International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 03 - Issue 01 || January 2017 || PP. 63-68

Estimation of position and orientation of the target point from image information for robot controlling

Anish Mon T.S[1], Prof. A.R Peter[2]
Department of Mechanical Engineering ,
NSS College Of Engineering, Palakkad.

Abstract: This paper presents the aspects of finding the coordinates of the robot end-effector target point. In manufacturing industries were using robots for many applications like assembling, drilling etc.in this processes object recognition, object tracking are necessary operations to perform. Object tracking in this area constitutes an important issue for vision assisted robot control. This paper describes the real time computation of relative pose estimation of robot end effector with respect to the targeted objects using image information and an algorithm to find position and orientation of the target. The algorithm should use as few feature points as possible during estimation such as centre of gravity (cog) of blobs, pixel coordinates in the image plane. The pose estimation is done through Dementhon linear approach and the iterations are done to get the pose of the objects till converges to the solution with lower error vector.
Keywords: pose, eye in hand, perspective projection.

Introduction

In computer integrated industrial automation, the environment that includes the robots need to respond with unknown situations, to make such system possible the first priority is given to collect information from the world surrounds, the sensor best suited for this is incorporating a vision camera. Nowadays it is very easy for acquiring the image and execution of that image from the camera. With the visual feedback or image from the camera is used to identify and locate the target. This could be done by localizing the target in the image by using image features; they can be points, circles, grey level regions, lines etc. The major task in visual feedback is to find the position and orientation of the target for controlling the robot end effector. There by compensating the positioning errors with the unstructured environment and enhancing the robot controlling performance.

Image based pose estimation

Image based pose estimation for controlling the robot is a closed loop control of the system. This controlling system consists of two processes; tracking and controlling, tracking finds the real time estimation and regularly updating the image features with the robot movement[2]. Depending upon the tracking, controlling procedure is generated. The image based tracking was first introduced by Weiss and Sanderson [8]. In image based pose estimation and controlling features are directly derived from the camera image. The simplest feature that can be taken from an image is a point on the image. The controlling is basically done based on the error between current and the required / desired target in the image features[3].

Compared with other vision assisted controlling of robot, image based visual serviong has certain advantages [4]. Since image based controlling of robot is carried out on the basis of image features directly it will reduce the required for computation, no need of 3D reconstruction of target objects and eliminate errors in the modelling.

Camera configuration in image based robot controlling

There are two basic configurationsavailable for camera mounting on the robot. One is installing the camera at the end of robot end effector (eye in hand). The other one is setting the camera globally i.e. the camera is mounted in the workspace (eye to hand) means the camera and the manipulator fixed separately.

In this paper we are dealing with eye in hand configuration, which used to take images from the neighbourhood of the manipulator.

International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 03 - Issue 01 || January 2017 || PP. 63-68

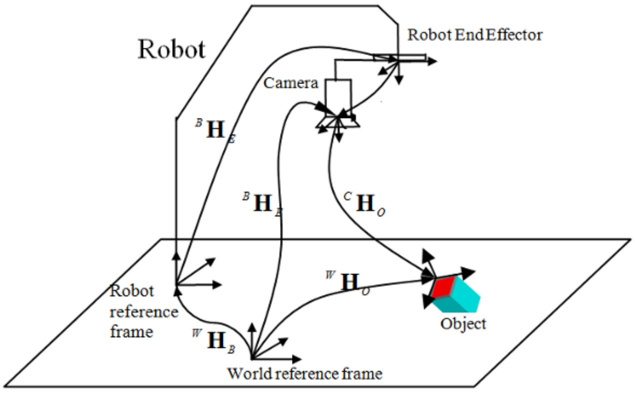


Fig. Eye in hand camera configuration [1]

Pinhole camera model and perspective projection

The pin hole camera model uses with the relationship between the 3D coordinates of a point in the object onto an image plane. Pinhole camera model does notuse lenses to focus light and aperture is described as a point. The advantages of using this camera model is geometric distortions or blurring of unfocused objects caused by lenses[5] and the camera model can uses as first order approximation of the projection of 3D image to 2D image.

Perspective projection is the mapping from 3D scene to a flat plane as it is seen by camera. The main characteristic of perspective projection is that, an object in the image plane becomes smaller as their distance from the camera increases[6].

Pose estimation

In image based pose estimation the feature of image such as points, circles, lines, grey level regions (blob) are used as the input for the pose estimation. Finding the position and orientation process includes the determination of rotation and transformation of the target with respect to the camera coordinates. The pose of the target is determined by treating the image features with the desired pose for that we are using linear and non linear approximate methods. Dementhon approach [7] is one of them. In Dementhon approach an initial volume of pose space is assumed and the whole correspondences compatible with this volume are taken into the formulation. Then the initially assumed volume of pose is reduced recursively till it can be calculated as a single pose[8].

Dementhon approach requires a point in the object with a known image be selected as the origin of the coordinate system. Consider a pinhole camera having 'f' as focal length and an image feature point p with Euclidean coordinates (x,y) and homogeneous coordinates ($\omega_x \omega_y \omega$). The point p is the perspective projection of the 3D object point P with homogeneous coordinates P = (X, Y, Z, 1) ^{T}[10]. Take the object reference origin as P_0 in the camera frame The Euclidean coordinates of P in the object frame are represented by the vector P = $(X, Y, Z)^T$ from P_0 to P.

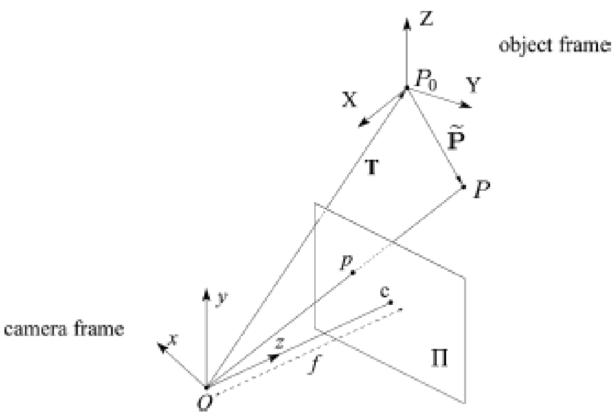


Fig. Camera geometry

The above figure shows camera with center of projection O, focal length f, image center c, and image plane $\prod$, projects object point P onto image point p. The translation between the camera frame and the object frame is T from the origin is at P_0 with respect to the camera frame.

The main task in this problem is to find an unknown coordinate transformation between the object and the camera, represented by a rotation matrix $R = [R_1R_2R_3]^T$and a translation vector $\mathbf{T} = (Tx , Ty , Tz)^T$. The vectors $R^T{}_1$,$R^T{}_2$, $R^T{}_3$ are the row vectors of the rotation matrix; they are the unit vectors of the camera coordinate systemexpressed in the object coordinate system. The translation vector **T** is the vector from the center of projection *O* of the camera to the origin P_0 of the object[7]. The coordinates of the image point p is shown to be relating with the object point P by

$$\begin{bmatrix} \omega x \\ \omega y \\ \omega \end{bmatrix} = \begin{pmatrix} fR^T{}_1 & fT_x \\ fR^T{}_2 & fT_y \\ R^T{}_3 & T_z \end{pmatrix} \begin{bmatrix} P \\ 1 \end{bmatrix}$$

We can multiply the elements of the perspective projection by a matrix $\frac{1}{Tz}$ since the image points are defined up to a multiplicative constant. We also introduce the scaling factor $s = \frac{f}{Tz}$. We obtain

$$\begin{bmatrix} \omega x \\ \omega x \end{bmatrix} = \begin{pmatrix} sR^T{}_1 & sT_x \\ sR^T{}_2 & sT_y \end{pmatrix} \begin{bmatrix} P \\ 1 \end{bmatrix} \qquad (1)$$

$$\omega = R_3 \bullet \frac{P}{T_z} + 1 \qquad (2)$$

Here the dot product of R_3.Pis the projection of vector to the camera axis and the P is defined in the object coordinate system. The vector R_3 is the unit vector along the axis of the camera. The dot product of R_3.P will be small when the depth ranges of the object being small respectto the object distance from the depthTz[11]. Therefore the ω can be equated to unity. Then the prospective projection gives the following transformation matrix.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{pmatrix} sR^T{}_1 & sT_x \\ sR^T{}_2 & sT_y \end{pmatrix} \begin{bmatrix} P \\ 1 \end{bmatrix} \qquad (3)$$

