Pose estimation of a tilted pellet using single view from robot mounted camera

Shraddha Chaudhary
Department of Electrical Engineering
Indian Institute of Technology Delhi
New Delhi, India

Chaudhary.shraddha18@gmail.com

Sumantra Dutta Roy
Department of Electrical Engineering
Indian Institute of Technology Delhi
New Delhi, India
sumantra@ee.iitd.ac.in

ABSTRACT

This paper proposes a novel approach for finding the orientation of the tilted pellet using a single view of the pellet. This is important for the automation of pick and place problem. A texture less isolated cylindrical object is used for the pick and place operation. In this approach the isolated pellets are segmented from the background, following which, its contour is estimated. Then based on the assumption that length and diameter of the pellet remains constant, the mathematical formulation of the pose estimation of the pellet is given, and using which the orientation of the tilted pellet is computed using just single view. This is then experimentally verified for the different orientations of the pellet and the results are within the acceptable levels of accuracy A brief overview of the error estimation has been included. This error in orientation of the pellet is due to the variance in the actual height of the pellet. Error Jacobian for the above inaccuracy was calculated, and was found to be within required limits.

Categories and Subject Descriptors

I.4.9 [Image Processing and Computer Vision]: Applications. I.2.9 [Artificial Intelligence]: Robotics - Commercial robots and applications. I.4.8 [Image Processing and Computer Vision]: Scene analysis- Object recognition. J.6 [Computer Applications]: Computer-aided engineering - Computer-aided manufacturing (CAM).

General Terms

Measurement, Performance, Experimentation

Keywords

Computer vision, Robotics, Pose, Monocular vision, Curve based matching

1. INTRODUCTION

In many manufacturing environments, work-pieces are supplied in plates. It is a common industrial problem to load machines with such pieces. Often, human labour is employed to load an automatic machine with work-pieces. Such jobs are monotonous

© 2015 Association for Computing Machinery. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

AIR'15, July 02 - 04, 2015, Goa, India ©2015 ACM. ISBN 978-1-4503-3356-6/15/07...\$15.00 DOI: http://dx.doi.org/10.1145/2783449.2783502 Ratan Sadanand
Department of Mechanical Engineering
Indian Institute of Technology Delhi
New Delhi, India
ratan.sadan@gmail.com

Santanu Chaudhury
Department of Electrical Engineering
Indian Institute of Technology Delhi
New Delhi, India
schaudhury@gmail.com

and do not enrich human life. While the cost of labour is increasing, human performance remains essentially constant. Furthermore, the environment in manufacturing areas is generally unhealthy. Also, when work-pieces are inserted into machines by hands, there is a possibility of the limbs being exposed to danger. This is where the concept of process automation using computer vision plays a vital role. Use of computer vision in such cases is an obvious solution.

Picking the work piece using a vision system would require the accurate estimation of its position and orientation (pose) using the data acquired by the vision sensor. Few of the known approaches for pose estimation include use of stereo vision [1] and range sensors [2]. Using the stereo set-up we can find the correspondence between the two images and thereby reconstruct the scene for further reference. However, generally objects used in industries are simple and featureless, which poses a problem to the researchers working in this area. In such a scenario, finding point correspondence for such objects becomes very difficult and even impossible in certain cases. In this paper we are proposing an approach that uses mono-vision model based recognition to solve 3D Pose estimation problem. The orientation of an isolated pellet was established from a single view of the camera, based on the known dimensions of the pellet and the workspace of the robot. But, the difficulty creeps in when we have to solve the 3D pose estimation problem using a single view, as 3D information of the object cannot be extracted directly from a single image. Since, we are taking just single view there is no alternative for establishing correspondence between two views and determine its change w.r.t to its horizontal plane Hence, in the image of the pellet looks like it is just oriented w.r.t to its vertical plane, therefore it becomes a non-trivial problem to solve for.

1.1 Related Work

There has been a lot of research in the area of pose estimation of a 3D object using single view point [8, 9], in which using a single view important features are captured. This is advantageous in the aspect that a single view mitigates the problem of spatial correspondence between the views from the different cameras or multiple viewpoints. Similarly, [5, 6] have used monocular vision but in their work also focused on the look-up-table (LUT) scheme, through which the data was learned and hence the pose of the object was deciphered. But this scheme is very time consuming and not memory-efficient as lot of data has to be stored in the LUT for single pose of the pellet. This paper extends the work

presented in [7]. Although the methodology proposed in [7] was able to reduce the time and memory consumed, it didn't cater to the problems related to the 3D configuration of the cylindrical pellet. In this paper we are proposing the mathematical formulation for 3D pellets and its experimental results which will help make the problem of 3D pellet orientation solvable, using a single view.

This paper is organised as follows. Section 2 discusses the problem in estimating the pose for the lying down pellet and tilted pellet, Section 3 discusses the mathematical formulation and Section 4 deals with the experimental results and discussions. Finally the conclusions are given in Section 5.

2. POSE ESTIMATION

Pose of an object consists of its position (X, Y, and Z coordinates) and its orientation. An object's in particular a pellet orientation could be w.r.t to vertical plane (2D configuration) and also with respect to its horizontal plane (3D Configuration) as shown in Figure 1. Left pellet is in 2D configuration and the right pellet is in 3D configuration. In this case Orientation with respect to Z is not being considered because pellet is symmetrical about Z plane.



Figure 1. 2D and 3D Orientation of the pellet.

2.1 Position estimation

In this section, position estimation for the pellet is discussed. Position is the location where object of interest has been placed in the workspace In this case position estimation refers to the X, Y and Z location of the pellet in the de4fined coordinate system. In order to determine the Position of the texture-less or featureless object ,it is a non-trivial problem because it becomes difficult to establish a correspondence between the object and image to find its position, hence to determine its X, Y and Z is in itself a problem of concern. In order to solve this problem we have used a contour based approach [7]. But the fact that all the pellets are lying on the table in an isolated manner can be exploited to get Zposition information by calibrating the KUKA Robot [11] end effector to the table, and thereby deriving the other two information - X-position and Y-position of the pellet- by calibrating the camera mounted on the robot to the workspace lying on the surface of the table using the Zhang's calibration [6]. In the proposed methodology, Z = 0 was considered

Where,

u, v are image co-ordinates.

 K_{3X3} represents internal camera matrix.

 R_{ij} and t_q are the external camera calibration matrix parameters. c is a constant.

x, y represents the position of the pellet in world co-ordinate system.

2.2 Segmentation

We have used S. K. Nayar's [10] reflectance ratio based segmentation using the formula as:

$$p = \frac{(I_1 - I_2)}{(I_1 + I_2)} \tag{2}$$

Where.

p = reflectance ratio of pixel 1

 I_1 = intensity value at pixel 1

 I_2 = intensity value at pixel 2

with a preprocessing step of histogram equalization on the regionof-interest which has in turn helped us in improving the contrast in an image and there enhancing the edges. Therefore enhancement of edges gives better localization on the pellets and thereby better results as compared with its previous work in [7]. Some results are shown below.



Figure 2. Original image and manual histogram equalized image

2.3 Orientation estimation

This section of the paper deals with the 2D and 3D orientation of the cylindrical pellet. Firstly, we have determined the position (X,Y), and Z coordinates) in Section 2.1 and together with the orientation of the pellet give us the complete pose of the pellet. The object's orientation is w.r.t to its vertical plane it is considered as the 2D pose (Position + 2D orientation) of an object whereas, when the object's orientation is w.r.t to its horizontal plane along with some 2D orientation its pose is considered as the 3D Pose (Position +2D orientation+3D orientation).Now, There are different steps involved, in order to find the orientation of the pellet, which are as shown in brief in the flowchart in the Figure 3.

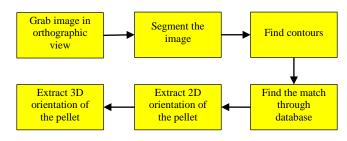


Figure 3. Flowchart for the 2D and 3D orientation estimation

2.3.1 2D Pose: orientation computation

For estimation of 2D orientation of the pellet as given in [7], initially, a database was created for the appearance of pellet in 2D when viewed from the camera. This process is performed offline and therefore it does not result in an overhead in processing time.

3D objects are stored in the database with the camera transformation matrix which helps us in converting the world coordinate of an object to the 2D co-ordinate by taking both translation and rotation into account [1]. After creating the database, next step is to segment the image as explained in the Section 2.2. Following this contour matching is done, which helps us find the contour or curve if the isolated pellet and thereby the curve id matched with the already existing database for orientation and hence we get the 2D orientation of the pellet [7].

2.3.2 3D Pose: orientation Computation:

This section of the paper introduces the novelty in the paper that describes about the orientation of the pellet in the 3D configuration. The orientation of an object in 3D is generally described using three angles - which maybe Euler angles, Roll-Pitch-Yaw conventions, etc. However, the cylindrical pellet under consideration is symmetrical about its axis. Hence only two angles are required to describe the orientation of the pellet in 3D. This is explained in detail in Section 3. The pellet is not completely lying on the table instead it is tilted. This case was not dealt in the analysis presented in [7]. Since the object we are working on is the texture less object, this becomes a problem to solve. In order to address this problem, we have used the orthographic projections. A correlation was established between the inclination of the pellets with the horizontal plane, and the projected area of the pellets from the top view. A parametric relation was derived between the inclination angle, diameter of the pellet, height of the pellet and the projected area of the pellet. Algorithm for the 3D orientation detection is explained in the Figure 4.

Algorithm #1: Brief description for calculating the orientation of tilted pellet.

- 1. Set camera at a constant height in orthogonal view.
- Measure the Height and Diameter (width) of the cylindrical object in image space.
- 3. Get the image of the tilted pellet from the same height.
- 4. Segment the Images.
- Now, area of the segmented image measured in image space, orientation of the tilted pellet can be determined as a function of height, width and area.

Figure 4. Algorithm for estimating the 3D orientation of the cylindrical object.

3. MATHEMATICAL FORMULATION AND VALIDATION

The height and diameter of the cylindrical pellet in the image space are h and d, respectively. It is inclined at an angle θ to the horizontal plane, and at angle ϕ with the vertical plane. The front view and top view of the pellet with the relevant dimensions are shown in Figure 5. Assuming that the image captured by the camera is free of the perspective distortions, the projection of the cylindrical pellet is shown in Figure 5.The calculations involved for obtaining the angles θ (inclination of the pellet from the horizontal plane) and ϕ (inclination of the pellet from the vertical plane) from a single view of the pellet is given below in subsection 3.1.

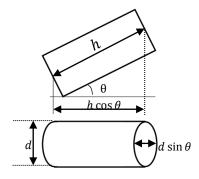


Figure 5. Orthographic projection of the tilted pellet.

3.1 Calculation of the 3D orientation of pellet

It is assumed that the pellets are similar in dimensions - i.e., the height h and diameter d of all the pellets are equal. The effect of minor variations in h and d, and the resultant errors are modelled in Section 3.2. For a given set of pellets, the values of h and d in image space are assumed equal. These are determined from the images of the pellet in two known orientations — when the pellet is lying down ($\theta = 0^o$) and when the pellet is standing vertical ($\theta = 90^o$), as shown in Figure 5. The projected area of the pellet can be written in terms of the orientation θ , diameter d, and height h as:

$$A(\theta) = hd\cos\theta + \frac{\pi}{4}d^2\sin\theta \tag{3}$$

The projected length of the pellet from end to end can be written as follows:

$$L(\theta) = h\cos\theta + d\sin\theta \tag{4}$$

Combining, Equations (4) and (5):

$$\begin{bmatrix} hd & \frac{\pi}{4}d^2 \\ h & d \end{bmatrix} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} A \\ L \end{bmatrix}$$
 (5)

$$\begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} = \begin{bmatrix} hd & \frac{\pi}{4}d^2 \\ h & d \end{bmatrix}^{-1} \begin{bmatrix} A \\ L \end{bmatrix}$$
 (6)

The RHS of the equation (6) is known from the measurements. The projected area A and the rojected length L can be measure from the thresholded image of the pellet. These measurements are done based on the properties of the thresholded region. For a cylindrical pellet with h/d ration greater than 1, the orientation of the major axis of the thresholded region of pellet will give the value of ϕ . For h/d ratio less than 1, the value of minor axis should be used to calculate the value of ϕ . The major or minor axes of the thresholded region has been calculated using MATLAB script. The method 'regionProps.Orientation' in MATLAB can be used to obtain the properties of the thresholded region. The thresholded image can then be rotated by an angle $-\phi$, so that the major axis is now aligned with the x-axis of the image. Following this, the projected length can be estimated using the length of the bounding box for the thresholded region of the rotated image. By measuring the area of the thresholded region, the projected area A can also be estimated. The values for h and d in the image space are determined. Once these parameters are known, the numerical values of the LHS in Equation (6) can be used to find θ . The value of θ can be determined using Atan2 function of MATLAB.

3.2 Error Estimation in θ

In this section, we have modelled the error in θ , because we have pellets of variable size. From Equation (6) it is visible that θ is a function of height h and diameter d, which implies that any variation in them, would lead to change in θ .

We need to calculate the distribution of the error in theta with variation in its dependent parameter, i.e., $p_e(\theta|h,d,l,A)$ can be modelled using the Jacobian as:

$$d\theta = \begin{bmatrix} \frac{\partial f}{\partial h} & \frac{\partial f}{\partial d} & \frac{\partial f}{\partial L} & \frac{\partial f}{\partial A} \end{bmatrix} \begin{bmatrix} dh \\ dd \\ dL \\ dA \end{bmatrix}$$
(7)

Where, f has been derived using equation (6) as follows:

$$\theta = A \tan 2(x, y) \tag{8}$$

$$\theta = 2 \tan^{-1} \frac{y}{\sqrt{x^2 + y^2} + x} \tag{9}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} = \begin{bmatrix} \frac{\pi d^2 h}{4} - Ad \\ hd^2 \left(\frac{\pi}{4} - 1\right) \\ \frac{hA - hdl}{hd^2 \left(\frac{\pi}{4} - 1\right)} \end{bmatrix}^{-1} \tag{10}$$

$$f = 2 \tan^{-1} \frac{y}{\sqrt{x^2 + y^2 + x}} \tag{11}$$

$$f = 2 \tan^{-1} \frac{hA - h \, dL}{\sqrt{(hA - h \, dL)^2 + (\frac{\pi}{4} d^2 L - Ad)^2 + (\frac{\pi}{4} d^2 L - Ad)^2}}$$
(12)

Error in estimation of theta due to any of its dependent parameter can be obtained using equation (7). The error estimation has been done using the Jacobian. Since the Jacobian has been determined using partial derivatives, the expressions are long. These have not been included due to limitations on space. For the sake of completion, calculations of $\frac{\partial f}{\partial h}$ is given in Appendix A. Using the Jacobian the error in estimated inclination of the pellet w.r.t variations in parameters h, d, A and L is studied and the results for same are presented in the Section 4.3.

3.3 Implementation in MATLAB

The calculation mentioned in Section 3.1 is implemented in MATLAB, using the steps mention in Algorithm#2.

Algorithm#2 determines the θ using MATLAB and assuming that area A and length L of the pellet remains constant, the 3D orientation of the pellet is determined based on the mathematical formulation given in subsection 3.1. We have also calculated the ϕ in order to validate our results with respect to the results that we have obtained from the steps mentioned in the subsection 2.3.1. It has been seen that value of ϕ that was obtained here are in consistent with the values obtained in subsection 2.3.1.

#Algorithm 2: MATLAB implementation algorithm to compute 3D orientation:

- 1. Acquire the image from the robot mounted camera
- 2. Threshold the image to detect the contour.
- 3. Detect the major axis of the detected contour. Calculate the inclination with the vertical plane (phi)
- 4. Rotate the image to align the major axis of the detected contour with the horizontal axis
- 5. Determine the orientation with the horizontal plane (theta) using the height and diameter values in the image space.

4. RESULTS AND DISCUSSION

The measurement of h and d in the image space in done initially using the images of pellet for the inclinations $\theta = 0^{\circ}$ (Lying down pellet, that is lying on the horizontal plane) and $\theta = 90^{\circ}$ (This is the pellet in the standing configuration, for which we get a circle as it orthographic projection) as shown in Figure 6.

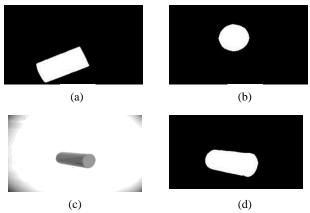


Figure 6. (a) Segmented Image of lying down pellet,
(b) Segmented image of standing pellet
(c) Tilted pellet
(d) Segmented image of tilted pellet.

Following this, the image of the pellet in an oblique orientation is acquired, and segmented to obtain a binary image.

4.1 Simulated Experiments

The parameters A and L as described in Equation (6) are measured and the inclination θ is estimated. In this section we are calculated the area and height based on the simulated image from CAD. In the Figure 7, it can be seen that the h and d for the pellet are changed continuously when there is a change in the inclination from the horizontal plane and keeping the 2D orientation \emptyset constant. Since our formulation emphasize on the area and length of the pellet as the main parameter so that variation has been captured in the Figure 7.

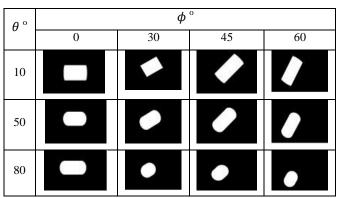


Figure 7. Variation in the tilted pellet height and Diameter with different heta and particular heta

The results for the measured value of θ compared to the actual value is shown in the Figure 8. In the Figure 8, x-axis represents the inclination w.r.t horizontal plane and y-axis represents the error in theta (θ).It can be concluded from the above graph that the 3D orientation is determined almost error free from 10 degrees to 90 degrees, although errors are little high although the average errors as depicted in the Table 1 below are within the limits of error and picking ability.

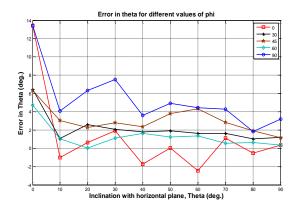


Figure 8. Graph depicting the error in the 3D orientation w.r.t to its actual orientation.

Table 1. Average error in the estimated orientation of the pellet:

| S. No. | Measured Values | Average Error |
|--------|------------------|---------------|
| 1 | Phi (Ø) | 0.0175 |
| 2 | Theta (θ) | 2.8331 |

In the Figure 8, it depicts that error profiles gives us a normal deviation and ran for 50 odd trials. With its mean concentrated near 2° of error and its variance is spread across with the standard deviation of 3.03°, which is a quite fine-tuned result as during the pickup of the tilted pellet firstly the Angle A of the KUKA Robot is oriented and thereafter the angle B has to be oriented w.r.t to the angle A and obliqueness in the pellet. Hence with the combination of these two we get the best pick up, which has not been discussed in the [7].

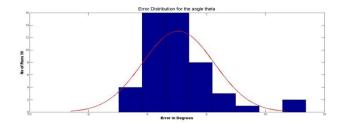


Figure 9: Error Profile for the Angle Theta (θ)

4.2 Actual experiments and repeatability measurements

For better and robust system, we have considered three cases for Phi (\emptyset), less than 90 degrees, equal to 90 degrees and greater than 90 degrees for different values of theta(θ). Since the discussed method considers isolated pellets at different orientation, the robot's repeatability has to be measured. For repeatability measurements, we placed the isolated pellet at different angles and orientations with respect to the workplace. In this experiment, the workplace was 200mmx 150mm. The camera was mounted on the end effector and the end effector's axis was orthogonal to the workplace. The pellet was placed at different known angles. The image was segmented and the angle (theta and phi) were obtained using the above mathematical formulation.



Figure 10. Experimental setup for actual pellet

In order to validate the calculated result, the end effector of the robot was aligned with the obtained orientation. Above mentioned procedure was repeated for 30 trials, 10 for each case of phi. The results are tabulated for different cases in the Table 2.

Table 2. Repeatability measurement for the Tilted pellet Trials

| S.No | Cases for Phi (Ø) | Success rate for 30 Trials |
|------|-------------------|----------------------------|
| | | in percentage |
| 1. | Ø < 90 | 90% |
| 2. | Ø = 90 | 85% |
| 3. | Ø > 90 | 95% |

Repeated trials prove that the proposed Mathematical formulation for the computation of the tilted pellet orientation would enable successful pick up of the pellet for any theta (θ) . To add to it, since the shape of the pellet is well known, approach formulated is well suited

4.3 Error variation in Theta (θ)

One pellet was calibrated with its 'h' and 'd' values (taken as reference pellet) and its theta was calculated using the algorithm#2.Pellet with variation in height was taken and

similarly theta was estimated. Now using the error estimation equation (7) for $\Delta\theta$ variation was calculated in theta depending upon the change in 'h'. Graph in Figure 11, depicts the variation. Above mentioned procedure was repeated for 8-9 pellets of different heights.

