Utilizing Sensor Data Redundancy to Gain Robustness in the Control of Calibration-Free Robots

André Maryniak and Volker Graefe
The Institute of Measurement Science, Bundeswehr University Munich
85577 Neubiberg, Germany

Abstract - In vision-based robot control typically more sensor data are available than the minimum necessary for computing robot control commands. We propose an approach for utilizing the redundant sensor data to improve the robustness of robot control. It is based on a novel method for taking into account all available measurements in solving the resulting over-determined system of equations for the robot control commands. The approach has been evaluated in computer simulations and in real-world manipulation experiments with a calibration-free robot. It has proven to be suitable for a real-time implementation and to lead to a more robust control than conventional methods.

Keywords: calibration-free manipulation, sensor data redundancy, robust robot control

1. INTRODUCTION & MOTIVATION

Robot vision systems have the potential to deliver rich information about the robot’s environment. This can be a great advantage, but it can also be difficult to utilize all this information and to handle it in real time. Here we present a method to utilize redundant sensor data for improving the robustness and adaptability of robot control. The method is of a general nature, but especially useful if no complete model of the robot and its sensor system exists, as in the case of a calibration-free robot. Accordingly, it will be introduced in such a context, where almost no model knowledge, and in any case absolutely no quantitative knowledge of the mechanical, kinematic, sensory and control characteristics of the robot, is to be used by the system or in its design.

In the field of vision-based control of robot manipulators much work has been done. For a good review the reader may refer to the Transactions on Robotics and Automation’s special issue on this topic (October 1996). Also some efforts in the more specific area of calibration-free manipulation are known, but to our knowledge the problem of utilizing sensor data redundancy in this context has not been addressed so far. While [Jägersand, Nelson 1995] tackle a different problem, they utilize the existing sensor data redundancy in the standard form of a least squares solution.

In [Graefe, Maryniak 1998] we have derived a concept to control a robot performing a grasping task based on the sensor-control Jacobian matrix, J, that contains the partial derivatives of image coordinates on control words. In that approach we modeled the gripper and the object as visible points; these points have to be made coincide in the real world in order to grasp the object. If, and only if, they coincide in the real world they also coincide in both images of the stereo camera arrangement we use. Therefore, the task can be described as making the two image points coincide in both images, regardless of any sensor or control characteristics. When \( d = (d_1, d_2, d_3, d_4)^T \) is the vector of distances between gripper and object in all dimensions used (x and y in both images, Figure 1), a control word vector, \( c^\circ \), satisfying

\[
\Delta d = J \Delta c
\]

would make the gripper coincide with the object if the system were linear and error-free.

Because of the non-linearity of the (unknown) function between control commands and image distances, outputting \( c^\circ \) will generally not make \( d \) become zero. However, it will decrease its magnitude, and feeding back the new \( d \) and, hence, iterating the process, will finally lead to a sufficiently small \( d \) such that the grasping task is being accomplished, as shown by [Graefe, Maryniak 1998].

In vision-based robot control, (1) is usually an over-determined system of equations. In our example it is a system of 4 equations for 3 unknowns. More generally, (1) is a system of \( m \) equations for \( n \) unknowns, where \( m \) is the dimension of the sensor data vector delivered by the robot’s sensor system, and \( n \) is the number of degrees of freedom of the robot; it is over-determined if \( m > n \).

In the example mentioned above [Graefe, Maryniak 1998] have solved a similar system of equations by arbitrarily discarding one of its 4 equations. This simple method was successful, but discarding measurements usually does not lead to optimal solutions. Moreover, especially if little knowledge about the exact configuration of the robot exists, it is difficult to decide which equations should be discarded. Therefore, a better method for computing \( c^\circ \) has been sought.

[Maryniak, Graefe 1998] have proposed one such approach for utilizing the redundant sensor data. Its key point is that, before computing the robot control commands, the

\[ \text{We denote vectors by underlining and matrices by capital letters.} \]
sensor data space is transformed into a space of the lowest dimensionality that is appropriate for the task. Thus, it avoids solving an over-determined system of equations. Nevertheless, all measurements are taken into account, which gives the robot an improved performance in many different setups.

The method presented here possesses an even better performance. Like [Maryniak, Graefe 1998], it takes advantage of the redundancy of sensor data, rather than discarding measurements, and it uses it for improving the robustness and adaptability of the robot control. Here however, the setup may essentially be arbitrary and unknown.

In order to understand the problem we are considering, the reader may for a moment imagine a setup of two identical cameras observing the scene. They are arranged exactly parallel to each other, so that the epipolar line is parallel to the x-axes of both image coordinate systems. In this case the two equations belonging to the y-coordinates of the cameras contain (disregarding noise) the same information; discarding one of them would cause no loss. If measurement noise is present, averaging the 2 equations, rather than omitting one of them, tends to improve the reliability of the result.

On the other hand, omitting one of the other two equations would remove all depth information and would (apart from noise) result in a system without full rank that cannot be solved without further constraints, since we have three equations, where two contain the same information, for three unknowns. If noise is present, a solution can usually be computed, but it has no relevance.

If the camera arrangement is not exactly as described, but only approximately so, the two y-equations contain some relatively noisy depth information, and omitting one of the x-equations leads to an ill-conditioned system which has a unique solution, but where slight variations of the input data cause strong changes of the solution. The solution will, therefore, be rather noisy and not very useful as a basis for robot control. If the camera arrangement is arbitrary and unknown, randomly omitting one of the four equations will lead to results of unpredictable quality.

There exist several methods for solving over-determined equations numerically. One method is to first solve an exactly determined case by arbitrarily omitting rows and then improving the result by considering further rows [Björck, Duff 1980]. Other methods include the use of the pseudo-inverse and provide solutions for a least-square residual [Björck 1996]. However, in such methods either all rows contribute equally to the final solution, or weights are being used which are based on additional information which is not available in our case of calibration-free robots in arbitrary setups. Therefore, we have developed a more appropriate method.

**2. THE SOLUTION**

We tackle this problem in the context of a manipulation task for a calibration-free vision-guided robot. An object that is resting at an initially unknown arbitrary location somewhere in the robot’s work space is to be grasped. The only sensors used are two uncalibrated video cameras in an uncalibrated stereo arrangement (Figure 2).

It is possible to generate a reduced system of equations that is not over-determined from (1) by omitting \(m-n\) equations. In fact, this may be done in \(k\) different ways, which leads to \(k\) different reduced systems

\[
\mathbf{G}_i^{(m-n)}, \mathbf{J}_i, \mathbf{G}_i, \mathbf{J}_i, 1,...,k
\]

yielding one solution \(\mathbf{c}_i^o\), each. Ideally, all these systems should have the same (correct) solution. Actually, each equation is affected by its own measurement noise, causing the solutions to differ from each other. Some of the systems may be ill-conditioned and yield solutions that mainly reflect the effects of noise. As the conditioning of such a reduced system improves, it becomes less error-prone and its solution becomes more reliable.

To take the best advantage of all available information, we compute the final control word vector, \(\mathbf{c}_i^o\), as a weighted average of all solutions of the reduced systems of equations (2):

\[
\mathbf{c}_i^o = \frac{\sum_j w_j \cdot (\mathbf{c}_i^o, G_{ij})}{\sum_j w_j}
\]

(4)

where the weights, \(w_j\), should reflect the “quality” of the corresponding equation. Since no further assumptions shall be made, no further knowledge is available, other than the systems of equations (2) themselves. Therefore, the weights can only be derived from them.

As an indicator for the quality, the condition number, \(\gamma\), can be used. It is a numerical measure for the conditioning of a problem [Schmeiøer, Schirmeyer 1976], that is in our case represented by the truncated sensor-control Jacobians, \(\mathbf{J}_i\). The larger the condition number is, the more sensitive is \(\mathbf{J}_i\) [Anderson et al. 1995], and, thus, the associated solution for the control word vector, \(\mathbf{c}_i^o\). The solution of the respective subsystem will increasingly be disturbed by (even small) measurement noise, which means a poor quality.

We investigated the two-norm condition number \(\gamma\). The two-norm\(^2\) represents the largest amplification of length that the respective matrix is able to induce on a vector [Press et al. 1995]. The condition number may, according to [The Math-Works 1993] and [Shahian, Hassul 1993], rather efficiently be computed by the ratio of the largest to the smallest singu-

\(^2\) For matrices, it is exactly termed the matrix norm subordinate to the vector two-norm [Bjöck 1996].
lar values. However, the computational expense is still considerable.

But there is another possibility. Usually, for an ill-conditioned system the absolute value of the determinant of the system matrix, in this case the truncated sensor-control Jacobian matrix, is very small compared to the one of a well-conditioned system. With an improved conditioning and, hence, a smaller condition number, the quality of a reduced system grows, and the absolute value of the corresponding system determinant as well. At the same time, the system becomes less error-prone.

[Golub, van Loan 1989] state that a bad condition number does not generally correspond to a small absolute value of the determinant. In fact, cases can be constructed for which the opposite is true; but in actual applications such cases are apparently not encountered.

Hence, for the calculation of the weights in (4) two alternatives exist, the reciprocal of the (two-norm) condition number

\[ w_i' = \frac{1}{\|J_i\|_2} \quad (5) \]

and the absolute value of the determinant

\[ w_i' = \frac{1}{|\text{det}(J_i)|} \quad (6) \]

of the truncated Jacobian.

In Matlab, for the calculation of the condition number about 20 times more floating point operations are necessary than for the calculation of the absolute value of the determinant. Although this should, generally, not be exactly the same for any particular implementation, it is obvious that for real-time applications preference should be given to the “Determinant Weighted Averaging” if the two possible weights are, indeed, similar. Saving computation time is especially important for later extensions to more degrees of freedom of the manipulator, or more image features, since the number of reduced systems, \( k \), according to (3) then grows rapidly.

Basically, in both cases all equations and, thus, all measurements are being used in computing the control word vector. This is an advantage because usually even a noisy or redundant measurement contains some useful information. On the other hand, if a subsystem lacks full rank it does, practically, not affect the final solution. Consequently, solutions of ill-conditioned subsystems enter the final solution only with a small weight. Both methods for computing \( w_i \), (5) and (6), allow problems with \( n > 2 \) to be treated.

3. EVALUATION

The method was first evaluated in simulations. Then it was implemented on a vision-guided articulated arm robot (Figures 2 and 4) and tested in real-world experiments. In both cases the sensor-control Jacobian had 4 rows, corresponding to the 4 elements of the image distance vector, \( d_x \) and 3 columns, corresponding to the robot’s 3 degrees of freedom that were actually used; \( m=4 \) and \( n=3 \).

3.1 Simulation

In a first series of simulations (not reported here in detail) it was verified that, indeed, the weights obtained from the determinants and the two-norm condition numbers are very similar. If a subsystem would be assessed suitable on the basis of its condition number, it would be assessed suitable on the basis of the absolute value of its system determinant as well, and the final results would be nearly identical.

Another series of simulations was performed in order to assess the individual solutions of the truncated systems and those of Determinant Weighted Averaging. For a camera arrangement with both cameras at approximately the same height the two systems truncated by either one of the \( y \)-equations yielded good results, while the ones of the two systems truncated by an \( x \)-equation were poor. The quality of the results of Determinant Weighted Averaging were of the same order as those of the systems truncated by a \( y \)-equation (due to their large determinants) while the ones of the systems truncated by an \( x \)-equation had hardly an influence.

More interesting is a series of simulations in which the cameras were arranged in an arbitrary fashion at grossly different heights, contrary to the implicit model used in [Graefe, Maryniak 1998] that assumed both cameras to be at the same heights.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{The position of the gripper in world coordinates (cf. Figure 2) obtained by a simulation when disturbing the distance vector by a uniformly distributed noise in the range of \( \pm 5 \) pixels and deriving weights from the absolute values of the determinants. Above, the solutions of the four truncated systems are shown. The absolute values of the determinants are 38.3, 381.2, 18.4, 314.3 giving normalized weights, \( w_i \), of 0.075, 0.494, 0.024, 0.407 respectively. To the left, the resulting solutions are shown which are obtained by Determinant Weighted Averaging of the individual solutions of the four subsystems.}
\end{figure}
height. This camera arrangement simulates the effect of some unmodeled modification of the sensor arrangement, as it may happen for any number of reasons.

In a situation like the one shown in Figure 2 the sensor-control Jacobian matrix, \( J \), and the image distance vector, \( d \), were measured on the real robot. A uniformly distributed, uncorrelated pseudo-random noise (±5 pixels) was then added to the elements of \( d \), and the individual solutions of the truncated systems, \( \xi^0 \), as well as the solution according to Determinant Weighted Averaging, \( \xi^0 \), were calculated. Figure 3 shows the results in world coordinates for 100 such simulations. (It should be stressed that the world coordinates are introduced for description purposes only; they are in no way used in the calculation of the control word vectors.)

Figure 3 shows two important facts: First, the solutions of the subsystems are different, although they should be identical. This is due to non-linearities and to the fact that the Jacobian had been measured and, therefore, contains errors, even without any additional noise from the simulation. Second, the results are to different degrees prone to variations in the vector components, \( d \), caused by the additional noise. Both facts are related to the quality of the subsystems.

It can clearly be seen that in the subsystems in which a \( y \)-coordinate, \( d_y \) or \( d_z \), is omitted the influence of noise is by far smaller than if an \( x \)-coordinate, \( d_x \) or \( d_z \), is omitted.

If one of the four reduced systems is randomly chosen, as proposed by Graefe, Maryniak 1998, it is a matter of luck whether the result is good or bad. Of course, if we knew the actual sensor arrangement, an intelligent choice could be made; but here we are dealing with the case that such knowledge is not available because the sensor arrangement may have been changed accidentally. The solutions obtained by Determinant Weighted Averaging are equal to or better than the individual ones (Figure 3). The standard deviations of the subsystems are different, although they should be identical. This is due to non-linearities and to the random modifications of the camera arrangement. In a first series of experiments the cameras were mounted at approximately the same height. In a second series of experiments a height offset between the cameras was introduced (Figure 2) to study the effects of such an unmodeled disturbance.

In a slight deviation from the concept described in the introduction, we perform the grasping in three phases: first the closed gripper is moved to an intermediate position, a few cm above the object, then the gripper is opened, and finally the object is grasped. To minimize the risk of a collision between the closed gripper and the object, the magnitude of motion is limited by using (7) instead of (4) for computing the control word vector when the intermediate position is being approached:

\[
\xi_j^0 = \frac{\sum_j (\xi_j^0, \Theta_j^0)}{\max_j \left( \frac{w_j}{500} \right)} .
\]

In the experiments a support of unknown height carrying a dark cylindrical object was placed arbitrarily on the work table, and a grasping process was performed. We observed whether an intermediate position somewhere above the object, suitable for initiating the final grasping, could be reached and how many control steps were necessary for this (Figure 5). Furthermore, we observed whether the grasping was successful or not. For a statistical evaluation we divided the work area into a set of 4*8 square fields (3.5 cm * 3.5 cm). Ten grasping experiments were performed for each field where the object was placed at arbitrary positions within the field, and averages were computed for each field.

Figure 5 shows the results for the first setup where the cameras were arranged at approximately the same height. The results of the experiments with the cameras at different heights is similar, apart from a slightly changed working area of the robot due to the different poses of the cameras.

<table>
<thead>
<tr>
<th></th>
<th>Standard deviation [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Det. Wgt. Aver.</td>
<td>1.68</td>
</tr>
<tr>
<td>system truncated by ( d_1 )</td>
<td>1.62</td>
</tr>
<tr>
<td>system truncated by ( d_2 )</td>
<td>1.44</td>
</tr>
<tr>
<td>system truncated by ( d_3 )</td>
<td>19.7</td>
</tr>
<tr>
<td>system truncated by ( d_4 )</td>
<td>1.43</td>
</tr>
</tbody>
</table>

**Table 1**
The gripper positions’ Standard deviations by Determinant Weighted Averaging and truncating measurements. (For the definition of the \( d_i \), cf. Figure 1.)

**3.2 Real-World Experiments**

We have implemented our solution on an articulated arm robot, Mitsubishi Movemaster RV-M2 (Figures 2 and 4), possessing 5 degrees of freedom, corresponding to the 5 joints \( J_1 \) to \( J_5 \) (Figure 4). \( J_1 \) to \( J_5 \) are the joints that were actively controlled in our experiments. To keep the gripper vertical, \( J_6 \) is controlled by a priori knowledge. \( J_6 \) (rotation of the gripper around its axis) is not used here because our objects are flat circular disks that can be grasped with any orientation of the gripper around its vertical axis. The cameras are attached to the robot on metal rods at the first link, causing a co-revolution of them with \( J_1 \). 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. In a first series of experiments the cameras were mounted at approximately the same height. In a second series of experiments a height offset between the cameras was introduced (Figure 2) to study the effects of such an unmodeled disturbance.
Table 2

Determinant Weighted Averaging compared to truncating measurements for the cameras arranged at approximately the same height (first series) and at different heights (second series). Ten grasping tasks per field were performed and averages over all squares are shown. If no values are available this is indicated by n.v.

<table>
<thead>
<tr>
<th></th>
<th>First series: cameras at approx. same height</th>
<th>Second series: cameras at different heights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>average number of successful approaches</td>
<td>average number of successful approaches</td>
</tr>
<tr>
<td></td>
<td>standard deviation of the number of control steps</td>
<td>standard deviation of the number of control steps</td>
</tr>
<tr>
<td>average number of control steps necessary</td>
<td>average number of control steps necessary</td>
<td></td>
</tr>
<tr>
<td>Det. Wgt. Aver.</td>
<td>2.01 0.41 9.73 9.73</td>
<td>1.08 0.14 8.91 8.91</td>
</tr>
<tr>
<td>system truncated by (d_x)</td>
<td>3.42 1.27 4.41 8.14</td>
<td>n.v. n.v. 0 0</td>
</tr>
<tr>
<td>system truncated by (d_y)</td>
<td>2.15 0.5 9.73 9.82</td>
<td>1.23 0.16 9.45 9.91</td>
</tr>
<tr>
<td>system truncated by (d_z)</td>
<td>3.4 1.07 4.41 7.27</td>
<td>1.35 0.04 3.1 4.2</td>
</tr>
<tr>
<td>system truncated by (d_d)</td>
<td>2.13 0.46 9.27 9.86</td>
<td>1.93 0.27 6.5 10</td>
</tr>
</tbody>
</table>

4. DISCUSSION

When comparing the results obtained by Determinant Weighted Averaging to the ones obtained by the previous method, the following differences are noticeable:

- The new method yields good and reliable results, even if the sensor arrangement is severely disturbed without the system being aware of it.
- For the older method the results are much less consistent, even if the cameras are in their correct position.

C Omitting an \(x\)-information \((d_x \text{ or } d_y)\), basically gives useful solutions, but they are often too inaccurate, causing the grasp to fail. Therefore, the average number of successful approaches to the intermediate position shows unacceptable results, even though visual feedback compensates for some of this inaccuracy. Moreover, in about half of all cases the final grasping (presently still executed without visual feedback) fails, since the accuracy required for an open-loop control of a nonlinear plant is not provided here. Also, for successful grasping processes the average number of control steps necessary and its standard deviation are significantly higher than for the systems truncated by a \(y\)-information.

C Omitting a \(y\)-information \((d_y \text{ or } d_z)\), the system still performs well in the first setup (cameras at equal height). With a disturbed setup, however, the performance of the system truncated by \(d_y\) decreases, as indicated by about
a third of misgrasps and a considerable number of necessary control steps. Out of the reduced systems, the system truncated by \( d \) gives the best results. This system could be the intelligent choice in the case of the disturbed setup, if the disturbance were known in advance. However, since it is not known, and since the setup could also have been disturbed in a totally different way, such an \textit{a priori} choice is not possible.

\(<\)Although the new method takes into consideration all individual solutions, its performance is not degraded by a negative influence from the ill-conditioned systems; this would have been possible, and even likely, for other methods, e.g., a straightforward application of a least-squares criterion. The main advantages of the new method are:

- C No advance choice of measurements (a matter of luck without \textit{a priori} information, and possibly difficult even with \textit{a priori} information) is necessary.
- C Arbitrary disturbances of the sensor arrangement are tolerated.
- C Since the weights are recomputed each time a new Jacobian is determined, the approach provides an automatic adaptation to changes in the sensor arrangement and to other factors affecting the conditioning of (2).

\(<\)The Determinant Weighted Averaging yields an increase of performance compared to using the best suitable reduced system(s):

- C In the first setup this is obvious from the smaller amount of necessary control steps, although the improvement is rather small.
- C In the second setup, with only 1.08 instead of 1.23 control steps necessary, the performance increase is more obvious. Although a lucky choice performs even slightly better here in terms of successful grasps (due to a better viewing perspective of one of the cameras), this is a matter of luck and, therefore, not relevant.

\(<\)As a result of the new approach the performance of the robot is improved in terms of

- C speed, since fewer control steps are necessary, and
- C reliability, since fewer iterations indicate a higher accuracy in the performance of the individual control commands.

5. CONCLUSIONS

We have developed a concept, Determinant Weighted Averaging, to take advantage of redundant sensory information in the control of calibration-free robots. The concept divides an over-determined system of equations for the numerical parameters of robot control commands into reduced systems, each of them providing a partial solution. These solutions are weighted dynamically according to the quality of the respective reduced system, and averaged to a final solution. For the calculation of the weights we chose the absolute value of the determinant of the respective reduced system.

The concept has been proven in real-world experiments for the task of visually guided manipulation of objects by a robot with a totally calibration-free camera-manipulator system.

Compared to earlier approaches based on truncating information, the new concept improves the performance of the robot in terms of speed and reliability, as well as in robustness and adaptability to arbitrary changes (caused by the task or caused by disturbances) of the robot’s sensor characteristics. The performance improvement occurs even in comparison to the best truncated systems. The need for a hard-to-perform choice of measurements is, thus, eliminated.

The approach is extensible to further degrees of freedom, which is going to be a near-future challenge for us, and to a higher degree of sensory redundancy as well.

ACKNOWLEDGMENT

We thank Mr. Minh Chinh Nguyen for making his image processing software available for our use, and Mrs. Andrea Rott (Institute of Mathematics and Computer Applications) for many fruitful discussions.

REFERENCES


