Towards an Anthropomorphic Robotical Hand-Eye Coordination*

Thomas Fink, Alexa Hauck, Georg Färber
Lab. for Process Control and Real-Time Systems
Technische Universität München, Germany
email: hauck@lpr.e-technik.tu-muenchen.de

ABSTRACT

Human grasping still outshines its robotical counterparts with respect to accuracy, speed, robustness, and flexibility. When trying to develop a robotical hand-eye system, it is therefore only natural to examine the results of neuroscience. In this paper, we examine the human hand-eye system concerning motion planning and control using robotical categories and strategies. From the results, we derive a system concept for an anthropomorphic robotical hand-eye coordination.

1. INTRODUCTION

Using sensor information to control robots has become a very popular field of research, since it promises to lead to the design of *autonomous robots*. In contrast to their preprogrammed industrial counterparts, autonomous robots must be able to deal with unexpected events such as obstacles or misplaced objects. This is especially important for *personal robots* because they operate in an environment that is not adapted to the needs of machines.

Nowadays, vision is by far the most commonly used sensor because CCD cameras are cheap and easy to use. Additionally, vision is an important human sense, and, therefore, the information received by the robot's vision system is easier for the human operator to understand.

In the field of visually controlled robot manipulators, research has mainly focussed on the control part, circumventing the problems of extracting and interpreting image features by using artificial features such as blobs or by selecting objects and features manually. This resulted in a large number of impressive servoing methods which cope very well with a specific problem (for a survey see [8], for a collection of articles on state-of-the-art results [23]). Yet, versatile hand-eye systems that could be labeled "autonomous" or "intelligent" still seem to be a distant goal.

On the other hand, human grasping still outshines its robotical counterparts with respect to accuracy, speed, robustness and flexibility. When trying to develop a robotical hand-eye system, it is therefore only natural to examine the results from neuroscience. Therefore, we started a joint research project with the *Neuropsychological Team for Skilled Motor Control* at the *Ludwig-Maximilians-Universität München, Germany*. Our goal is to use information gained from the analysis of human reaching and grasping to design and implement a robotical hand-eye system.

The principal idea is to use a set of modules, each specialized in a different phase or situation that occurs when reaching for an object. In the *vision part* of the system, such modules are charged with tasks such as object detection and recognition, pose estimation, visual tracking, and motion prediction. This modular structure is supported by studies from the neurosciences that describe selective disorders that result from accidents. Mai et al. for example, describe the case of a patient suffering from ataxia, the inability to coordinate voluntary muscular movements, who cannot grasp stationary objects but can catch moving ones [28]. Goodale and Servos present evidence that visual mechanisms for perceiving and grasping an object are functionally and neurally distinct [14].

Numerous neurological studies also suggest that objects have multiple cerebral representations that contain the information relevant for different tasks. In addition to the classical distinction between object identification ("what") and the localization of an object ("where") [30], Jeannerod gives evidence for the existence of a "pragmatic representation" that encodes all attributes of an object that are necessary for controlling a movement towards it [26]. This principle shall be employed in our robotical hand-eye system by using a hierarchical object data-base containing sensor- and task-specific information about the objects to grasp [21]. This object data base provides the *model knowledge* necessary for hand-eye coordination together with information about the system's configuration stemming from calibration processes.

In the *robotical part* of the system, which is responsible for motion planning and control of the robot manipulator, visual information must be translated into suitable control sequences for the manipulator. Here, two main approaches can be found in literature: *Open-loop systems* determine the object's pose from the visual input as accurately as possible and then move the robot appropriately ("look-then-move"). On the other hand, *visual servoing systems* use a continuous feedback of visual information to guide the robot ("look-and-move"). The common view is that these two approaches are mutually exclusive. A closer look at the results from neuroscience reveals, though, that in human grasping both strategies are combined [37].

This lead us to propose a hybrid structure for the robotical part [20]: First, the object to grasp is identified (for example by its silhouette [2]) and its rough 3D position is determined. This

*The work presented in this paper is supported by the *Deutsche Forschungsgemeinschaft* as part of the Special Research Program "Sensorimotor – Analysis of biological systems, modelling, and medical-technical application" (SFB 462).

is used to generate a trajectory that brings the gripper into the vicinity of the object. The remaining gap is closed by minimizing the distance between gripper and object in the image of a free-standing camera.

In this paper, we will extend this concept. Sec. 2 contains a survey of state-of-the-art robotical hand-eye coordination and formulates categories by which hand-eye systems can be classified. These categories are then used in Sec. 3 to interpret current knowledge about the human hand-eye system. In Sec. 4, these results are integrated into a system concept for an anthropomorphic robotical hand-eye system as a first step towards implementation.

2. ROBOTICAL HAND-EYE COORDINATION

Following to the taxonomy by Sanderson and Weiss [33], servoing architectures can be divided into four major categories according to two criteria:

- **Visual information**: The categories of *image-based* and *position-based* respectively concern the level of abstraction of visual information that is used for robot control. In the former, image features are used directly, while in the latter they are converted to information in Cartesian coordinates (pose estimation).
- **Control architecture**: According to whether an internal feedback loop for joint level control is present or not, systems fall into the categories of *dynamic control* and *direct control*.

In order to establish a more detailed distinction between hand-eye systems, we propose to use four additional criteria:

- **Hand-eye configuration**: According to whether cameras are mounted on the manipulator or not, we speak of *eye-in-hand* or *free-standing* camera systems.
- **Visual control strategy**: Visual information can be used for open-loop (*look-then-move*) or closed-loop control (*look-and-move*).
- **Focus of attention**: While *endpoint-open-loop* systems only use visual information of the object to grasp, *endpoint-closed-loop* systems observe both the target object and the robot manipulator.
- **Camera configuration**: One can distinguish *monocular* and *binocular* vision systems.

Hand-eye configuration

Robotic hand-eye systems can first of all be classified according to their physical configuration: In **eye-in-hand camera systems**, cameras are mounted on the robot arm. If the camera is directly fixed to the manipulator tool its position relative to the end-effector is known and constant. If the camera is mounted onto another limb of the robot arm, at least one joint is situated between the camera and the tool. Hence, the position is no longer constant, but depends on the current angle of the intermediate joints. Since the target object is automatically zoomed while the gripper is approaching it, the eye-in-hand camera system has the advantage of a continuously increasing resolution, thus improving its accuracy. Disadvantages of this configuration are the effects of perspective, mechanical vibrations of the manipulator, and the problem that, in the final phase of reaching, the target object may leave the field of view of the camera. Furthermore, the pose of the target relative to the camera changes while the manipulator (with the camera) is moved. Therefore, even stationary objects have to be tracked in the image. In recent years, lightweight video cameras have reduced the problem of additional weight that changes the robot's dynamics. Eye-in-hand systems are employed for example in [12, 11, 7, 5, 38, 9, 10, 11, 3].

In **free-standing camera systems** (see for example [17, 22, 25, 1]), cameras are usually fixed in the workspace, apart from the robot. Thus, the image of a stationary object is not altered by the movement of the manipulator. Contrary to the eye-in-hand camera systems, the target object can easily be kept in the field of view, yet the target may become occluded by the gripper. Possible extensions to this approach are, first, the use of cameras with zoom lenses to zoom into the target object, or secondly, to use cameras that are free-standing but not stationary. In the latter case, the cameras can be moved in a way that occlusion is avoided and spatial resolution is improved (for examples see [39, 35].

Visual control strategy

In the field of visually controlled robot manipulators, two main approaches can be distinguished: The first approach proposes a sequential structure, with an open-loop between hand and eye. The term **look-then-move** signifies the separation between the act of "looking" and the act of "moving" the gripper: first, the object's position is determined from the visual input as accurately as possible, then the robot is moved manipulator appropriately. This approach lends itself very well to integration in manufacturing, with the visual input replacing the knowledge about the exact position of the parts to handle. However, the accuracy of the operation depends heavily on the accuracy of the visual sensor, the manipulator, its controller, and on the sensor to robot calibration. A classical look-then-move hand-eye system is for example described in [3]; Allen et al. close the loop to track and grasp a moving object, yet in the case of stationary objects their system falls into the category look-then-move [1].

Most robotical hand-eye systems cited in the following fall into the category of **look-and-move** or **visual servoing** systems. They use a continuous feedback of visual information to guide the robot. In these systems, accuracy is increased not by using more refined and expensive subsystems but by closing the control loop with visual information. This approach furthermore promises to lead to calibration-free systems. However, due to the closed-loop control, the trajectory of the movement is not known in advance. Thus, the target may become invisible in the course of the motion. For a introduction to visual servoing systems see [24], recent examples can be found in [6, 9, 10, 17, 22, 25].

Control architecture

Hand-eye systems can be further distinguished by their control architecture. **Dynamic control** systems use a hierarchical structure; the robot is internally stabilized with the help of a second control loop which employs encoder feedback from the robot's internal joint angle sensors. This separation of the visual controller from the robot kinematics and dynamics permits to view the robot as an ideal Cartesian motion device which is not affected by problems like oscillations and singularities. In other words, the visual controller can assume idealized axis dynamics, because of the high sampling rates of the internal feedback loop. This simplifies the control design problem considerably. Since many robots provide an interface for Cartesian inputs or incremental position commands, implementation is simple and portable.

Direct control systems have no hierarchical control architecture. The visual servo controller takes over the job of the internal robot joint controller, computing the state of the joints directly from the visual information. However, the relatively low sampling rates of the vision process make direct visual control of a robot end-effector an extremely challenging control problem. Thus, current hand-eye systems use a hierarchical control structure.

Visual information

Having detected the target object in the video image, the question arises which visual information is to be used to control the robot. The **position-based** approach first estimates the position of the target object relative to the camera in Cartesian coordinates, based on a geometrical model of the robot, its reachable work space (task space), and the target object. The error signal for the robot controller is therefore defined in Cartesian coordinates. As most robots provide a Cartesian interface and because operating in 3D-space can be understood intuitively by the system designer, programming is facilitated. However, an exact determination of the pose of the object to grasp relative to the manipulator requires an accurate hand-eye (gripper-camera) calibration and a precise pose estimation. Look-then-move systems usually are position-based; position-based visual servoing systems can be found in [1, 10, 35, 38].

In **image-based**, also called *feature-based* systems, the error signal is defined in terms of image features, and therefore is directly measured in the image coordinate system. Therefore, computational costs are significantly reduced and the system becomes less sensitive to errors in camera calibration and system modelling. However, the computation of robot motion on the basis of image features takes place in a less intuitive projection of the task space, depending on the chosen image features and the method determining the distance between the features. As this process is non-linear and its parameters highly correlated, it presents a significant challenge to control design and has proven to be difficult to analyze theoretically. Pioneering work in the field of image-based visual servoing stems from Weiss [36]; further examples for image-based hand-eye systems can be found in [12, 7, 9, 32, 17, 22, 25].

Hybrid system which combine position-based and image-based control are described in [5, 6].

Focus of attention

Systems can be further classified regarding their focus of attention: **Endpoint open-loop** systems (EOL) only observe the target object and get no visual information of the actual position of the manipulator. Determining and controlling the position of the gripper is exclusively based on a combination of joint angle sensors, internal knowledge of the end-effector kinematics and on the camera-object calibration. Thus, the positioning accuracy of the system depends heavily on the accuracy of the calibration. Eye-in-hand systems usually employ this approach.

Endpoint closed-loop systems (ECL) compensate hand-eye calibration errors by observing both the target object and the robot end-effector. However, the simultaneous tracking of the target and the gripper places constraints on the field of view of the system. Visual servoing systems with free-standing cameras mostly use this approach.

Note, that this categorization only applies to systems that actually grasp objects; in [9, 7, 10, 12, 11, 6], systems are described that visually position a camera relative to an object.

Camera configuration

Finally, hand-eye systems differ with respect to the number of cameras employed. The principal problem when using only one camera (**monocular** system) is that depth information is lost due to the projection of the scene onto the image plane. Therefore, additional information is needed to determine the pose of an object or to guide the manipulator in 3D. This information can stem from a geometrical model of the object to grasp (see e.g. [3, 35]), from other images taken at a different time (e.g. [10, 9]), or at a different place (e.g. [19, 31]. Eye-in-hand systems usually employ only one camera because the additional weight on the manipulator changes its dynamics.