Inference that can be drawn from the Figure 11 is that with the variation in the height or length of the pellet, there is a variation but, important observation that is being revealed in this Figure 10 is that, even for larger variation in height maximum variation in theta is 5° only.

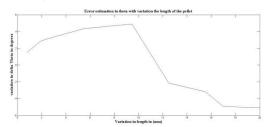


Figure 11. Error estimate in theta with different pellet lengths

5. CONCLUSIONS AND FUTURE WORK

The algorithm that has been formulated is independent of the size of the pellet. It works well with the different sizes of the pellet as shown in the Figure 7.Combination of the Angle A (roll) and Angle B (pitch) of the KUKA robot helps us in giving a better and robust pick of the lying down as well as the Tilted pellet.

The method will work successfully when the h/d ratio is greater than 1. If the h/d ratio is smaller than 1, then the minor axis must be used to determine the inclination with the vertical plane. Since the h/d ratio is known prior, based on the value, the major or minor axis can be used to detect the orientation with the vertical plane (phi) are required. For a small h/d ratio and when theta is large, it is the minor axis that must be used instead of the major axis.

The average error obtained as in the case of phi (\emptyset) and theta (θ) calculations were 0.0175° and 2.8331° respectively for data collected over 50 trials. The success rate with respect to repeatability has been high in range of 85% to 95% for 30 trials conducted on different range of phi (\emptyset) .

We are trying to mathematically formulate for the Cuboidal and triangular prisms so that the pickup can be generalized for all kind of shapes.

6. ACKNOWLEDGMENTS

The financial support to the first two authors from BRNS/BARC Mumbai under the "Programme for Autonomous Robotics Lab" at IIT Delhi is sincerely acknowledged.

Appendix A:For example:Partial derivative of 'f' with (13) respect to 'h can be computed as follows':

$$\frac{\partial f}{\partial h} = \frac{\frac{A - L * d}{\#3} - \frac{\#2 * (A - L * d)}{\#3}}{\frac{\#2}{(\#3^2)} + 1}$$

Where #1,#2,#3 are:

$$#1 = \sqrt{#2 + \frac{L^2 * pi^2 * d^4}{16}}$$

$$#2 = (A * h - L * d * h)^2$$

(14)

$$#3 = #1 - A * d + \frac{pi*L*d^2}{4}$$

7. REFERENCES

- [1] Hartley, R. and Zisserman, A., 2001. *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge.
- [2] Simon, D. A., Hebert, M., and Kanade T., 1993. Real-time 3-D Pose Estimation Using a High-Speed Range Sensor, *Robotics Institute, Paper 460*.
- [3] Johansson, B. and Moe, A., 2005. Patch-duplets for object recognition and pose estimation, *The 2nd Canadian Conference on Computer and Robot Vision*, 2005. pp. 9-16.
- [4] Wu, C., Fraundorfer, F., Frahm, J. M., and Pollefeys, M., 2008. 3D model search and pose estimation from single images using VIP features, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* Workshops, pp.1-8.
- [5] Kragic, D., and Vincze, M., 2009. Vision for Robotics. Found. Trends Robot 1, pp. 1-78.
- [6] Zhang, Z., 2000. A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no.11, pp.1330-1334.
- [7] Tiwan, P., Boby, R. A., Roy, S. D., Chaudhury, S., and Saha, S. K., 2013. Cylindrical Pellet Pose Estimation in Clutter using a Single Robot Mounted Camera. *Proceedings of Conference on Advances In Robotics*.
- [8] Breitenmoser, A., Kneip, L., and Siegwart, R., 2011. A monocular vision-based system for 6D relative robot localization, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.79-85.
- [9] Siddiqui, M. and Medioni, G., 2010. Human pose estimation from a single view point, real-time range sensor, IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp.13-18.
- [10] Nayar, S. K. and Bolle, R. M., 1993. Computing reflectance ratios from an image, *Pattern Recognition, IEEE International Conference on*, pp. 999-1007.
- [11] KUKA KR5 Arc, Accessed on December 2014 http://www.kukarobotics.com/en/products/industrial robots/low/kr5 arc/