From this expression it can be said that it defines a scaled orthographic projection of the object point P. where ‘s’ is the scaling factor for this projection of point[12]. When the scaling factor s will be equal to one, then the equation becomes a transformation of points from an object coordinate system to camera coordinate system and if the value of s is other than one the image is scaled and approximates to a perspective image of the object because thescaling factor is inversely proportional to the depth distance from the centre of projection of camera (*O*) to the origin of the object point$P_0$and it can be written as, $s = f / T_z$.

Hence the equation can be rewritten as

$$\begin{bmatrix} X & Y & Z & 1 \end{bmatrix} \begin{pmatrix} sR_1 & sR_2 \\ sR_X & sT_y \end{pmatrix} = \begin{bmatrix} \omega x & \omega y \end{bmatrix} \qquad (4)$$

In this equation the corresponding homogeneous coordinate ω of each point in the image p with coordinates (x,y) became found from the previous steps and ω is known. There by it is possible to find ωx , ωy .From the equation (4) it is clear that the image components ω x & ω y and the object orientation XYZ of the point P are known and the only unknown components are the pose components sR_1, sR_2, sT_x, sT_y. These components can be found by considering N object points P_k where k= 1,2…..N. With the Euclidean coordinates $P_k = (X_k, Y_k\ Z_k)^T$ and their projection of points on the image plane as pk and the homogenous components of

ωk, from these points we can write two linear system of equations. Solve these equations for the unknown components of pose i.e. sR_1, sR_2 and sT_x, sT_y.

Since the pose components are calculated, from these values it is possible to find the s, R_1,R_2 by applying that R_1 and R_2 are unit vectors and after that R_3 can be found by the cross product of $R_1$and R_2.

$$s = (|sR_1||sR_2|)^{1/2}$$

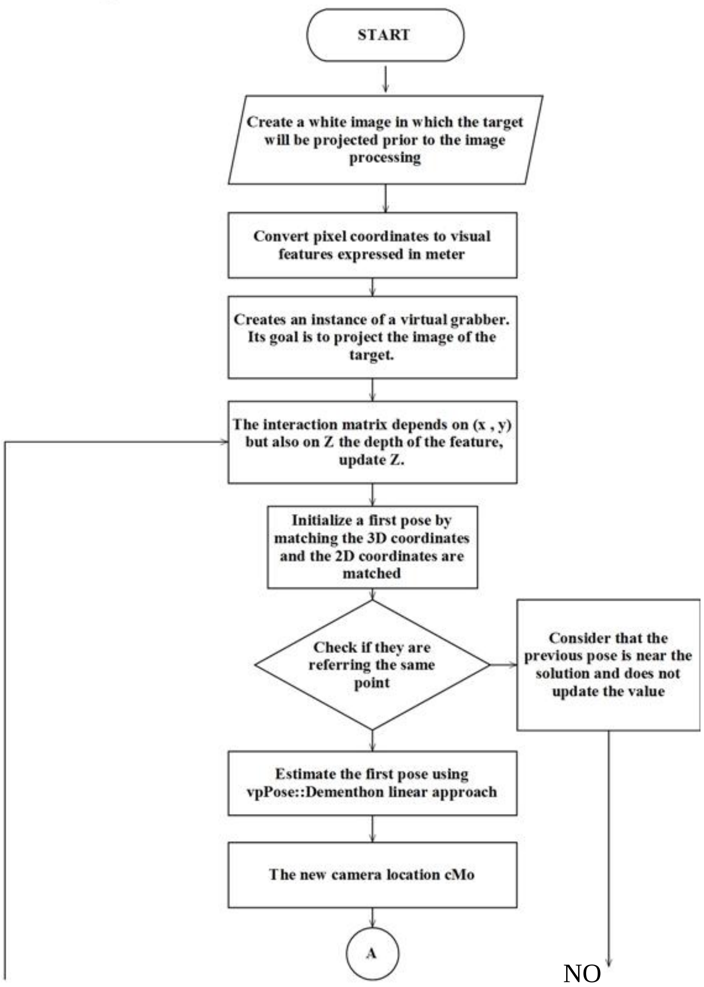
$$R_1 = (sR_1)/s, \mathrm{R}_2 = (sR_2)/s$$

$$R_3 = R_1 \times R_2$$

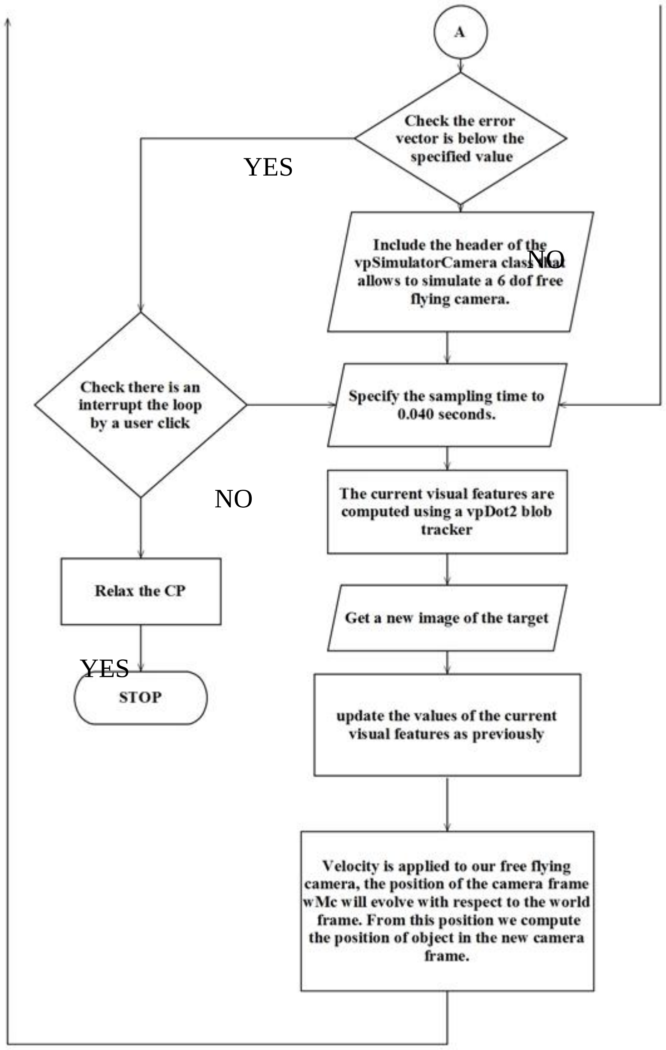
$$T_x = (sT_x)/s, T_Y = (sT_y)/s, T_z = f/s$$

The homogenous components in the equation (4)canbe calculated by consider N points, the homogenous components are $(\omega k \quad xk, \omega k \quad yk)$ corresponding to the point on the image plane P_k, then we can assume that ωk = 1 for every point in. So that it projects a scaled projection and it is good for solving the homogeneous components because we are known with the pose components for this assumption andthere by using equation (2) better results of wk can be found after that solving the linear system of equations of equation (4) refines the pose of thepoints [20].By repeated iteration of this solution, the results will be constant and exact pose of the pointwill be obtained.

Pose Estimation Algorithm



YES



Conclusion

This paper described the algorithm for estimating the pose of the objects as seen by the camera image information using the Dementhon linear approach. The pose calculation combines the efficient iterative processes in the algorithm and this algorithm can be used for robot positioning and tracking in manufacturing industries as a vision feedback in the vision assisted robots. This algorithm can perform well in clutter and noisy environments because it uses image filtering techniques and blob tracking for object detection. Future works that can perform in this area is by involving in this algorithm to work with the lines instead of points in the image.

References

[1]. Wen-Fang Xie1, Zheng Li1, Claude Perron2 and Xiao-Wei Tu1, Switching Control of Image Based Visual Servoing in an Eye-in-Hand System using Laser Pointer,
[2]. Baird, H.S. 1985. *Model-Based Image Matching Using Location*.MIT Press: Cambridge, MA.
[3]. Beis, J.S. and Lowe, D.G. 1999.
[4]. Indexing without invariants in 3D object recognition. *IEEE Trans. Pattern Analysis and MachineIntelligence*, 21(10):1000–1015
[5]. Beveridge, J.R. and Riseman, E.M. 1995. Optimal geometric modelmatching under full 3D perspective. *Computer Vision and Image Understanding*, 61(3):351-364.
[6]. Burns, J.B.,Weiss, R.S., and Riseman, E.M. 1993. View variation of point-set and line-segment features. *IEEE Trans. Pattern Analysisand Machine Intelligence*, 15(1):51–68.

[7]. Cass, T.A. 1992. Polynomial-time object recognition in the presence of clutter, occlusion, and uncertainty. In *Proc. European Conf.on Computer Vision*, Santa MargheritaLigure, Italy, pp. 834–842.
[8]. Cass, T.A. 1994. Robust geometric matching for 3D object recognition.In. *Proc. 12th IAPR Int. Conf. on Pattern Recognition*,Jerusalem, Israel, vol. 1, pp. 477 –482.
[9]. Cass, T.A. 1998. Robust affine structure matching for 3D object recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(11):1265–1274.
[10]. DeMenthon, D. and Davis, L.S. 1993. Recognition and tracking of 3D objects by 1D search. In *Proc. DARPA Image UnderstandingWorkshop*, Washington, DC, pp. 653–659.
[11]. DeMenthon, D. and Davis, L.S. 1995. Model-based object posein 25 lines of code. *International Journal of Computer Vision*,15(1/2):123–141.
[12]. Ely, R.W., Digirolamo, J.A., and Lundgren, J.C. 1995. Model supportedpositioning. In *Proc. SPIE, Integrating PhotogrammetricTechniques with Scene Analysis and Machine Vision II*
[13]. Orlando,FL.Fiore, P.D. 2001. Efficient linear solution of exterior orientation.*IEEE Trans. on Pattern Analysis and Machine Intelligence*,23(2):140–148.
[14]. Fischler, M.A. and Bolles, R.C. 1981. Random sample consensus:A paradigm for model fitting with applications to image analysisand automated cartography. *Comm. Association for ComputingMachinery*, 24(6):381–395.
[15]. Geiger, D. and Yuille, A.L. 1991. A common framework for imagesegmentation. *International Journal of Computer Vision*,6(3):227–243.
[16]. Gold, S. and Rangarajan, A. 1996.Agraduated assignment algorithmfor graph matching. *IEEE Trans. on Pattern Analysis and MachineIntelligence*, 18(4):377–388.
[17]. Gold, S., Rangarajan, A., Lu, C.-P., Pappu, S., and Mjolsness, E.1998. New algorithms for 2D and 3D point matching: Pose estimationand correspondence. *Pattern Recognition*, 31(8):1019–1031.
[18]. Morokoff, W.J. and Caflisch, R.E., 1994. Quasi-random sequencesand their discrepancies. *SIAM Journal Scientific Computing*,15(6):1251–1279.
[19]. Murase, H. and Nayar, S.K. 1995. Visual learning and recognitionof 3-D objects from appearance. *Int. Journal of Computer Vision*,14(1):5–24.
[20]. Olson, C.F. 1997. Efficient pose clustering using a randomized algorithm.*Int. Journal of Computer Vision*, 23(2):131–147.
[21]. Procter, S. and Illingworth, J. 1997. ForeSight: Fast object recognitionusing geometric hashing with edge-triple features. In *Proc.Int. Conf. on Image Processing*, vol. 1, Santa Barbara, CA, pp.889–892.
[22]. Sinkhorn, R. 1964.Arelationship between arbitrary positive matricesand doubly stochastic matrices. *Annals Mathematical Statistics*,35(2):876–879.