Binocular systems employ two cameras, thus are able to reconstruct the depth information which is lost in the projection of the scene on the image plane without the use of a geometrical object model. Position-based systems can determine depth via triangulation; in image-based systems, depth information is integrated implicitely by driving the error signals in two images to zero. The disadvantage of binocular systems is that feature extraction has to be performed for two images at a time; additionally, it must be assured that the features extracted in the two images correspond to the same physical features of the object. Most free-standing hand-eye systems use two cameras.

3. HUMAN HAND-EYE COORDINATION

After formulating criteria by which hand-eye systems can be classified, we now examine current models of the human hand-eye system and try to fit them into the robotical categories. Remaining on a high level of abstraction helps to avoid the problem that the human system cannot easily be copied at functional or even algorithmical level because of the different “hardware”.

Hand-eye configuration

Regarding this criterion, obviously no literature survey is necessary. The human hand-eye system employs a free-standing camera system and makes use of active vision methods such as vergence and focus control, visually guided saccades and smooth pursuit eye (and head) movements.

Visual control strategy

In 1899 already, Woodworth proposed that a reaching movement consists of two components: an *"initial impulse propelling the hand towards the target"* and a *"current control to home in on the final position via successive approximations"* [37]. The former was found to be dependent on visual information only in the beginning, to program the movement; the latter depends on primarily visual feedback during motion. In robotical terms this means that both look-then-move and look-and-move strategies are employed.

Woodworth's results stem from experiments that analyzed the accuracy of reaching movements with varying availability of visual information. In contrast, Mai and Marquardt examined the kinematical difference between what they call "automated" and "controlled" movements [27]: In an experiment originally designed to investigate writing disabilities, test persons are first required to write the letter "a" for several times with normal writing velocity. Then, the task is to redraw the letters. Under normal writing conditions, the acceleration profiles for each letter are smooth, velocity profiles are bell-shaped; when redrawing the letter, though, the acceleration profile shows recurrent acceleration and deceleration phases corresponding to a multi-peaked velocity profile.

The notion that the required accuracy of a movement affects the corresponding trajectory was supported by Milner [29] who measured the trajectories of human subjects inserting a pin into a hole. For small holes and therefore high precision requirements, the velocity profile showed small oscillations at the end corresponding to a sequence of submovements. Burdet presents a detailed model of reaching movements which explains these experimental results [4].

Control architecture

As estimates of the human "visual reaction time", that is the time after which changes in a trajectory corresponding to visual input appear, vary between $100ms$ and $250ms$, it is not surprising that models for visually controlled arm movements propose a hierarchical control architecture, thus falling into the category of dynamic control systems. In [4] for example, the internal feedback loop is explained by the visco-elasticity of the muscles.

Visual information

It is still an open question whether the human hand-eye system converts the visual information directly into motor programs or via an intermediate reference frame which is similar to Cartesian space. Many models of human reaching movements that successfully explain measured trajectories assume that the visually determined target location is given in Cartesian coordinates [4], thereby describing position-based systems. On the other hand, Stein presents evidence for image-based hand-eye coordination by describing a distributed system of transformation algorithms that directly convert sensor in motor information and vice versa [34].

Focus of attention

Goodale et al. prooved that observation of the hand is not required to be able to adjust movements in the case of target displacements [15]. They also showed that one does not even need to perceive the displacement. Yet, vision of the hand before or during motion significantly increases the accuracy of reaching movements [13].

Camera configuration

When closing one eye and trying to grasp an object, everybody will confirm that binocular vision is not a prerequisite for successful hand-eye coordination. However, Goodale and Servos showed that the availability of binocular cues before or during a movement increases its accuracy and efficiency [16].

Summarizing, results from neuroscience indicate that for human hand-eye coordination both open-loop and dynamic closed-loop visual control strategies are employed, the latter being important if a high endpoint accuracy is required. This also applies to the criteria focus of attention and camera configuration: The more complex alternatives endpoint-closed-loop and binocular vision are not necessary but increase the efficiency and the precision of movements. The question whether visual imformation is transformed into Cartesian coordinates is not answered convincingly yet.

4. SYSTEM CONCEPT

The former sections showed that human strategies for hand-eye coordination can be described in robotical terms. Unfortunately, two facts prevent us from directly copying the human system: First, some questions remain open, such as the form of visual information used for motion control; in these cases, we a free to choose a suitable robotical strategy. Secondly, due to the difference in "hardware", parts of the human hand-eye system may not be realizable on a robot. This leads to decisions based on the current state of technology.

Hand-eye configuration

In order to mitigate the problem of different hardware we plan to implement and test algorithms on the robot system depicted in Fig. 1, which consists of a 6DOF arm (*amtec*) and a pan-tilt head with two colour cameras (*RWI stereo vision system*).

Visual control strategy

A look at the human hand-eye system shows that the classical look-then-move strategy is very useful for moving the hand towards the object to grasp. As standard robot manipulators only provide a two-finger gripper, the question of endpoint accuracy cannot be neglected, thus visual feedback is necessary at least in the final phase of the reaching movements to correct modelling and measurement errors. Yet, numerous results of neuroscience show that visual information is also incorporated into the highly automated and preplanned reaching movement. Thus, the clear division into a reaching and a grasping phase as proposed in [20] cannot be maintained. Instead, a trajectory generation scheme encompassing both phases has to be designed. An additional advantage of this hybrid structure is that the look-then-move trajectory can be planned in such a way that the target remains visible throughout the movement.

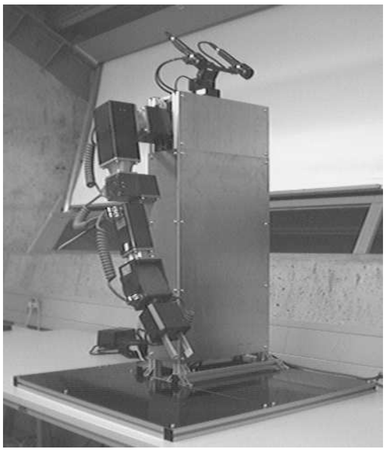

Figure 1: Hand-eye system **MinERVA**

Control architecture

Due to the relatively slow visual feedback rates, in both the human hand-eye system and state-of-the-art robotical ones direct visual control is not feasible (yet).

Visual information

Concerning this criterion, the literature in neuroscience is ambiguous. Therefore, we are free to select robotical strategies that can be realized on our robot system. Because the target position for the look-then-move trajectory does not need to be very accurate, it can be computed using a standard pose estimation algorithm combined with a coarse Cartesian hand-eye calibration. For the fine positioning in the vicinity of the target, pose estimation without the help of image processing hardware is too slow; thus, image-based methods such as the one proposed by Hager [18] are more suitable.

Focus of attention

To plan and control a movement that brings the manipulator into the vicinity of the target, the current position of the end-effector can be computed from the joint angles. Errors in the kinematic model of the robot can be compensated by the visual servoing process. For the latter, vision of the hand is necessary to allow image-based control.

Camera configuration

In the human hand-eye system, binocular vision is necessary for achieving high endpoint accuracy. Fortunately, Hager [18] and Hollinghurst [22] demonstrated that binocular vision can be successfully employed for image-based visual servoing with free-standing cameras, without the need for image processing hardware. For the initial estimation of the target position, monocular vision is sufficient if an object model is available.

Summarizing, we propose to use a dynamic position-based endpoint-open-loop look-then-move control module to generate a trajectory that brings the manipulator into the vicinity of the object to grasp, with a superimposed dynamic image-based endpoint-closed-loop visual servoing control to compensate modelling and image processing errors in the final phase of the movement. Cartesian coordinates of the target position are to be provided by a monocular model-based pose estimation module, while the visual servoing is to use a binocular camera system.

5. CONCLUSION

The goal of the work described in this paper was to develop a concept for an anthropomorphic robotical hand-eye system. To achieve this, we first gave a short survey of the state-of-the-art of robotical hand-eye coordinations and formulated criteria by which robotical hand-eye systems can be classified. We then analyzed human strategies for hand-eye coordination and showed that they can easily be described in robotical terms. Unfortunately, two facts prevent us from directly copying the human system: First, some questions such as the form of visual information used for motion control, remain open; here, further research is necessary. Secondly, due to the difference in “hardware”, strategies found in the human hand-eye system, may not be realizable on a robot. Thus, our concept for an anthropomorphic hand-eye system is to be seen as a specification that must be continuously refined and adapted in the progress of research.

Further work will concentrate on the implementation and testing of our system concept on our robot **MinERVA** which provides an anthropomorphic hand-eye configuration. For the design of the hybrid motion planning and control module, we will examine the current models of human reaching and compare them to visual servoing methods, with the goal of developing a common control module for both strategies.

REFERENCES

[1] P. K. Allen, A. Timcenko, B. Yoshimi, and P. Michelman. Hand-Eye Coordination for Robotic Tracking and Grasping. In K. Hashimoto, editor, *Visual Servoing*, pages 33–69. World Scientific Publishing Company, 1993.

[2] T. Bandlow, A. Hauck, T. Einsele, and G. Färber. Recognising objects by their silhouette. In *IMACS Conf. on Comp. Eng. in Systems Appl. (CESA'98)*, Apr. 1998.

[3] S. Blessing, S. Lanser, and C. Zierl. Vision-based Handling with a Mobile Robot. In M. Jamshidi, F. Pin, and P. Dauchez, editors, *International Symposium on Robotics and Manufacturing (ISRAM)*, volume 6, pages 49–59. ASME Press, 1996.

[4] E. Burdet. *Algorithms of Human Motor Control and their Implementation in Robotics*. PhD thesis, ETH Zürich, 1996.

[5] A. Castano and S. Hutchinson. Visual Compliance: Task-Directed Visual Servo Control. *IEEE Trans. on Robotics and Automation*, 10(3):334–342, June 1994.

[6] F. Chaumette, E. Malis, and S. Boudet. 2D 1/2 visual servoing with respect to a planar object. In *Proc. Workshop on New Trends in Image-Based Robot Servoing, In assoc. with IROS'97*, pages 45–52, Sept. 1997.

[7] F. Chaumette, P. Rives, and B. Espiau. Positioning of a Robot with respect to an Object, Tracking it and Estimating

its Velocity by Visual Servoing. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'91)*, pages 2248–2253. IEEEE, 1991.

[8] P. I. Corke. Visual Control of Robot Manipulators – A Review. In K. Hashimoto, editor, *Visual Servoing*, pages 1–31. World Scientific Publishing Company, 1993.

[9] A. Crétual and F. Chaumette. Positioning a Camera Parallel to a Plane Using Dynamic Visual Servoing. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'97)*, pages 43–48, Sept. 1997.

[10] N. Daucher, M. Dhome, J. Lapresté, and G. Rives. Speed Command of a Robotic System by Monocular Pose Estimate. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'97)*, pages 55–62, Sept. 1997.

[11] B. Espiau, F. Chaumette, and P. Rives. A New Approach to Visual Servoing in Robotics. *IEEE Trans. on Robotics and Automation*, 8(3):313–326, June 1992.

[12] J. T. Feddema. *Real Time Visual Feedback Control for Hand-Eye Coordinated Robotic Systems*. PhD thesis, Purdue University, 1989.

[13] C. Ghez, J. Gordon, M. F. Ghilardi, and R. Sainburg. Contributions of Vision and Proprioception to Accuracy in Limb Movements. In M. S. Gazzaniga, editor, *The Cognitive Neuroscience*, pages 549–564. MIT Press, 1995.

[14] M. A. Goodale, A. D. Milner, L. S. Jakobson, and D. P. Carey. A neurological dissociation between perceiving objects and grasping them. *Nature*, 349:154–156, 1991.

[15] M. A. Goodale, D. Pelisson, and C. Prablanc. Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature*, 320:748–750, 1986.

[16] M. A. Goodale and P. Servos. Visual Control of Prehension. In H. N. Zelaznik, editor, *Advances in Motor Learning and Control*, chapter 5. Human Kinetics, 1996.

