Calibration-Free Robots

Volker Graefe*
Institute of Measurement Science
Bundeswehr University Munich
85577 Neubiberg, Germany

Abstract

Robots not needing any quantitative models of themselves or of the world and, thus, no calibration of their sensors, their actuators or their kinematics promise great advantages over conventional robots in terms of robustness, adaptability and cost of ownership. Concepts for such robots are introduced and results obtained in real-world experiments are reported. Similarly to infants and young animals, such robots learn relationships between internal motion control commands on one hand and resulting effects on data produced by their sensors through interactions with the world on the other hand. On the basis of such knowledge they reach goals in the world by controlling their motions in such a way that equivalent goals in their internal sensor data space are reached. Experiments show that object manipulation and an automatic build-up of skill by such robots has been accomplished.

Keywords: calibration-free robots, calibration-free vision, learning

1 Introduction

Conventional robots depend heavily on quantitatively correct models of their own characteristics and of the relevant parts of their environments. Generating those models and maintaining their accuracy in an ever-changing and often unpredictable world is difficult and expensive.

Robots depending neither on inbuilt quantitative models nor on pre-defined numerical values of any parameter would never need any explicit calibration and, therefore, promise great advantages in terms of robustness and cost of ownership. Such robots are called “calibration-free.” Some types of quantitative model knowledge are trivial to provide, such as the number of motors to be controlled or the number of sensors used by the robot. If a robot is provided with this type of knowledge, but not with any other quantitative knowledge, it is still considered calibration-free.

In some cases it is difficult to distinguish between truly calibration-free robots and self-calibrating ones. While truly calibration-free robots do not use any quantitative model knowledge, self-calibrating ones use some quantitative model knowledge, but they acquire the necessary knowledge automatically, either during a distinct calibration phase or as a side effect of their normal operation. In the sequel self-calibrating robots will be included in the term “calibration-free.” Truly calibration-free robots are robust, not only against modeling errors, but also against later changes of parameter values, because they do not use such data. Self-calibrating robots can be similarly robust, if they do not limit themselves to an initial self-calibration, but continue to determine the actual values of the relevant parameters during their normal operation.

Self-calibration is a kind of learning. To adapt to changes, a self-calibrating robot should never cease to learn. Also, it should be able to forget or modify previously acquired knowledge that has become irrelevant due to changed circumstances.

The completely calibration-free robot may be an idealization, or it may in some cases be so difficult to realize that a realization is not practicable. However, robots that are not calibration-free, but require only a coarse, approximate, and easy-to-perform calibration provide similar advantages as calibration-free robots in terms of robustness and cost of ownership. Therefore, they are of practical interest, too, and their design may be based on largely the same concepts as those used for calibration-free robots. An example of a coarsely calibrated robot would be a vision-guided vehicle with a forward-looking camera where the actual orientation of the camera is allowed to deviate by, say, 20 degrees in any direction from its assumed ideal orientation.

* E-Mail: Graefe@unibw-muenchen.de URL: http://www.unibw-muenchen.de/campus/LRT6/

Concepts for the realization of vision-guided calibration-free robots will be presented in the sequel and an introduction into some related problems will be given. Among the problems that will be discussed are:

- Manipulation: How can a robot manipulate objects without inbuilt knowledge of its optical and kinematic parameters?
- Knowledge representation and learning: How can a robot learn by doing? What should be learned? How should it be represented? How can the learning speed be optimized? How should forgetting and relearning be organized in the interest of an adaptation to changing situations?

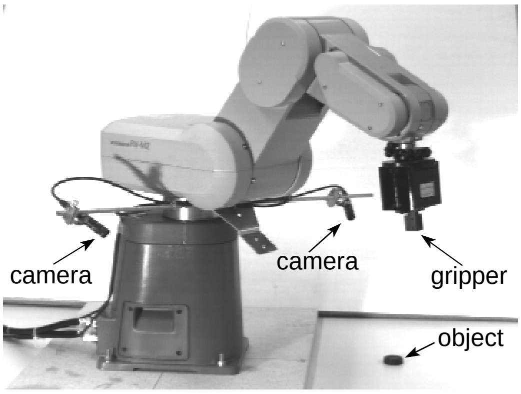

Figure 1: Manipulator robot consisting of an articulated arm (5 degrees of freedom) with a two-finger gripper and a stereo vision system.

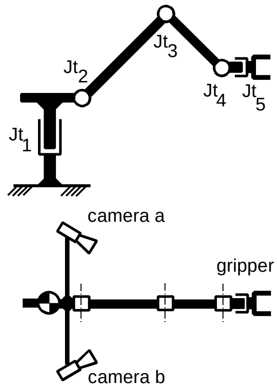

Figure 2: Schematic diagram of the articulated arm and the camera arrangement.

Although not all of these questions can be answered yet, some solutions will be presented and results of real-world experiments with calibration-free robots will be reported.

2 Robot Control

2.1 An Example: Object Manipulation

A manipulator arm equipped with a pair of video cameras and performing object manipulation may serve as an example of a calibration-free robot (Figures 1 and 2). The manipulation task studied is simple, but characteristic: An object that is placed at some initially unknown location in the robot's work space is to be grasped and picked up by the robot's 2-finger gripper. To simplify the explanation of the calibration-free approach, it will be assumed initially that the objects to be manipulated are of rotational symmetry (e.g., cylinders or spheres) with a vertical axis. Only 3 of the robot's 5 degrees of freedom (Jt_1, Jt_2, and Jt_3) need then to be considered; Jt_4 may be controlled in such a way as to always keep the gripper in a vertical orientation, and the rotation of the gripper around its vertical axis (Jt_5) is irrelevant due to the rotational symmetry of the object.

2.2 The Traditional Model-Based Approach

Concept

A classical approach for controlling such a robot would evaluate the camera images according to the well-known methods for stereo evaluation (a kind of inverse perspective transformation). This transformation requires all internal and external camera parameters (location, orientation, focal length, principal point, etc.) to be known with great accuracy. The stereo evaluation would then deliver the location of the object to be grasped relative to some coordinate system that is defined by the orientations and locations of the cameras. As an alternative, a carefully aligned projector illuminating the scene with a specific light pattern could be used in combination with one or two cameras. In both cases the coordinates of the object in a ground-based coordinate system would then be determined by an appropriate coordinate transformation critically depending on a substantial number of parameters.

Finally, the joints of the robot would be controlled in such a way as to move the gripper to that point in the ground-based coordinate system which was determined to coincide with the location of the object. This control implies a coordinate transformation (inverse kinematic transformation) between the ground-based coordinate system and the joint angles, which requires accurate knowledge of the robot's dimensions, kinematics, joint angles and control characteristics.

Advantage

The main advantage of model-based control is that it lends itself to the application of classical control theory and, thus, may be considered a conventional and well-understood approach.

Shortcomings

The control of vision-guided manipulators according to the traditional model-based approach depends critically on an accurate calibration of the camera parameters and the arm kinematics. Model-based control breaks down when there is no accurate quantitative agreement between reality and the models. Differences between reality and models may come about easily; an error in one of the many coefficients that are part of the numerical models may be sufficient for causing a failure of the robot. Among the many possible causes for discrepancies are initial calibration errors, aging of components, changes of environmental conditions, such as temperature, humidity, electromagnetic fields or illumination, maintenance work and replacement of components, to mention only a few. Consequently, robots depending on model-based control usually work only in carefully controlled environments and need frequent recalibrations, in addition to a cumbersome and expensive initial calibration.

2.3 Calibration-Free Robot Control

Concept

To overcome the shortcomings of model-based robot control many researchers have striven to develop robot control concepts that avoid, or at least reduce, the necessity of modeling and calibration. [Cooperstock, Milios 1993] use a set of neural networks to realize a robot able, after a training phase, to approach an object and grasp it without requiring a calibration. [Yoshimi, Allen 1994] use an uncalibrated camera system to perform a peg-in-hole task utilizing rotational invariance in a specific setup. [Hollinghurst, Cipolla 1994] avoid using an exact camera model by performing a self-calibration at four known positions. Doing this, they can handle calibration errors of the cameras as well as linear modeling errors of the manipulator.

In the sequel a method introduced by [Graefe 1995] will be described. It was developed independently of a similar work by [Conkie, Chongstitvatana 1990]; it differs from that work by using the motor control word space instead of the joint space, thus avoiding the need of calibrating the joints. In contrast to [Cooperstock, Milios 1993] it does not require any initial training of the robot.

A key point of the method introduced by [Graefe 1995] is the way how the robot and the world are modeled (Figure 3). The robot controller models everything outside itself as one "black box" of unknown internal structure. The black box includes the entire external world, but also the sensors, actuators and other components of the robot. The robot controller can send control commands to that black box and receives sensor data from it. We humans know that control commands cause the robot's actuators to move, and the state of the robot and of objects in the world to change, and that the sensor data have some meaning and reflect the state of objects in the world. The robot controller does not know any such things. All it knows is that some sensor data will change (actually, due to independent motions of external objects, but from the robot's point of view seemingly at random) even if the controller does not send any control commands, other sensor data will change in a more or less systematic, and thus predictable, way as a consequence of the control commands the controller issued. Its world model consists of a set of associations

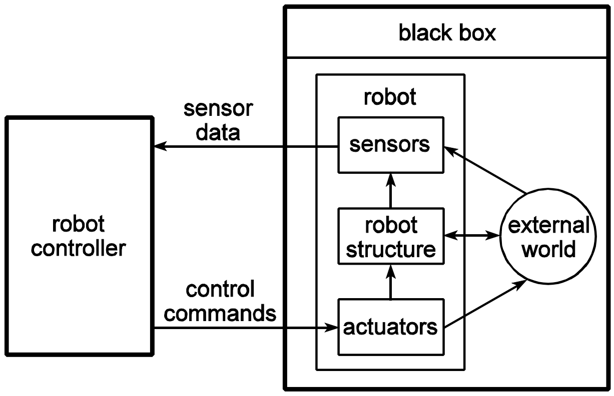

Figure 3
The controller of a calibration-free robot and its model of the world and of the robot; the internal structure of the "black box" is unknown to the robot controller.

between control commands on one hand, and resulting sensor data, or changes in sensor data, on the other hand. It may acquire such knowledge either by being programmed accordingly, or by learning.

When we humans assign a task to a robot we tend to define a goal in terms of a desired state of the world, e.g., a box being placed on a pallet, or an object being grasped by the robot. For the controller of a calibration-free robot such task descriptions or goals are meaningless. It understands as goals to be achieved certain combinations of sensor data, and in pursuing such goals it will issue control commands intended to bring the desired sensor data about. Taking a grasping operation as an example, some sensor data may, in human terms, reflect the locations and orientations of the object and the gripper. The robot could then accomplish the grasping by first making the sensor data relating to the locations and orientations of the open gripper and the object identical, then sending a control command that causes the gripper to close, and finally evaluate sensor data that reflect the degree of success of the grasping operation, if such data are available.

The robot control method introduced by [Graefe 1995] is based on such concepts. It eliminates the need for a calibration of the robot and of the vision system, it uses no world coordinates, and it comprises an automatic adaptation to changing parameters. It is based on the utilization of laws of projective geometry that always apply, regardless of camera characteristics, on machine learning for the acquisition of knowledge concerning the relationships between control commands and subsequent changes in sensor data, and on a direct transition from visual input information to control commands (without any models of the robot and the external world, as indicated in Figure 3). Different forms of learning and knowledge representation have been studied in connection with this method, allowing either the rapid adaptation to changes of the system parameters or the gradual improvement of skills by an accumulation of knowledge.

The learning concept that we have implemented for calibration-free grasping has been inspired by nature. While the robot watches its end effector with its cameras, like a playing infant watches his hands with his eyes, it executes "test motions" by sending more or less arbitrary control commands to its motors. By observing the resulting changes in the camera images it "learns" the relationships between such changes in the images and the control commands that caused them. After having executed test motions in a variety of configurations the robot is able to move its end effector to any position in the camera images that corresponds to a physically reachable position in the world. If, in addition to the end effector, an object is visible in the images, the end effector can be brought to the object in both camera images and, thus, in the real world.

Based on this concept a robot can grasp objects without any knowledge of its kinematics or its camera parameters. In contrast to other approaches with similar goals, but based on neural nets, no training is needed before the manipulation may start.

It should be possible to have the robot learn the relationships between motor control commands and resulting image motions by having it observe its own motions during its ordinary operation, instead of having it execute specific test motions. This would lead to a faster operation, but it has not been tested yet.

Advantages

In sharp contrast to the classical approach, ours requires no knowledge regarding:

- the exact locations of the cameras
  (except that the cameras should be located at some distance from the work plane of the robot, preferably in an opposite arrangement)

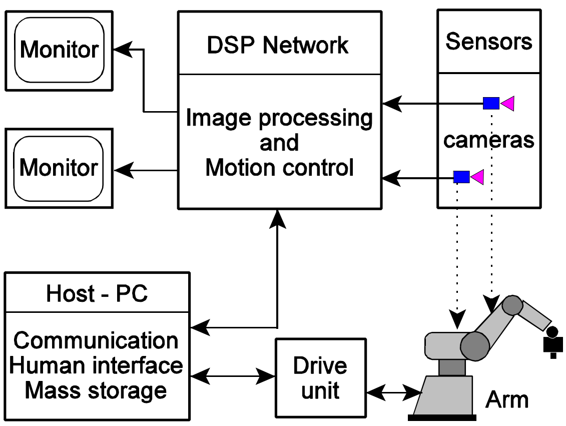

Figure 4
The experimental setup, including the arm, the host computer and the DSP network for stereo vision and motion control

- the exact viewing directions and the internal parameters of the cameras
  (except that both cameras should have the actual work space of the robot in their fields of view)
- the dimensions, kinematics and joint angles of the robot
  (except that, for practical reasons, we presently assume that the robot is of the articulated arm type, and that the general type of the gripper and the number of degrees of freedom of the system are known)
- the quantitative relationships between the control words sent to the motor controllers and the resulting motions
  (except that these relationships are assumed to be "smooth").

Shortcomings

The concept as such has no known shortcomings. However, in the present implementation it is still necessary to execute test motions in addition to the actually desired motions of the robot, which slows down the operation. Also, some problems related to the initialization of the robot when it is switched on, and to the optimization of learning and knowledge representation have not been fully solved, yet.

3 Realization

3.1 Experimental Setup

We have implemented the concept on a robot consisting of an articulated arm (Mitsubishi Movemaster RV-M2) with 5 degrees of freedom, corresponding to the 5 joints, Jt_1 to Jt_5 (Figures 1 and 2), and a stereo vision system as the only sensor. The cameras have been attached to the robot on metal rods at the first link, so that they rotate around the (vertical) axis of Jt_1 together with the arm. They are mounted in a rather unstable way to make the impossibility of any calibration or precise adjustment obvious, and to allow easy random modifications of the camera arrangement. Figure 4 shows the setup used for our experiments.

In the context of this particular work, we are only interested in the problem of robot control without model knowledge, not in grasping strategies, collision avoidance or complex vision problems. Therefore, we have structured the environment for our experiments in such a way that both the gripper and the object are easy to locate in the images, and that the object is easy to grasp by a gripper approaching it from above without any risk of collisions with the table or other objects. For the purpose of describing the concept of our approach we model both the gripper and the object as visible points. The robot's vision system recognizes and tracks the gripper and the object in the images, assigns reference points to them, and continuously delivers the image coordinates of both points in the images of both cameras (Figure 5). The grasping may be considered to have been accomplished when the two points, gripper and object, coincide in the real world. Regardless of any camera characteristics, this is the case if, and only if, they coincide in the images of both cameras. The rendezvous between the gripper and the object may, therefore, be brought about by any sequence of robot motions that makes the two points coincide in the images(!) of both cameras simultaneously.

3.2 Implementation of the Control Concept

In implementing the concept we made the following assumptions:

- The robot's gripper and the object to be grasped are visible for both cameras; the vision system recognizes them and delivers continuously their image coordinates in both camera images.
- The internal and external camera parameters (optical characteristics, locations, and viewing directions) are unknown.
- The motions of the cameras are unknown. The cameras may either be at rest (located somewhere outside the robot), or they may move together with some part of the robot (not necessarily the end effector); they do not move independently, and they do not move relative to each other.
- The dimensions and joint angles of the robot are unknown; there is no kinematic redundancy.
- The robot's motions are controlled by numerical control words, one for each motor. The functional dependencies of world coordinates of the gripper on control words are unknown. Also, the functional dependencies of image coordinates on world coordinates are unknown, since we do not use quantitative models. However, we assume that a zero control word causes no motion, while a control word of small magnitude causes a small motion, and a control word of larger magnitude causes a larger motion.

Two implementations were realized: the first one utilizing qualitative knowledge of the robot's kinematics treated the rotation of the arm around its vertical axis separately from the motions of the arm joints, Jt_2 and Jt_3, while the second one did not use such knowledge and treated all degrees of freedom in a unified way.

The First Implementation

To be picked up, an object must be located in that plane in which the gripper moves when the joints, Jt_2, Jt_3, or Jt_4 move (the work plane of the robot). The task of picking up an object may, therefore, be decomposed into two subtasks: a rotation of the robot around its vertical axis until the work plane coincides with the object to be picked up, and a motion of the gripper within the work plane.

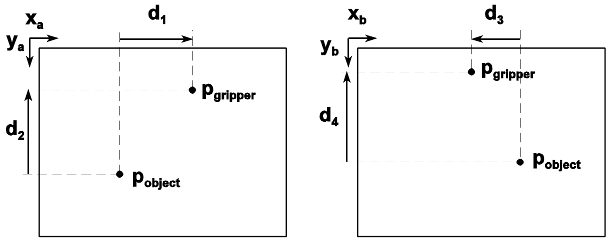

Figure 5
The gripper and the object (modeled as points) in the images of two cameras, a (left) and b (right)

The control for the motion within the work plane will be described first. Let us assume for the moment that the work plane already coincides with the object. The task is then to modify the angles of the joints, Jt_2 and Jt_3, (Figure 1), in such a way that the gripper coincides with the object. The image from either one of the cameras may be used for accomplishing this goal. Let us assume, as an example, that the camera, a, is being used. It performs a one-to-one mapping between an area of the robot's work plane and the camera's image plane. The images of the object and of the gripper coincide if, and only if, the object and the gripper are at the same location. To grasp the object it is, therefore, sufficient to make the image coordinates of the gripper identical to the image coordinates of the object. (It is a key point of the calibration-free approach that we are concerned only with the robot-internal image data, but not with any quantitative relationships between those image data and the states of external objects.)

In a calibration-free robot the gain coefficients relating the magnitudes of motor control words to the resulting changes in the image coordinates of the gripper are initially unknown. They may be determined by sending control words to both joints to make them execute small test motions and observing the resulting changes in the images. The gain coefficients relating the 2 control words to the resulting changes in both image coordinates in both camera images form a 2*2 Jacobian matrix. Since this Jacobian relates sensor data directly to motor control words it is called sensor-control Jacobian [Graefe, Maryniak 1998]. Having determined the elements of the sensor-control Jacobian, it is easy to compute those control words which would make the gripper coincide with the object, if the system were linear. In reality, due to the non-linearity of the system and other reasons, executing a motion according to the computed control words will not bring the gripper exactly to the object, but at least the distance between the gripper and the object will be reduced. Repeating the procedure once or twice makes the distance small enough for grasping the object.

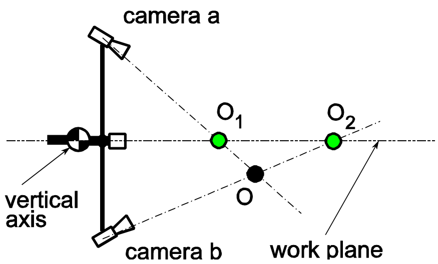

Figure 6
Disparity of apparent object locations, O_1 and O_2, corresponding to an object, O, outside of the robot's work plane

Up to this point it has been assumed that the object was initially located in the work plane. Therefore, only one of the two cameras was necessary, and it did not matter which one of them was actually used. Their images, and the sensor-control Jacobians, are different, but the control words computed in step 1 of the rendezvous procedure above are the same, regardless of which camera's image is used. The reason is that there is only one correct motion of the gripper in the world and, thus, only one correct set of control words.

Now let us assume that the object is not located in the work plane. In this case, the angle of joint Jt_1, must be modified. Whether such a situation exists, may be determined by computing the control words according to step 1 of the rendezvous procedure once each with the image from each camera, without executing a motion. In the rendezvous procedure it is assumed implicitly that any object seen is located in the robot's work plane. If actually the object is not located in the work plane, the two cameras will see the object at different locations in the work plane, as Figure 6 shows, and consequently the two sets of control words will be different. Controlling the joints, Jt_2 and Jt_3, based on the image of the camera, a, would, therefore, generate a motion of the gripper toward O_1, the projection of O onto the work plane as seen by the camera, a. Similarly, basing the motion control on the image of the other camera, b, will cause a motion of the gripper toward O_2. The two sets of motion control words will be identical if, and only if, O is located in the work plane, because only then will O_1, O_2, and O be at the same location.

If the two sets of control words computed on the basis of the two images are different, the angle of the first joint, Jt_1, should be modified by issuing a suitable control word. The correct sign and magnitude of the control word which makes the work plane coincide with the object may be determined from the disparity between the two sets of control words for the joints, Jt_2 and Jt_3, computed from the images of the two cameras. An iterative method similar to the one used for controlling Jt_2 and Jt_3 may be used for this.

The Second Implementation

The second implementation differs from the first one by treating all degrees of freedom in the same way, without using any knowledge of the robot's kinematics. We assume that the motion of the robot is controlled by a set, or vector, $\underline{c}$, of numerical control words (one control word for each motor) that is sent to the actuators.

If $\underline{d} = (d_1, d_2, d_3, d_4)^T$ (Figure 5) is the vector of image distances between the gripper and the object in all dimensions used (x and y in both images), then the task of making the gripper rendezvous with the object is equivalent to the task of making $\underline{d}$ vanish.

