. & Zelinsky, A. (2002). Vision in and out of vehicles: Integrated driver and road scene monitoring, *Proceedings of the 8th International Symposium on Experimental Robotics (ISER)*, 2002, Sant Angelo d'Iscia, Italy

Bodor, R.; Morlok, R. & Papanikolopoulos, N. (2004). Dual-camera system for multi-level activity recognition, *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2004, Sendai, Japan

Brooks, R. A.; Breazeal, C.; Marjanovic, M.; Scasselati, B. & Williamson, M. M. (1999). The Cog Project: Building a Humanoid Robot, In: *Computation for Methaphors, Analogy, and Agents*, C. Nehaniv, (Ed.), Springer, Germany

Chaumette, F.; Hashimoto, K.; Malis, E. & Martinet, P. (2004). TTP4: Tutorial on Advanced Visual Servoing, Tutorial Notes, IEEE/RSJ IROS, 2004

Corke, P. I. (1994). Visual Control of Robot Manipulators – A Review, In: *Visual Servoing*, K. Hashimoto, (Ed.), World Scientific, 1994

Cowan, N. (2002). Binocular visual servoing with a limited field of view, In: *Mathematical Theory of Networks and Systems*, Notre Dame, IN, USA, 2002

Dickmanns, E. D. (2003). An advanced vision system for ground vehicles, *Proceedings of the International Workshop on In-Vehicle Cognitive Computer Vision Systems (IVC2VS)*, 2003, Graz, Austria

Dorai, C.; Wang, G.; Jain, A. K. & Mercer, C. (1998). Registration and Integration of Multiple Object Views for 3D Model Construction, In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 1, 1998

Elder, J. H.; Dornaika, F.; Hou, B. & Goldstein, R. (2004). Attentive wide-field sensing for visual telepresence and surveillance, In: *Neurobiology of Attention*, L. Itti, G. Rees & J. Tsotsos, (Eds.), 2004, Academic Press, Elsevier

Flandin, G.; Chaumette, F. & Marchand, E. (2000). Eye-in-hand/eye-to-hand cooperation for visual servoing, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2003

Hager, G. D. (1995). Calibration-free visual control using projective invariance, *Proceedings of the 5th International Conference on Computer Vision (ICCV)*, 1995

Hall, D. L. & Llinas, J. (2001). *Handbook of Multisensor Data Fusion*, CRC Press, 2001, Boca Raton, FL, USA

Hartley, R. I. & Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2004, NY, USA

Hayman, E. (2000). *The use of zoom within active vision*, Ph.D. Thesis, University of Oxford, 2000, Oxford, UK

Hollighurst, N. & Cipolla, R. (1994). Uncalibrated stereo hand-eye coordination, In: *Image and Vision Computing*, Vol.12, No. 3, 1994

Horaud, R.; Knossow, D. & Michaelis, M. (2006). Camera cooperation for achieving visual attention, In: *Machine Vision and Applications*, Vol. 15, No. 6, 2006, pp. 331-342

Hosoda, K.; Moriyama, H. & Asada, M. (1995). Visual servoing utilizing zoom mechanism, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 1995

Hutchinson, S.; Hager, G. D. & Corke, P. I. (1996). A tutorial on visual servo control, In: *IEEE Transaction on Robotics and Automation*, Vol. 12, No. 5, 1996

Jankovic, N. D. & Naish, M. D. (2005). Developing a modular spherical vision system, *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1246-1251, 2005, Barcelona, Spain

Kelly, R.; Carelli, R.; Nasisi, O.; Kuchen, B. & Reyes, F. (2000). Stable visual servoing of camera-in-hand robotic systems, In: *IEEE Transactions on Mechatronics*, Vol. 5, No. 1, 2000

Kragic, D. & Christensen, H. I. (2002). *Survey on Visual Servoing for Manipulation*, Technical Report, Stockholms Universitet, ISRN KTH/NA/P-02/01-SE, CVAP259, 2002

Kühnlenz, K. & Buss, M. (2005). Towards multi-focal visual servoing, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2005

Kühnlenz, K. & Buss, M. (2006). A multi-camera view stabilization strategy, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2006

Kühnlenz, K. (2007). *Aspects of multi-focal vision*, Ph.D. Thesis, Institute of Automatic Control Engineering, Technische Universität München, 2007, Munich, Germany

Kühnlenz, K.; Bachmayer, M. & Buss, M. (2006). A multi-focal high-performance vision system, *Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA)*, 2006, Orlando, FL, USA

Lipiello, V.; Siciliano, B. & Villani, L. (2005). Eye-in-hand/eye-to-hand multi-camera visual servoing, *Proceedings of the IEEE International Conference on Decision and Control (CDC)*, 2005

Malis, E. (2001). Visual servoing invariant to changes in camera intrinsic parameters, *Proceedings of the 8th International Conference on Computer Vision (ICCV)*, 2001

Malis, E.; Chaumette, F. & Boudet, S. (2000). Multi-cameras visual servoing, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2000

Maurer, M.; Behringer, R.; Furst, S.; Thomanek, F. & Dickmanns, E. D. (1996). A compact vision system for road vehicle guidance, *Proceedings of the 13th International Conference on Pattern Recognition (ICPR)*, 1996

Morooka, K. & Nagahashi H. (2006). A Method for Integrating Range Images in Different Resolutions for 3-D Model Construction, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2006

Nelson, B. & Khosla, P. (1993). *The resolvability ellipsoid for visually guided manipulation*, Technical Report, CMU-RI-TR-93-28, The Robotics Institute, Carnegie Mellon University, 1993, Pittsburgh, PA, USA

Nelson, B. & Khosla, P. (1995). An extendable framework for expectation-based visual servoing using environment models, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 1995

Pellkofer, M. & Dickmanns, E. D. (2000). EMS-Vision: Gaze control in Autonomous vehicles, *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2000, Dearborn, MI, USA

Sagawa, R.; Nishino, K. & Ikeuchi, K. (2001). Robust and Adaptive Integration of Multiple Range Images with Photometric Attributes, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001

Soucy, M. & Laurendeau, D. (1992). Multi-Resolution Surface Modelling from Multiple Range Views, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 1992

Ude, A.; Gaskett, C. & Cheng, G. (2006). Foveated Vision Systems with Two Cameras Per Eye, *Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA)*, 2006, Orlando, FL, USA

Vijayakumar, S.; Inoue, M. & Souza, A. D. (2004). *Maveric – Oculomotor experimental vision head*, http://homepages.inf.ed.ac.uk/svijayak/projects/maveric/index.html, 2004

Vision Systems: Applications

Edited by Goro Obinata and Ashish Dutta

Computer Vision is the most important key in developing autonomous navigation systems for interaction with the environment. It also leads us to marvel at the functioning of our own vision system. In this book we have collected the latest applications of vision research from around the world. It contains both the conventional research areas like mobile robot navigation and map building, and more recent applications such as, micro vision, etc.The fist seven chapters contain the newer applications of vision like micro vision, grasping using vision, behavior based perception, inspection of railways and humanitarian demining. The later chapters deal with applications of vision in mobile robot navigation, camera calibration, object detection in vision search, map building, etc.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds