VISION BASED IN-SITU CALIBRATION OF ROBOTS WITH APPLICATION IN SUBSEA INTERVENTIONS

M.Sc. Thesis

Thesis presented to obtain the M.Sc. title at the Mechanical Engineering Department at PUC-Rio.

Advisor: Marco Antonio Meggiolaro

Rio de Janeiro
September 2007
Trond Martin Augustson

CALIBRAGEM VISUAL IN SITU DE MANIPULADORES ROBÔTICOS COM APLICAÇÃO EM INTERVENÇÕES SUBMARINAS

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Mecânica do Departamento de Engenharia Mecânica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Marco Antonio Meggiolaro
Orientador
Pontifícia Universidade Católica do Rio de Janeiro

Prof. Mauro Speranza Neto
Pontifícia Universidade Católica do Rio de Janeiro

Prof. Raul Queiroz Feitosa
Pontifícia Universidade Católica do Rio de Janeiro

Prof. Fernando Cesar Lizarralde
Universidade Federal do Rio de Janeiro

Prof. José Eugenio Leal
Coordenador Setorial do Centro Técnico Científico – PUC-Rio

Rio de Janeiro, 3 de setembro de 2007
Trond Martin Augustson

Graduated in applied physics at The University of Bergen in 1998. He has worked with seismic surveying until entering the master program at PUC-Rio in 2005.
Thanks to

- My advisor Marco Antonio Meggiolaro, for help and support;

- Professor Raul Feitosa, for the contribution on computer vision;

- PUC-Rio for the opportunity and the great academic environment that constitute the basis of this work;

- CENPES/PETROBRAS for the support, including the submarine camera and information on the TA-40 manipulator;

- My family, however distant, supporting me during my work.
Abstract

The majority of today’s industrial robots are programmed to follow a predefined trajectory. This is sufficient when the robot is working in a fixed environment where all objects of interest are situated in a predetermined position relative to the robot base. However, if the robot’s position is altered all the trajectories have to be reprogrammed for the robot to be able to perform its tasks. Another option is teleoperation, where a human operator conducts all the movements during the operation in master-slave architecture. Since any positioning errors can be visually compensated by the human operator, this configuration does not demand that the robot has a high absolute accuracy. However, the drawback is the low speed and low accuracy of the human operator scheme. The manipulator considered in this thesis is attached to a ROV (Remote Operating Vehicle) and is brought to its working environment by the ROV operator. Every time the robot is repositioned, it needs to estimate its position and orientation relative to the work environment. The ROV operates at great depths and there are few sensors which can operate at extreme depths. This is the incentive for the use of computer vision to estimate the relative position of the manipulator. Through cameras the differences between the actual and desired position of the manipulators is estimated. This information is sent to controllers to correct the pre-programmed trajectories. The manipulator movement commands are programmed off-line by a CAD system, without need even to turn on the robot, allowing for greatest speed on its validation, as well as problem solving. This work includes camera calibration and calibration of the structure of the manipulator. The increased accuracies achieved by these steps are merged to achieve in-situ calibration of the manipulator base.

Key Words

Robotics; Calibration; Computer vision; SIFT; Pattern recognition; Automation; Stereopsis
Resumo


A maioria dos robôs industriais da atualidade são programados para seguir uma trajetória pré-definida. Isto é suficiente quando o robô está trabalhando em um ambiente imutável onde todos os objetos estão em uma posição conhecida em relação à base do manipulador. No entanto, se a posição da base do robô é alterada, todas as trajetórias precisam ser reprogramadas para que ele seja capaz de cumprir suas tarefas. Outra opção é a teleoperação, onde um operador humano conduz todos os movimentos durante a operação em uma arquitetura mestre-escravo. Uma vez que qualquer erro de posicionamento pode ser visualmente compensado pelo operador humano, essa configuração não requer que o robô possua alta precisão absoluta. No entanto, a desvantagem deste enfoque é a baixa velocidade e precisão se comparado com um sistema totalmente automatizado. O manipulador considerado nesta dissertação está fixo em um ROV (Remote Operating Vehicle) e é trazido até seu ambiente de trabalho por um teleoperador. A cada vez que a base do manipulador é reposicionada, este precisa estimar sua posição e orientação relativa ao ambiente de trabalho. O ROV opera em grandes profundidades, e há poucos sensores que podem operar nestas condições adversas. Isto incentiva o uso de visão computacional para estimar a posição relativa do manipulador. A diferença entre a posição real e a desejada é estimada através do uso de câmeras submarinas. A informação é enviada aos controladores para corrigir as trajetórias pré-programadas. Os comandos de movimento do manipulador podem então ser programados off-line por um sistema de CAD, sem a necessidade de ligar o robô, permitindo rapidez na validação das trajetórias. Esse trabalho inclui a calibragem tanto da câmera quanto da estrutura do manipulador. As melhores precisões absolutas obtidas por essas metodologias são combinadas para obter calibração in-situ da base do manipulador.

Palavras-Chave

Robótica; Calibragem; Visão computacional; SIFT; Reconhecimento de padrões; Automação; Visão estéreo
## Summary

1 Introduction 20
   1.1. Motivation 20
   1.2. Work objectives 20
   1.3. Work description 21
   1.4. Organization of the Thesis 24

2 Kinematic Modeling for Calibration of Manipulators 26
   2.1. Introduction 26
   2.2. Basic Concepts of Kinematics 28
   2.3. The Denavit-Hartenberg Convention 29
   2.4. Classic Manipulator Calibration 32
   2.5. Elimination of Redundant Errors 36
   2.6. Physical Interpretation of the Redundant Errors 38
   2.7. Partial Measurement of End-Effector Pose 39
   2.8. Inverse Kinematics 40
   2.8.1. Solvability 40
   2.9. Experimental Procedures 41

3 Computer Vision 47
   3.1. Introduction 47
   3.2. Mathematic Camera Models 48
      3.2.1. The Pinhole Model 48
      3.2.2. Intrinsic Parameters 49
      3.2.3. Extrinsic Parameters 51
   3.3. Camera Calibration 52
      3.3.1. Radial Distortion 56
      3.3.2. Sophisticated calibration 58
      3.3.3. Coordinate extraction 61
      3.3.4. Nonmaximum Suppression 61
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2. Laboratory Experiments</td>
<td>108</td>
</tr>
<tr>
<td>5.2.1. Camera Calibration</td>
<td>110</td>
</tr>
<tr>
<td>5.2.2. Experiments with the X-Y Table</td>
<td>113</td>
</tr>
<tr>
<td>5.3. Calibration of an Underwater Camera</td>
<td>123</td>
</tr>
<tr>
<td>5.4. Position Estimation using the Underwater Camera</td>
<td>125</td>
</tr>
<tr>
<td>5.5. Camera Calibration performed Underwater</td>
<td>132</td>
</tr>
<tr>
<td>6 Conclusions and Suggestions</td>
<td>134</td>
</tr>
<tr>
<td>6.1. Conclusions</td>
<td>134</td>
</tr>
<tr>
<td>6.2. Suggestions for future work</td>
<td>135</td>
</tr>
<tr>
<td>7 References</td>
<td>136</td>
</tr>
<tr>
<td>Appendix A</td>
<td>139</td>
</tr>
</tbody>
</table>
List of figures

Figure 1 - Repeatability and absolute accuracy ........................................... 21
Figure 2 - Coordinate systems of the manipulator .................................... 27
Figure 3 - Denavit-Hartenberg parameters [5] ......................................... 30
Figure 4 - Translation and rotation with effects of errors in the i-th link [6] ........................................................................................................ 32
Figure 5 - Generalized errors for the i-th link. $\varepsilon_{p,i}$, $\varepsilon_{s,i}$, $\varepsilon_{r,i}$ represent the rotation around the x, y and z-axes respectively. [6] ............... 33
Figure 6 – Error compensation block diagram [6] .................................. 35
Figure 7 – Combination of translational linear errors [6] ......................... 38
Figure 8 – Simplified combination of error [6] ......................................... 39
Figure 9 – Finding the rotation axis of joint 2 (Z₁), side view ............... 42
Figure 10 – Finding the rotation axis of joint 1 (Z₀), upper view ............ 42
Figure 11 - The trajectory of the probe forms a plane that is found by a least square approximation .............................................................. 43
Figure 12 - Angles between the laser tracker reference frame and the normal plane [12] ............................................................. 44
Figure 13 - Schematic representation of the pinhole model ................. 48
Figure 14 – Geometry of the pinhole model [14] ..................................... 49
Figure 15 - CCD layout. (a) shows an ideal square, (b) shows that the scale in x and y direction can differ, (c) shows that the axes might not be perpendicular [14]. ................................................. 49
Figure 16 - The modified pinhole model. [14] .......................................... 50
Figure 17 - The image center is not always in the middle of the sensor since the lens normal does not intersect with the middle of the sensor panel............................................................................. 50
Figure 18 – Transformation of world coordinates to camera coordinates. [14] ......................................................................................... 52
Figure 19 – Calibration rig ...................................................................... 53
Figure 20 - Transformation from world coordinates to picture coordinates. [14] ......................................................................................... 53
Figure 21 - Types of radial distortion

Figure 22 - Principle of barrel distortion. The black coordinates represent the image coordinates for a camera without distortion. The violet coordinates show the distorted image coordinates.

Figure 23 - The 4 directions used in nonmaximum suppression

Figure 24 - Output of the nonmaximum suppression algorithm

Figure 25 - Maxima and minima in the Difference-of-Gauss are compared to its 26 neighbors [3]

Figure 26 - Histogram of key-point orientation [3]

Figure 27 - Keypoint calculation process. Each 4x4 element of gradients (left) is referred to as a bin. For each bin a histogram of 8 directions is calculated (right) [3]

Figure 28 - Distribution of relative keypoint orientation

Figure 29 - Triangulation

Figure 30 - Stereo Triangulation

Figure 31 - Initial procedure to estimate the position of the reference camera relative to the keypoints. Creating a set of 3D coordinates, $^1p$.

Figure 32 - Finding a second set of corresponding coordinates, $^2p$

Figure 33 - TA40 and the miniature robot used as master

Figure 34 - TA-40 and coordinate systems [1]

The frame center $O_4$ of joint 5 is located 747mm along the $z_4$ axis from $O_4$, giving $d_4=747$. Since the frame centers position along the common normal is zero, $a_4=0$. The $z_4$ axis is rotated $-90^\circ$ relative to $z_3$, giving $\alpha_4=-90^\circ$.

Figure 35 - A 2D interpretation of the frames $O_2,O_3$ and $O_4$. Frame $O_5$ coincides with frame $O_4$ [24]

Figure 36 - Position error of end-effector after calibration

Figure 37 - Rotation error at the end effector after calibration

Figure 38 - Position error of link 5 after calibration

Figure 39 - Rotation error of link 5 after calibration
Figure 40 – x-y table in the Robotics laboratory at PUC......................... 109
Figure 41 – Calibration rig....................................................................... 109
Figure 42 – Edges of the calibration rig. The estimated corners are marked in red. ................................................................. 110
Figure 43 – Figure showing the edges of the image in white. The image to the left shows the initial k-mean line parameters. The image on the right shows the improved estimate after only three iterations. The estimated coordinate of the corner is marked in red. The yellow and green lines show the estimated lines. .................................................................................. 111
Figure 44 – Relative movement between extracted coordinates and the projected coordinates................................................................. 113
Figure 45 - The robot used as reference object in the experiment........... 114
Figure 46 – Coordinate system of the x-y table ...................................... 114
Figure 47 – Method to estimate the set of coordinates, \( ^2p \). The reference image is marked in yellow. The position of one of the two cameras is to be estimated relative to the reference image. .... 118
Figure 48 – Triangulation to estimate the coordinate set \( ^1p \) relative to the origin. The reference image is marked in yellow. The images closest to the reference image and the two cameras used to estimate the coordinate set \( ^2p \) were not used. .................. 118
Figure 49 – RMS position error as a function of intraocular distance using RANSAC and quaternion rotation estimation....................... 120
Figure 50 – Graph showing the average RMS position error as a function of the intraocular distance after the samples with the worst error ratio had been removed......................................................... 121
Figure 51 - RMS position error as a function of intraocular distance using RANSAC with LMS rotation estimate...................................... 122
Figure 52 - RMS position error as a function of intraocular distance using RANSAC with LMS rotation estimate and eliminating the coordinates with a high error ratio. .................................................. 123
Figure 53 – The corrected image coordinates relative to their respective extracted image coordinates ................................................. 125
Figure 54 – Attachment support for the underwater camera. ................. 126
Figure 55 – Position accuracy as a function of the intraocular
distance using all 72 coordinates and quaternion rotation
estimation. ........................................................................................................ 128
Figure 56 - Position accuracy as a function of the intraocular
distance after the coordinates with a large error ratio had been
eliminated........................................................................................................ 129
Figure 57 - Position error as a function of intraocular distance with
the underwater camera using all 72 coordinates and LMS
rotation estimation.......................................................................................... 130
Figure 58 - Position accuracy as a function of the intraocular
distance after the coordinates with a large error ratio had been
eliminated........................................................................................................ 131
Figure 59 – Underwater calibration. The green points are the
corrected image coordinates and the red lines show their
respective image coordinates ........................................................................ 133
Figure A.1 – Coordinate systems of the TA-40...................................... 139
List of Tables

Table 1 – Denavit-Hartenberg parameters ............................................................... 94
Table 2 – Errors from simulation ........................................................................... 103
Table 3 – Camera calibration parameters ................................................................. 112
Table 4 – Position parameters for the experiment ................................................. 115
Table 5 – Position error as a function of intraocular distance using
RANSAC ........................................................................................................... 119
Table 6 – Position error as a function of intraocular distance using
RANSAC and error ratio elimination .................................................................. 120
Table 7 – Position error as a function of intraocular distance using
RANSAC together with the least mean square estimated
rotation matrix. .................................................................................................... 121
Table 8 – Position error as a function of intraocular distance using
RANSAC together with the least mean square estimated
rotation matrix and eliminating the coordinates with error ratio.................... 122
Table 9 – Calibration parameters for the underwater camera in air,
where \( s_u \) is the aspect ratio, \( f \) is the focal length, \( k_1 \) and \( k_2 \) are
the radial distortion coefficients, \( T_1 \) and \( T_2 \) are the tangential
distortion coefficients and \((u_0, v_0)\) denotes coordinates of the
image center. ...................................................................................................... 124
Table 10 – Position parameters for the experiment with the
underwater camera ............................................................................................... 127
Table 11 – Position error as a function of intraocular distance using
the underwater camera, all 72 coordinates, and quaternion
rotation estimation .......................................................................................... 128
Table 12 – Position accuracy as a function of the intraocular
distance after the coordinates with a large error ratio had been
eliminated .......................................................................................................... 129
Table 13 – Position error as a function of intraocular distance using
the underwater camera using all 72 coordinates and LMS
rotation estimation .......................................................................................... 130
Table 14 - Position accuracy as a function of the intraocular distance after the coordinates with a large error ratio had been eliminated.

Table 15 – Calibration parameters for the underwater camera in air, where $s_u$ is the aspect ratio, $f$ is the focal length, $k_1$ and $k_2$ are the radial distortion coefficients, $T_1$ and $T_2$ are the tangential distortion coefficients and $(u_o, v_o)$ denotes coordinates of the image center.
List of Variables

Aₐ – Observation matrix used in quaternion rotation estimation
Cₐ – Observation matrix for plane estimation
Cₐᵣ – Observation matrix for arc coordinates projected onto the estimated plane
Cₜ – Observation matrix for translated coordinates used to estimate circle parameters
Gₑ – Non singular Identification matrix where the redundant errors have been eliminated.
₇₀ₐ – homogeneous matrix 4x4 that describes the orientation and position of the manipulator end-effector relative to its base as a function of the angles of the links ₀ₗ and the generalized errors ₑ.
ₐᵢ – Denavit-Hartenberg parameter: length of the common normal between two adjacent links
ₐ – quaternion vector
ₐ₀,ₐ₁,ₐ₂,ₐ₃ – quaternion vector components of ₐ.
ₐₚ – parameter for an estimated 3D plane
ₐₗ – line parameter used to estimate a line through least square approximation
ₐₚ – parameter for an estimated 3D plane
ₐₙ – quaternion vector
ₐ₀,ₐ₁,ₐ₂,ₐ₃ – quaternion vector components
ₐₙ – parameter used to estimate a circle from translated arc parameters
ₐₙ – vector used in quaternion rotation estimation
ₐₚ – line parameter used to estimate a line through least square approximation
ₐₚ – parameter for an estimated 3D plane
ₐₚ – line parameter used to estimate a line through least square approximation
ₐₚ – parameter for an estimated 3D plane
ₐₚ – estimated radius of projected circle
dₘₙₒₚ – variable used in deduction of triangulation principles
dₚ – parameter for an estimated 3D plane
\( d_i \) – Denavit-Hartenberg parameter: distance between the origin \( O_i \) and \( H_i \)

\( d_{\text{int}} \) – Intraocular distance between two camera centers

\( d_n \) – distance limit that denotes the maximum allowed distance from the nearest corresponding coordinate

\( d_{r,j} \) – a coordinates’ distance from the cluster center \( \hat{i}_s p \)

\( e_0 \) – matrix containing the edge angle of an image

\( i \) – complex quaternion unit vector

\( j \) – complex quaternion unit vector

\( k \) – complex quaternion unit vector

\( \mathbf{k} \) – quaternion rotation axis given by \( \mathbf{k} = [i \ j \ k] \)

\( n_p \) – normal vector for estimated plane

\( \hat{p} \) – coordinate in the normalized in the normalized image plane

\( ^1p \) – 3D coordinate set relative to the origin

\( ^2p \) – 3D coordinate set corresponding to \( ^1p \) with a different reference frame

\( \hat{i}_s p \) – translated coordinate set, used in RANSAC algorithm

\( \hat{i}_s p \) – estimated center of the translated coordinate set \( \hat{i}_s p \)

\( q \) – quaternion vector

\( q_0, q_1, q_2, q_3 \) – quaternion vector components of \( q \)

\( r_{r,j} \) – error ratio of a coordinate, denoting its position error relative to its distance from the center of the cluster.

\( r_{\text{lim}} \) – error ratio limit, denoting the accepted error ratio \( r_{r,j} \)

\( r_c \) – estimated radius of projected circle

\( \hat{u} \) – normalized x-coordinate of an image relative to the image center

\( \hat{v} \) – normalized y-coordinate of an image relative to the image center

\( \hat{u} \) – image x-coordinate corrected for radial distortion

\( \hat{v} \) – image y-coordinate corrected for radial distortion

\( u_t \) – output vector from least square estimate of a circle

\( v_p \) – least square vector estimated by a least square of \( C_a \)

\( x_a \) – x coordinate of an arc projected onto plane
$x_i$ – coordinate of joint i in the Denavit-Hartenberg notation

$x_q$ – substitution parameter used in quaternion deduction

$x_r$ – estimated x-coordinate of circle center

$x_t$ – translated x-coordinate used to estimate circle parameters

$x_{tri}$ – coordinate defining a point used in triangulation

$y_a$ – y-coordinate of an arc projected onto plane

$y_q$ – substitution parameter used in quaternion deduction

$y_r$ – estimated y-coordinate of circle center

$y_t$ – translated x-coordinate used to estimate circle parameters

$y_{tri}$ – coordinate defining a point used in triangulation

$y_i$ – coordinate of joint i in the Denavit-Hartenberg notation

$z_i$ – coordinate of joint i in the Denavit-Hartenberg notation

$z_{tri}$ – coordinate defining a point used in triangulation

$C_a$ – observation matrix of arc coordinates relative to the laser tracker

$C'_a$ – matrix containing the projected coordinates of $C_a$

$G$ – matrix containing the edge magnitude of an image

$G_x$ – matrix containing the edge magnitude in x direction of an image

$G_y$ – matrix containing the edge magnitude in y direction of an image

$S_x$ – Sobel filter mask for detecting edges in x direction

$S_y$ – Sobel filter mask for detecting edges in y direction

$L_o$ – observation matrix for camera calibration

$M$ – projection matrix of the pin hole camera

$M_v$ – projection matrix in vector form

$^cO$ – camera reference frame

$^wO$ – world reference frame

$^cP$ – coordinate relative to the camera reference frame

$^wP$ – coordinate relative to the world reference frame

$Q$ – skew matrix used to calculate vector product

$R$ – rotation matrix defining the relative rotation between two views

$\Delta X$ – difference between desired position of the manipulator end-effector and the actual position.
$\Delta X_i$, matrix containing differences between desired position of the manipulator end-effector and the true measured position.

$J_t$ - The matrix $6m \times 6(n+1)$ formed by $m$ Identification Jacobians, called the Total Identification matrix

$\alpha$ – magnification factor in x direction for the pin hole model [pixels]

$\beta$ – magnification factor in y direction for the pin hole model [pixels]

$\alpha_i$ – Denavit-Hartenberg parameter: the angle between the joint axes in the right hand sense.

$\hat{e}$ - vector containing the estimated generalized errors

$e'$ - vector where the redundant errors are incorporated in the non redundant errors

$\varepsilon_{x,i}$ - Generalized error of joint $i$ along x-axis

$\varepsilon_{y,i}$ - Generalized error of joint $i$ along y-axis

$\varepsilon_{z,i}$ - Generalized error of joint $i$ along z-axis

$\varepsilon_{p,i}$ - Generalized rotational error of joint $i$ around x-axis

$\varepsilon_{q,i}$ - Generalized rotational error of joint $i$ around y-axis

$\varepsilon_{r,i}$ - Generalized rotational error of joint $i$ around z-axis

$\gamma_1, \gamma_2$ - multiplication factors that determines at what coordinate the closest mutual point is found for two lines in 3D

$\theta_i$ – Denavit-Hartenberg parameter: the angle between the $x_{i-1}$ axis and the common normal $H_iO_i$ measured along the z-axis

$\theta_{p,x}$ – estimated angle around the $x$-axis, between the normal plane and the reference frame.

$\theta_{p,y}$ – estimated angle around the $y$-axis, between the normal plane and the reference frame.

$\theta_q$ – quaternion rotation angle

$\theta_{x_0}$ - camera bias angle around $x$-axis

$\theta_{y_0}$ - camera bias angle around $y$-axis

$\theta_{z_0}$ - camera bias angle around $z$-axis

$\zeta$ – substitution variable used in estimation of quaternion rotation angle

$\rho$ – substitution variable used in estimation of quaternion rotation angle