If J is the sensor-control Jacobian matrix containing the partial derivatives of individual image distances, d_i, on control words, c_k, a control word vector, $\underline{c}°$, satisfying

$$-\underline{d} = J \cdot \underline{c}° \qquad \textbf{(1)}$$

would make $\underline{d}$ equal to $\underline{0}$ if the system were linear. Because of the non-linearity of the (unknown) function between control commands and image distance, outputting $\underline{c}°$ will generally not make $\underline{d}$ become zero. However, it will decrease its magnitude, and iterating the process will finally lead to a sufficiently small image distance vector, $\underline{d}$, such that the grasping task is being accomplished (for more details see [Graefe, Maryniak 1998]).

One problem with this method is that there are 4 sensor data and, thus, 4 equations for computing the 3 control words forming the control word vector, $\underline{c}°$. Consequently, the system of equations, (1), is over-determined. This problem may be solved in various ways, e.g., by omitting one of the equations [Graefe, Maryniak 1998], by a transformation from the original 4-D sensor data space into a 3-D data space [Maryniak, Graefe 1998], or by utilizing the redundancy inherent in the over-determined system for improving the robustness of the robot control [Maryniak, Graefe 1999].

3.3 Experimental Results

Both methods were successful in enabling the robot to grasp objects, although no knowledge regarding the optical or kinematic parameters of the robot, or of the locations of the objects, was used. Even arbitrary rotations of the cameras were tolerated while the robot was operating. Key to this unusual robustness is the renunciation of quantitative model knowledge (i.e., camera model and kinematic model of the arm) and the direct transition from image coordinates to motor control commands, without using any inverse perspective or kinematic transformations.

The gripper approached the object in a series of steps with some test motions in-between ("look-think-move" approach). The grasping process was, therefore, relatively slow (about 50 s).

3.4 Discussion

The viability of the concept and of both implementations was demonstrated. In one respect, however, the robot is not completely calibration-free. To avoid the risk of the gripper colliding with the object during the approach, the gripper actually approaches an intermediate position, a few cm above the object, before finally grasping it. In our implementation this has been realized by implicitly using the knowledge that the y-coordinates in the images (Figure 5) correspond approximately to "down" in the world, and that an offset of a certain number of pixels in the vertical image coordinates of the object corresponds to a suitable distance in the world between the intermediate position and the object. Consequently, a rotation of a camera by more than about 20 ° around its optical axis or a change of the focal length by more than about 50% might not be tolerable. In these respects the robot may be considered coarsely calibrated. If the intermediate position had been defined as, say, "one gripper diameter away from the object in the direction of the gripper's axis" the robot could be considered calibration-free in this respect. However, this has not been implemented yet.

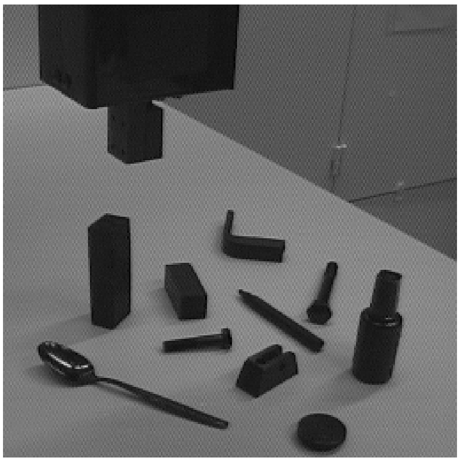

Figure 7
Objects that could be manipulated

In both implementations described the gripper is controlled in such a way that its axis always remains approximately vertical. One reason for this is that approaching the object from above and with a vertically oriented gripper minimizes the risk of collisions, for instance, with the table. Although the accuracy of the gripper's orientation is not critical, the attempt to maintain a specific orientation of the gripper relative to the gravity vector does imply a kind of coarse calibration. Avoiding it seems to be difficult, though. Probably it would require the robot to have some additional general knowledge of the world, of preferred directions in it, and of ways for recognizing them in camera images.

3.5 Additional Degrees of Freedom

While in the initial implementations of the concept only objects of rotational symmetry could be handled, the first implementation was augmented by [Nguyen 1997] to allow a greater variety of objects to be manipulated (Figure 7 shows examples). To grasp such objects, the gripper must not only be in the right position, but also in the right orientation. If the objects are either lying flat on a table or standing upright, the gripper's axis may be vertical, and a rotation of the gripper around Jt_5 is sufficient for obtaining a suitable orientation of the gripper.

The correct orientation of the gripper may be reached by first bringing the gripper to a location where it is near the object in the camera images. Then it is rotated until the edges of the gripper are parallel to the main axis of the object in the image. Since the gripper is in a vertical orientation, and since the main axis of the object to be grasped is either vertical or horizontal, this will occur simultaneously in both images. Even if the edge of the gripper is exactly parallel to the object axis in the images, due to perspective distortion this does not mean that they are exactly parallel in the world, too. However, if the object is near the gripper, the angle difference is small enough to allow the object to be grasped.

Objects of arbitrary shape in arbitrary orientations generally cannot be grasped by a robot with only 5 degrees of freedom. We are currently investigating certain special cases where objects as shown in Figure 7 may be grasped in an almost arbitrary orientation by a calibration-free robot with 5 degrees of freedom.

4 Learning

4.1 Concept

For a calibration-free robot according to the concept shown in Figure 3, learning means that the robot controller learns relationships between control commands and subsequent sensor data changes, for instance in the form of the location-dependent sensor-control Jacobian. In the implementations described above, the relationships between control commands and the resulting motions of features in the camera images that were learned before each motion step by executing test motions were forgotten immediately after the motion step. This made the system completely insensitive to parameter changes between motion steps, but it also prevented the robot from accumulating experiences and thereby improving its skills over time.

[Xie et al. 1997] have presented a method for knowledge representation that allows a robot to remember which control word vectors cause the gripper to move to specific locations in the camera images. Once the robot has moved the gripper to a specific location, it can then move it to the same location again later in a single step without performing test motions. If the gripper is to be moved not to exactly the same location as before, but to a point in a neighborhood of previously reached points, suitable control words for a one-step motion may be estimated by interpolation. If execution of the interpolated control word vector fails to move the gripper to the desired location with sufficient accuracy, the local sensor-control Jacobian is determined by test motions, and on this basis the goal is reached.

Whenever a grasping experiment is performed at a "new" location within the robot's work space, the control words belonging to that location are added to the knowledge base. As a consequence, the robot's knowledge becomes more comprehensive, and the operation of the robot gets faster as more grasping experiments are performed.

4.2 Implementation

The 4-D sensor data space (2 coordinates in each of 2 images, 512^4 points altogether) is divided (arbitrarily) into 8^4 cells (hypercubes) of 64^4 points each. The knowledge base is realized as a table allowing for 9^4 (= 6561) records, corresponding to the 9^4 corners of those cells. (It does not matter that some of the corner points do not correspond to physically reachable points in the world.) Each record belonging to a specific corner point in the 4-D image space may contain as data that control word vector which, when executed, makes the image of the gripper move to that point.

Initially the knowledge base contains no data, and the motion of the gripper is controlled by executing test motions according to the method described above. When the gripper has reached its destination, the control words corresponding to that location are assigned to all 16 corner points of the cell containing the goal point.

If later the gripper is to be moved to a point within a cell that has been visited before, the records for all 16 corners of the cell contain data. In this case a control word vector is computed by interpolation between the control word vectors belonging to the corner points of the cell. If the resulting motion brings the gripper to the desired location, we are finished. If not, a corrective motion is executed according to one of the methods described above, and corrections are applied to the data records belonging to the corner points of the cell. The magnitude of the corrections applied to the control word vector belonging to a corner point depends on the distance between the corner point and the gripper location (for details see [Xie et al. 1997]).

If the gripper is to be moved into a cell that has not been visited before, records for some of the corner points of that cell will contain no data. In this case the motion is performed according to one of the test-motion-based methods described above, and new or corrected data will be entered into the records belonging to the corner points of the cell.

4.3 Experimental Results and Discussion

[Xie et al. 1997] have shown, that it is, indeed, possible to realize permanent learning and an automatic skill improvement ("learning by doing") by this method. With each experiment the motions of the robot become faster and smoother, and after a few successful grasps the time required for grasping an object reduces to only a few seconds.

Machine learning is often associated with neural nets. However, compared to neural nets, tables as we use them have several advantages:

- It is obvious what has been learned.
- No training phase is necessary; the robot can start working and learning even with an empty knowledge base.
- Adding knowledge for additional locations to the knowledge base does not interfere with the existing knowledge (the so-called incremental learning is realized without any special effort).
- During the operation of the robot table entries may continuously be independently updated to average out measurement noise or to adapt to changes in the world (another form of incremental learning).

Scalability could be a problem with this approach. If we included the direction of the main axis of the object to be grasped in the images of both cameras in our sensor data space, we would need a 6-D table with about 9^6 (= 531,441) entries. It would probably take an unreasonable amount of time, before a substantial part of such a large table would be filled with data by learning and before a skill improvement could be observed. We are, therefore, experimenting with other forms of knowledge representation, for instance, a combination of individual 1-D or 2-D tables for the individual degrees of freedom of the robot.