[17] G. Hager and Z. Dodds. A Projective Framework for Constructing Accurate Hand-Eye Systems. In *Proc. Workshop on New Trends in Image-Based Robot Servoing, In assoc. with IROS'97*, pages 71–82, Sept. 1997.

[18] G. D. Hager. A Modular System for Positioning Using Feedback from Stereo Vision. *IEEE Trans. on Robotics and Automation*, 13(4):582–595, 1997.

[19] K. Hashimoto, T. Kimoto, T. Ebine, and H. Kimura. Manipulator Control with Image-Based Visual Servo. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'91)*, pages 2267–2272, 1991.

[20] A. Hauck and G. Färber. Hybrid Hand-Eye Coordination with a Single Stationary Camera. In *4th Int. Conf. on Control, Automation, Robotics and Computer Vision (ICARCV'96)*, pages 1715–1719, 1996.

[21] A. Hauck and N. O. Stöffler. A Hierarchical World Model with Sensor- and Task-Specific Features. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'96)*, pages 1614–1621, 1996.

[22] N. J. Hollinghurst. *Uncalibrated Stereo and Hand-Eye Coordination*. PhD thesis, Department of Engineering, University of Cambridge, Jan. 1997.

[23] R. Horaud and F. Chaumette, editors. *Workshop on New Trends in Image-Based Visual Servoing*. In association with IROS'97, Sept. 1997.

[24] S. Hutchinson, G. D. Hager, and P. I. Corke. A Tutorial on Visual Servo Control. *IEEE Trans. on Robotics and Automation*, 12(5):651–670, Oct. 1996.

[25] M. Jägersand. Image Based Visual Simulation and Tele-Assisted Control. In *Proc. Workshop on New Trends in Image-Based Robot Servoing, In assoc. with IROS'97*, pages 33–44, Sept. 1997.

[26] M. Jeannerod. The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17:187–245, 1994.

[27] N. Mai and C. Marquardt. Analyse und Therapie motorischer Schreibstörungen. In *Psychologische Beiträge*, volume 37, pages 538–582, 1995.

[28] N. Mai, S. Przywara, J. Hermsdörfer, and C. Marquardt. Behandlung der Posttraumatischen Ataxie. In Kuratorium ZNS, editor, *Tagungsbericht Forschung und Praxis der Neurologischen Rehabilitation – 10 Jahre Kuratorium ZNS*, pages 241–256, 1995.

[29] T. E. Milner. A model for the generation of submovements requiring endpoint precision. *Neuroscience*, 49(2):487–496, 1992.

[30] M. Mishkin, L. G. Ungerleider, and K. A. Macko. Object vision and spatial vision. Two cortical pathways? *Trends in Neuroscience*, 6:414–417, 1983.

[31] B. J. Nelson and P. K. Khosla. Visually Guided Manipulation Using Active Camera-Lens Systems. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'97)*, pages 437–442, 1997.

[32] P. Rives and J.-J. Borrelly. Underwater Pipe Inspection Task using Visual Servoing Techniques. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'97)*, pages 63–68, Sept. 1997.

[33] A. C. Sanderson and L. E. Weiss. Image-based visual servo control using relational graph error signals. In *Proc. of the IEEE*, pages 1074–1077, 1980.

[34] J. F. Stein. The representation of egocentric space in the posterior parietal cortex. *Behavioral and Brain Sciences*, 15:691–700, 1992.

[35] M. Tonko, J. Schurmann, K. Schäfer, and H.-H. Nagel. Visually Servoed Gripping of a Used Car Battery. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'97)*, pages 49–54, Sept. 1997.

[36] L. Weiss. *Dynamic Visual Servo Control of Robots: an Adaptive Image-Based Approach*. PhD thesis, Carnegie-Mellon University, 1984.

[37] R. Woodworth. The accuracy of voluntary movement. *Psychological Review*, 3:1–114, 1899.

[38] P. Wunsch and G. Hirzinger. Real-Time Visual Tracking of 3D Objects with Dynamic Handling of Occlusion. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'97)*, pages 2868–2873, Apr. 1997.

[39] J. Y. Zheng, Q. Chen, and S. Tsuji. Active camera guided manipulation. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'91)*, pages 632–638, 1991.