Distributing the support points (the corners of the hypercubes) equally over the sensor data space is probably not the best possible approach. It may be better to let the system automatically assign support points in an irregular pattern where they are actually needed.

5 Conclusions and Outlook

Calibration-free vision-guided robots requiring no quantitative knowledge about their sensors and their kinematics can be realized. In contrast to conventional robots, they do not perform any operations that depend critically on substantial numbers of hard-to-determine parameters, such as inverse perspective or inverse kinematic transforms. Instead, they use direct transformations between sensor data (e.g., image coordinates) and motor control words. The parameters describing the transformations are initially learned, and then continuously updated, by the robot during its interactions with the world.

Although calibration-free robots do not make explicit references to the real world, they can, nevertheless, reach user-defined goal states in the real world. They do it by approaching equivalent goal states in their sensor data space. Grasping is an example: the user wants the robot to grasp an object. The robot performs this task without reference to world coordinates or other entities in the world. Instead, it moves its arm in such a way that the image coordinates of the gripper become identical to the image coordinates of the object, and then closes the gripper.

By renouncing the utilization of world coordinates and by not attempting to model quantitatively the robot, the world, and the interactions between the robot and the external world, calibration-free robots achieve a high degree of robustness and adaptability. These characteristics will be especially useful for future service and personal robots that, unlike today's industrial robots must cooperate with ordinary humans and work in unstructured environments without having access to frequent maintenance and professional support.

A number of questions are still open. Among them are:

- What is the best form of knowledge representation for calibration-free robots, especially in respect of an optimization of the learning speed?
- What is the best way for adapting the knowledge base to gradually changing conditions in the world? Would an individual Kalman filter for each stored control word be a good idea?
- Under certain circumstances, e.g., after replacing a camera, it may be best to discard the stored knowledge completely, rather than updating it gradually; under other circumstances, e.g., when the viewing direction of a camera has been changed, it may be better to modify all stored data at once in a systematic way; how can the robot recognize such circumstances?
- How can a robot, when it is switched on for the first time and has no quantitative knowledge of its own characteristics, begin to move safely and to learn something about itself without having an accident before it has actually learned to coordinate its motions?

Although the calibration-free approach has been introduced in the context of manipulator robots, it is likely that it is equally suitable for the navigation of mobile robots. In fact, some ideas leading into this direction have already been conceived, but they have not been implemented yet.

Significant advantages stand to be gained by a widespread utilization of calibration-free robots. They certainly justify some effort to develop this concept further and to solve the remaining open questions related to it.

References

Conkie, A., Chongstitvatana, P. (1990): An Uncalibrated Stereo Visual Servo System. DAI Research Paper No. 475, University of Edinburgh.

Cooperstock, J. R.; Milios, E. E. (1993): Self-Supervised Learning for Docking and Target Reaching. Robotics and Autonomous Systems. Vol. 11, 1993, pp 243-260.

Graefe, V. (1995): Object- and Behavior-oriented Stereo Vision for Robust and Adaptive Robot Control. International Symposium on Microsystems, Intelligent Materials, and Robots. Sendai, pp 560-563.

Graefe, V.; Maryniak, A. (1998): The Sensor-Control Jacobian as a Basis for Controlling Calibration-Free Robots. IEEE International Symposium on Industrial Electronics, ISIE '98. Pretoria, pp 420-425.

Hollinghurst, N; Cipolla, R. (1994): Uncalibrated Stereo Hand-Eye Coordination. Image and Vision Computing. Volume 12/3, 1994. pp 187-192.

Maryniak, A.; Graefe, V. (1998): Transforming Sensor Data to Increase Robustness in the Control of Calibration-Free Robots. IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS '98. Victoria, pp 1834-1839.

Maryniak, A.; Graefe, V. (1999): Utilizing Sensor Data Redundancy to Gain Robustness in the Control of Uncalibrated Robots. Proceedings, IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS '99. Kyongju, Oct. 99.

Nguyen, M.-C.; Graefe, V. (1997): Object Manipulation Controlled by Uncalibrated Stereo Vision. The Second Chinese Congress on Intelligent Control and Intelligent Automation, CWCICIA'97. Xian, Vol. 1, pp. 77-83.

Xie, Q.; Graefe, V.; Vollmann, K. (1997): Using a Knowledge Base in Manipulator Control by Calibration-Free Stereo Vision. IEEE International Conference on Intelligent Processing System. Beijing, pp 1307-1311.

Yoshimi, B. H.; Allen, P. K (1994): Active, Uncalibrated Visual Servoing. IEEE International Conference on Robotics and Automation. San Diego, Volume 4, pp 156-161.